



INSTITUT DE FRANCE
Académie des sciences

Comptes Rendus

Mathématique


Jose A. A. Andrade and Pushpa N. Rathie

Exact Posterior distribution of risk ratio in the Kumaraswamy–Binomial model

Volume 361 (2023), p. 1063-1069

Published online: 7 September 2023

<https://doi.org/10.5802/crmath.469>

 This article is licensed under the
CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.
<http://creativecommons.org/licenses/by/4.0/>



Les Comptes Rendus. Mathématique sont membres du
Centre Mersenne pour l'édition scientifique ouverte
www.centre-mersenne.org
e-ISSN : 1778-3569



Statistics / Statistiques

Exact Posterior distribution of risk ratio in the Kumaraswamy–Binomial model

Jose A. A. Andrade^{*,a} and Pushpa N. Rathie^b

^a Department of Statistics and Applied Mathematics, Federal University of Ceara, 60455-670, Fortaleza-Ce, Brazil

^b Department of Statistics, University of Brasilia, 70910-900, Brasilia-DF, Brazil
E-mails: ailton@ufc.br, pushpanrathie@yahoo.com

Abstract. In categorical data analysis, the 2×2 contingency tables are commonly used to assess the association between groups and responses, this is achieved by using some measures of association, such as the contingency coefficient, odds ratio, risk relative, etc. In a Bayesian approach, the risk ratio is modeled according to a Beta-Binomial model, which has exact posterior distribution, due to the conjugacy property of the model. In this work, we provide the exact posterior distribution of the relative risk for the non-conjugate Kumaraswamy–Binomial model. The results are based on special functions and we give exact expressions for the posterior density, moments, and cumulative distribution. An example illustrates the theory.

2020 Mathematics Subject Classification. 62C10.

Manuscript received 12 November 2022, accepted 29 January 2023.

1. Introduction

The 2×2 contingency tables is one of the most prominent forms of displaying statistical data, while the tables are very intuitive and easy to interpret, important information also can be extracted from them. In fact, verifying whether an attribute is more frequent in a group rather than another is a question which we often want to investigate. For instance, the association between binary variables is an essential tool for decision-making in many areas, such as classification, clinical trials, cohort studies, etc. In those studies, several measures can be used, to name a few: *the odds ratio, contingency coefficient, ϕ coefficient*, etc. One of the most widely used measure of association is the *risk ratio* (or *relative risk*). In Bayesian analysis, the groups in a 2×2 table are modeled as independent Binomial distributions with unknown probabilities of success, and the corresponding parameters are modeled as Beta distributions with some fixed hyperparameters, conveniently chosen to represent our prior knowledge about the odds of success. Thus, those measures usually are estimated through a Beta-Binomial model, which is a *conjugate* structure, in the sense that the posterior distribution belongs to the same family as the prior distribution.

* Corresponding author.

From this setting, the posterior distribution of the risk ratio can be obtained in exact form (Aitchison & Bacon-Shone [1]). The Beta-Binomial model is widely used to obtain posterior estimates for some measures of association, Smith et al. [11] obtain some hypothesis test involving the probabilities of success, Aitkin & Chadwick [2] provides estimates for several measures of association.

Another perspective may find the Beta distribution rather restrictive, since other prior distributions can better represent our knowledge about the parameters, however using other prior distributions may require approximate methods such as MCMC, numerical integration, etc., since there might not have a conjugate structure. This difficulty makes researchers to choose the facility of the Beta-Binomial model. Nevertheless, the well-established concept, that only conjugate prior distributions (or some specific reference priors) could yield exact posterior distributions was broken by Andrade & Rathie [4], where it was showed that a great variety of nonconjugate models can still produce exact posterior distributions, in the sense that the posterior distribution and its quantities can be explicitly written in a computable form. In this new field, Andrade & Rathie [4] and Andrade et al. [6] propose very wide classes of exact posterior nonconjugate models for a scale and location parameters, respectively. A generalization for location-scale structures is provided by Andrade & Rathie [5]. The main tool used in these works are the special functions, in particular the Fox's *H-function* [8], which generalizes most of the distributions in Statistics.

In the context of Binomial models, Andrade [3] proposes an alternative to the Beta-Binomial model, the *Kumaraswamy-Binomial* model, in which the posterior quantities are also obtained in exact form. The theory provides exact expressions for the posterior moments, cumulative posterior and predictive distributions. Due to some friendly mathematical properties (e.g. cdf, moments, median, etc), the Kumaraswamy distribution (Kumaraswamy [9]) has been widely used as an alternative to Beta distribution (see for instance, Andrade et al. [6] and Fletcher & Ponnambalam [7]). In this work, we use the results of Andrade [3], concerning the nonconjugate Kumaraswamy-Binomial model, in order to obtain the exact posterior distribution and its quantities of the risk ratio in 2×2 contingency tables.

The text is organized as the following. Section 2 provides some preliminary definitions and results useful to the theory. In Section 3 the main results are presented, in which the explicit posterior distribution and the posterior quantities are provided. In Section 4 an illustrative example given, in which the exact results are compared with those obtained by the MCMC method. In Section 5 some general comments consider possible generalizations of the theory.

2. Preliminaries

Andrade [3] gives the exact posterior quantities for the Kumaraswamy-Binomial model. Firstly, the results are expressed through *unnormalized moments*, that is, if a random variable $X \sim f(x)$ such that $f(x) = c \times h(x)$, where c is the normalizing constant and $h(x)$ is the kernel of the distribution, the unnormalized moment of order $r \in \mathbb{Z}$ is defined as $I(r) = \int_{\mathcal{X}} x^r h(x) dx$. Once we obtain an expression for $I(r)$ we can access most of the posterior quantities, such as the normalizing constant itself, the posterior moments, etc.

Consider the Kumaraswamy-Binomial model

$$\begin{cases} y_1, \dots, y_n | \theta \sim \text{Ber}(\theta) \text{ iid} \\ \theta \sim \text{Kum}(\alpha, \beta), \end{cases} \quad (1)$$

where $0 < \theta < 1$ is the parameter of interest, α, β are hyperparameters and $\text{Kum}(\alpha, \beta)$ stands for the Kumaraswamy distribution $p(\theta) = \alpha\beta\theta^{\alpha-1}(1-\theta^\alpha)^{\beta-1}$. Andrade [3] gives an exact expression for the unnormalized posterior moment of order r , given by

$$I_y(r) = \alpha\beta\Gamma(\beta)\Gamma(n-y+1) \sum_{h=0}^{\infty} \frac{(-1)^h \Gamma(r+y+\alpha(h+1))}{h! \Gamma(r+n+\alpha(h+1)+1) \Gamma(\beta-h)}, \quad (2)$$

where $\alpha > 1$ and $y = \sum_i y_i$.

An important result (Luke [10]) uses the fact that, in general, for $0 < b < 1$ and $c > 0$, using the binomial theorem, we have

$$\int_0^b z^{c-1}(1-z)^d dz = \sum_{h=0}^{\infty} \frac{(-d)_h}{h!} \int_0^b z^{c+r-1} dz = \sum_{h=0}^{\infty} \frac{(-d)_h}{h!} \frac{b^{c+h}}{c+h}, \tag{3}$$

where $(-d)_h$ is the Pochhammer notation, $(-d)_h = \frac{(-d)(-d-1)\dots(-d-h+1)}{h!} = \frac{\Gamma(h-d)}{\Gamma(-d)}$.

3. Posterior distribution of the risk ratio

We consider an experiment where n_1 and n_2 subjects are randomly allocated into the groups 1 and 2, say treatment 1 and 2, the responses to the treatments are marked as ‘‘Success/Failure’’ (1/0), thus a general 2×2 contingency table is of the form

Group	Response	
	1	0
Treat. 1	θ_1	$1 - \theta_1$
Treat. 2	θ_2	$1 - \theta_2$

The probabilities of success within the two groups are θ_1 and θ_2 ($0 < \theta_1, \theta_2 < 1$). Thus, we let $P(X_i = 1|\theta_1) = \theta_1$ ($i = 1, \dots, n_1$) and $P(Y_j = 1|\theta_2) = \theta_2$ ($j = 1, \dots, n_2$). Note that X_i and X_j are independent, as well as θ_1 and θ_2 , hence we consider the Binomial model

$$\begin{cases} x_1, \dots, x_{n_1} | \theta_1 \sim \text{Ber}(\theta_1) \text{ iid} \\ y_1, \dots, y_{n_2} | \theta_2 \sim \text{Ber}(\theta_2) \text{ iid} \\ \theta_j \sim \text{Kum}(\alpha_j, \beta_j), j = 1, 2. \end{cases} \tag{4}$$

where α_j, β_j are hyperparameters for the Kumaraswamy distribution $p_j(\theta_j) = \alpha_j \beta_j \theta^{\alpha_j-1} (1 - \theta^{\alpha_j})^{\beta_j-1}$ ($j = 1, 2$). The kernel of the posterior distribution is given by

$$\begin{aligned} p(\theta_1, \theta_2 | x, y) &\propto L(\theta_1, \theta_2) p_1(\theta_1) p_2(\theta_2) \\ &\propto \theta_1^x (1 - \theta_1)^{n_1-x} p_1(\theta_1) \times \theta_2^y (1 - \theta_2)^{n_2-y} p_2(\theta_2), \end{aligned} \tag{5}$$

where $x = \sum_i x_i$, $y = \sum_i y_i$ and $L(\theta_1, \theta_2)$ is the likelihood function. Clearly, the marginal posterior distributions will not be in the Kumaraswamy family, that is, the likelihood and prior distribution will not combine their kernels, thus in some regular Bayesian view, some approximate method would be required.

In our approach, we first obtain the unnormalized moment of order r and s , that is

$$\begin{aligned} I_{x,y}(r, s) &= \int_0^1 \int_0^1 \theta_1^r \theta_2^s p(\theta_1, \theta_2 | x, y) d\theta_1 d\theta_2 \\ &= \int_0^1 \theta_1^{x+r} (1 - \theta_1)^{n_1-x} p_1(\theta_1) d\theta_1 \times \int_0^1 \theta_2^{y+s} (1 - \theta_2)^{n_2-y} p_2(\theta_2) d\theta_2 =: I_x(r) \times I_y(s), \end{aligned}$$

where $I_x(r)$ and $I_y(s)$ are given by (2) and the computable representations (2), replacing conveniently α and β by α_j , and β_j ($j = 1, 2$), respectively. The posterior normalizing constant is obtained by evaluating the series up to a required precision, the normalizing constant is $I_{x,y}(0, 0) = I_x(0) \times I_y(0)$, hence the exact posterior distribution is obtained dividing (5) by $I_{x,y}(0, 0)$, which can be written in terms of the *binomial theorem*,

$$\begin{aligned} p(\theta_1, \theta_2 | x, y) &= \frac{\alpha_1 \beta_1 \alpha_2 \beta_2}{I_x(0) I_y(0)} \sum_{j=0}^{n_1-x} \sum_{\ell=0}^{n_2-y} \binom{n_1-x}{j} \binom{n_2-y}{\ell} (-1)^{j+\ell} \theta_1^{x+\alpha_1+j-1} \theta_2^{y+\alpha_2+\ell-1} \\ &\quad \times (1 - \theta_1^{\alpha_1})^{\beta_1-1} (1 - \theta_2^{\alpha_2})^{\beta_2-1} \end{aligned} \tag{6}$$

In order to make inferences about ρ , we need to obtain the posterior distribution of ρ , thus we consider the posterior distribution (6) to obtain the following exact posterior quantities.

Theorem 1 (Cumulative posterior distribution of the risk ratio). *Let $\rho = \theta_1/\theta_2$ be the risk ratio. Considering (6), the cumulative posterior distribution of ρ is given by*

(i) For $0 < \rho_0 < 1$:

$$P(\rho < \rho_0|x, y) = \frac{\beta_1\beta_2}{I_x(0)I_y(0)} \frac{\Gamma(\beta_2)}{\Gamma(1-\beta_1)} \sum_{j=0}^{n_1-x} \sum_{\ell=0}^{n_2-y} \binom{n_1-x}{j} \binom{n_2-y}{\ell} (-1)^{j+\ell} \rho_0^{x+j+\alpha_1} \times \sum_{h=0}^{\infty} \frac{\rho_0^{\alpha_1 h} \Gamma(1-\beta_1+h) \Gamma\left[\frac{y+\ell+x+j+\alpha_1+\alpha_2}{\alpha_2} + \frac{\alpha_1 h}{\alpha_2}\right]}{h! \left[\frac{x+j+\alpha_1}{\alpha_1} + h\right] \Gamma\left[\frac{y+\ell+x+j+\alpha_1+\alpha_2+\alpha_2\beta_2}{\alpha_2} + \frac{\alpha_1 h}{\alpha_2}\right]}. \quad (7)$$

(ii) For $\rho_0 > 1$:

$$P(\rho < \rho_0|x, y) = 1 - \frac{\alpha_2\beta_1\beta_2\Gamma(\beta_1)}{I_x(0)I_y(0)\Gamma(1-\beta_2)} \sum_{j=0}^{n_1-x} \sum_{\ell=0}^{n_2-y} \binom{n_1-x}{j} \binom{n_2-y}{\ell} (-1)^{j+\ell} \rho_0^{-(y+\ell+\alpha_2)} \times \sum_{h=0}^{\infty} \frac{\rho_0^{-\alpha_2 h} \Gamma(1-\beta_2+h) \Gamma\left[\frac{y+\ell+x+j+\alpha_1+\alpha_2+\alpha_1 h}{\alpha_1}\right]}{h!(y+\ell+\alpha_2+\alpha_2 h) \Gamma\left[\frac{y+\ell+x+j+\alpha_1+\alpha_2+\alpha_2 h+\beta_1\alpha_1}{\alpha_1}\right]}. \quad (8)$$

Proof. (i). For $0 < \rho_0 < 1$: Let $\rho = \theta_1/\theta_2$ and $\theta_2 = \theta_2$, using (6),

$$P(\rho < \rho_0|x, y) = P(\theta_1 < \rho_0\theta_2|x, y) = \int_0^1 \int_0^{\rho_0\theta_2} p(\theta_1, \theta_2|x, y) d\theta_1 d\theta_2 = \frac{\alpha_1\alpha_2\beta_1\beta_2}{I_x(0)I_y(0)} \sum_{j=0}^{n_1-x} \sum_{\ell=0}^{n_2-y} \binom{n_1-x}{j} \binom{n_2-y}{\ell} (-1)^{j+\ell} \times \int_0^1 \theta_2^{y+\alpha_2+\ell-1} (1-\theta_2^{\alpha_2})^{\beta_2-1} \int_0^{\rho_0\theta_2} \theta_1^{x+\alpha_1+j-1} (1-\theta_1^{\alpha_1})^{\beta_1-1} d\theta_1 d\theta_2. \quad (9)$$

Let $\theta = \theta_1^{\alpha_1}$, then

$$A(\theta_2) = \int_0^{\rho_0\theta_2} \theta_1^{x+\alpha_1+j-1} (1-\theta_1^{\alpha_1})^{\beta_1-1} d\theta_1 = \frac{1}{\alpha_1} \int_0^{(\rho_0\theta_2)^{\alpha_1}} \theta^{\frac{x+j}{\alpha_1}} (1-\theta)^{\beta_1-1} d\theta = \frac{1}{\alpha_1} \sum_{h=0}^{\infty} \frac{(1-\beta_1)_h (\rho_0\theta_2)^{\alpha_1(\frac{x+j}{\alpha_1}+h+1)}}{h! \left(\frac{x+j}{\alpha_1} + h + 1\right)},$$

where the latest identity is due to (3).

Thus, replacing $A(\theta_2)$ in (9), we have

$$P(\rho < \rho_0|x, y) = \frac{\alpha_1\alpha_2\beta_1\beta_2}{I_x(0)I_y(0)} \sum_{j=0}^{n_1-x} \sum_{\ell=0}^{n_2-y} \binom{n_1-x}{j} \binom{n_2-y}{\ell} (-1)^{j+\ell} \times \sum_{h=0}^{\infty} \frac{(1-\beta_1)_h}{h!} \frac{\rho_0^{x+j+\alpha_1+\alpha_1 h}}{x+j+\alpha_1+\alpha_1 h} \int_0^1 \theta_2^{y+\alpha_2+\ell+x+j+\alpha_1+\alpha_1 h-1} (1-\theta_2^{\alpha_2})^{\beta_2-1} d\theta_2$$

The result follows by evaluating the integral on θ_2 as before.

(ii). For $\rho_0 > 1$: Note that $P(\rho < \rho_0|x, y) = 1 - P(\rho > 1/\rho_1|x, y)$, where $\rho_1 = 1/\rho_0$, hence $0 < \rho_1 < 1$, then we consider

$$\begin{aligned}
 P(\rho > 1/\rho_1|x, y) &= P(\theta_2 < \rho_1\theta_1|x, y) = \int_0^1 \int_0^{\rho_1\theta_1} p(\theta_1, \theta_2|x, y) d\theta_2 d\theta_1 \\
 &= \frac{\alpha_1\alpha_2\beta_1\beta_2}{I_x(0)I_y(0)} \sum_{j=0}^{n_1-x} \sum_{\ell=0}^{n_2-y} \binom{n_1-x}{j} \binom{n_2-y}{\ell} (-1)^{j+\ell} \\
 &\quad \times \int_0^1 \theta_1^{x+\alpha_1+j-1} (1-\theta_1^{\alpha_1})^{\beta_1-1} \int_0^{\rho_1\theta_1} \theta_2^{y+\alpha_2+\ell-1} (1-\theta_2^{\alpha_2})^{\beta_2-1} d\theta_2 d\theta_1. \quad (10)
 \end{aligned}$$

The result follows by the same strategy as before, that is, replacing $\rho_1 = 1/\rho_0$ and using $P(\rho < \rho_0|x, y) = 1 - P(\rho > \rho_0|x, y)$. \square

Corollary 2 (Posterior density of the risk ratio). *Considering (1), the posterior distribution of the risk ratio is given by*

(i) For $0 < \rho_0 < 1$:

$$\begin{aligned}
 p(\rho|x, y) &= \frac{\alpha_1\beta_1\beta_2}{I_x(0)I_y(0)} \sum_{j=0}^{n_1-x} \sum_{\ell=0}^{n_2-y} \binom{n_1-x}{j} \binom{n_2-y}{\ell} (-1)^{j+\ell} \frac{\Gamma(\beta_2)}{\Gamma(1-\beta_1)} \\
 &\quad \times \sum_{h=0}^{\infty} \frac{\rho^{x+j+\alpha_1+\alpha_1h-1} \Gamma(1-\beta_1+h) \Gamma\left[\frac{y+\ell+x+j+\alpha_1+\alpha_2}{\alpha_2} + \frac{\alpha_1h}{\alpha_2}\right]}{h! \Gamma\left[\frac{y+\ell+x+j+\alpha_1+\alpha_2+\alpha_2\beta_2}{\alpha_2} + \frac{\alpha_1h}{\alpha_2}\right]}. \quad (11)
 \end{aligned}$$

(ii) For $\rho_0 > 1$:

$$\begin{aligned}
 P(\rho|x, y) &= \frac{\alpha_2\beta_1\beta_2\Gamma(\beta_1)}{I_x(0)I_y(0)\Gamma(1-\beta_2)} \sum_{j=0}^{n_1-x} \sum_{\ell=0}^{n_2-y} \binom{n_1-x}{j} \binom{n_2-y}{\ell} (-1)^{j+\ell} \\
 &\quad \times \sum_{h=0}^{\infty} \frac{\rho^{-(y+\ell+\alpha_2+\alpha_2h+1)} \Gamma(1-\beta_2+h) \Gamma\left[\frac{y+\ell+x+j+\alpha_2+\alpha_1+\alpha_2h}{\alpha_1}\right]}{h! \Gamma\left[\frac{y+\ell+x+j+\alpha_2+\alpha_2h+\alpha_1+\beta_1\alpha_1}{\alpha_1}\right]}. \quad (12)
 \end{aligned}$$

Proof. Straightforward by differentiating the (7) and (8) with respect to ρ_0 . \square

Corollary 3 (Posterior moments of the risk ratio). *Considering Corollary (2), the posterior moment of order $r \in \mathbb{Z}$ of ρ is given by*

$$\mathbb{E}[\rho^r|x, y] = \sum_{j=0}^{n_1-x} \sum_{\ell=0}^{n_2-y} \sum_{h=0}^{\infty} \left[\frac{A_{j\ell}(h)}{x+j+\alpha_1h+r+\alpha_1} + \frac{B_{j\ell}(h)}{y+\ell+\alpha_2r-r+\alpha_2} \right], \quad (13)$$

where

$$A_{j\ell} = \frac{\alpha_1\beta_1\beta_2\Gamma(\beta_2)}{I_x(0)I_y(0)\Gamma(1-\beta_1)} \binom{n_1-x}{j} \binom{n_2-y}{\ell} (-1)^{j+\ell} \frac{\Gamma(1-\beta_1+h) \Gamma\left[\frac{y+\ell+x+j+\alpha_1+\alpha_2}{\alpha_2} + \frac{\alpha_1h}{\alpha_2}\right]}{h! \Gamma\left[\frac{y+\ell+x+j+\alpha_1+\alpha_2+\alpha_2\beta_2}{\alpha_2} + \frac{\alpha_1h}{\alpha_2}\right]} \quad (14)$$

and

$$B_{j\ell}(h) = \frac{\alpha_2\beta_1\beta_2\Gamma(\beta_1)}{I_x(0)I_y(0)\Gamma(1-\beta_2)} \binom{n_1-x}{j} \binom{n_2-y}{\ell} (-1)^{j+\ell} \frac{\Gamma(1-\beta_2+h) \Gamma\left[\frac{y+\ell+x+j+\alpha_2+\alpha_1+\alpha_2h}{\alpha_1}\right]}{h! \Gamma\left[\frac{y+\ell+x+j+\alpha_2+\alpha_2h+\alpha_1+\beta_1\alpha_1}{\alpha_1}\right]}. \quad (15)$$

Proof. Straightforward by integrating

$$\begin{aligned}
 \mathbb{E}[\rho^r|x, y] &= \int_0^1 \sum_{j=0}^{n_1-x} \sum_{\ell=0}^{n_2-y} \sum_{h=0}^{\infty} A_{j\ell}(h) \rho^{r+x+j+\alpha_1h-1} d\rho \\
 &\quad + \int_1^{\infty} \sum_{j=0}^{n_1-x} \sum_{\ell=0}^{n_2-y} \sum_{h=0}^{\infty} B_{j\ell}(h) \rho^{r-(y+\ell+\alpha_2+\alpha_2h+1)} d\rho
 \end{aligned}$$

□

4. Example

In this section we illustrate the theory through an illustrative example, consider the following 2×2 contingency table, considering either the data or the probabilities, that is

Group	Response		\Rightarrow	Group	Response	
	Yes	No			Yes	No
Treat. 1	2	18		Treat. 1	θ_1	$1 - \theta_1$
Treat. 2	3	8	Treat. 2	θ_2	$1 - \theta_2$	

Here $\theta_1 = P(\text{Response} = \text{Yes} | \text{Treat.1})$ and $\theta_2 = P(\text{Response} = \text{Yes} | \text{Treat.2})$ are the parameters that will be used to estimate the risk ratio. We apply Model (4), noting that $n_1 = 20, x = 2, n_2 = 11$ and $y = 3$.

It follows that the unnormalized marginal posterior moment of order r of θ_i ($i = 1, 2$) is given by

$$I_{D_i}(r) = \alpha\beta\Gamma(\beta)\Gamma(n_i - D_i + 1)H_2^1 \frac{1}{2} \left[1 \left| \begin{matrix} (1 - (r + D_i + \alpha), \alpha), (\beta, 1) \\ (0, 1), (-r + n_i + \alpha), \alpha \end{matrix} \right. \right], \tag{16}$$

and the corresponding computable expression is given by

$$I_{D_i}(r) = \alpha\beta\Gamma(\beta)\Gamma(n - D_i + 1) \sum_{h=0}^{\infty} \frac{(-1)^h \Gamma(r + D_i + \alpha(h + 1))}{h! \Gamma(r + n + \alpha(h + 1) + 1) \Gamma(\beta - h)}, \tag{17}$$

where $D_1 = x, D_2 = y$ and $\alpha > 1$ and $\beta > 0$ are the hyperparameters of the Kumaraswamy distribution. We define the same prior distribution $\theta_i \sim \text{Kum}(\alpha = 1.1, \beta = 1.1)$ ($i = 1, 2$), which are flat (nearly uniform) prior distributions.

As the results, we provide comparisons of the posterior quantities obtained from MCMC algorithm (OpenBugs package with 10,000 cycles and 8,001 burned in) with the exact form equivalents. In Table 1, one can notice that the posterior quantities of the posterior expectation, variance, cumulative distribution are similar in the two methods. Notice that the cumulative posterior distribution can be used to obtain credible intervals. According to (2), the posterior density of ρ assumes different expression for $0 < \rho \leq 1$ and $\rho > 1$, thus we separate the MCMC simulated values according to these intervals. Note that the comparison between exact and approximated densities in Figure 1 shows some similarities, however our result is exact.

Table 1. Comparison of the exact posterior estimates with their equivalent obtained by MCMC (10,000, discarding 8,000)

Posterior Quantity	Estimates	
	MCMC	Exact
$E[\rho x, y]$	0.5526	0.4881
$\text{Var}[\rho x, y]$	0.1526	0.0739
$P(\rho \leq 0.8 x, y)$	0.8120	0.8167

5. Concluding remarks

In categorical data analysis there are many measures of association to handle 2×2 tables, the most prominent of them are the odds ratio and risk ratio. In Bayesian approaches, the risk ratio is commonly modeled through Beta-Binomial model, which is rather straightforward to apply, due

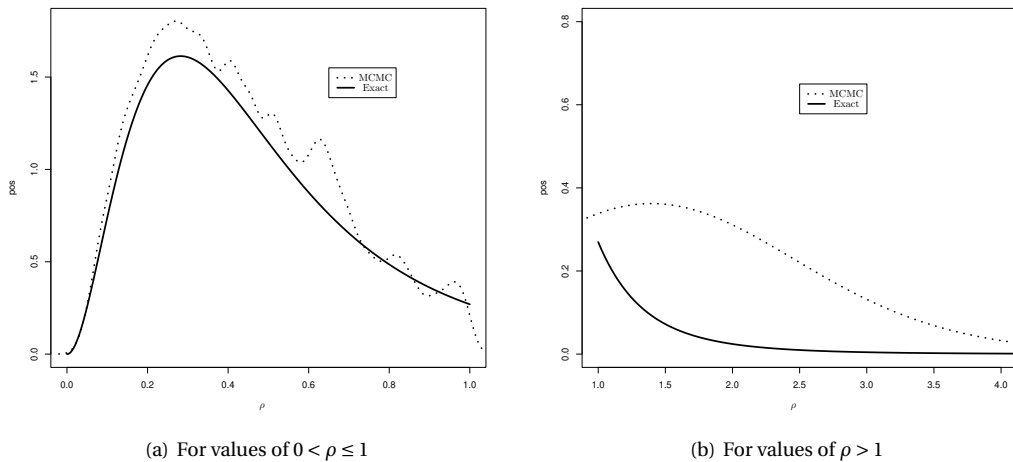


Figure 1. Exact and approximated posterior density of the risk ratio ρ .

to the conjugacy of the model. In this work, we establish results based on special functions that bring the same facility of the conjugate models, but with the Kumaraswamy distribution.

Future work will search for some exact models involving the odds ratio, which will increase the complexity of the model. Note that, once we have set the Binomial and prior distributions for the proportion of favorable answers, we will have to obtain a posterior distribution for the odds ratio, we will need to use special functions to assess the distribution of the odds ratio, this may lead to some very complex distribution.

References

- [1] J. Aitchison, J. Bacon-Shone, "Bayesian Risk Ratio Analysis", *Am. Stat.* **35** (1981), no. 4, p. 254-257.
- [2] M. Aitkin, T. Chadwick, "Bayesian analysis of 2×2 contingency tables from comparative trials", 2003.
- [3] J. A. A. Andrade, "Exact posterior computation for the binomial-Kumaraswamy model", *Adv. Comput. Math.* **46** (2020), no. 6, article no. 80 (13 pages).
- [4] J. A. A. Andrade, P. N. Rathie, "On exact posterior distributions using H-functions", *J. Comput. Appl. Math.* **290** (2015), p. 459-475.
- [5] ———, "Exact Posterior Computation in Non-Conjugate Gaussian Location-Scale Parameters Models", *Commun. Nonlinear Sci. Numer. Simul.* **53** (2017), p. 111-129.
- [6] J. A. A. Andrade, P. N. Rathie, R. B. A. Farias, "Exact Bayesian computation using H-functions", *Comput. Appl. Math.* **38** (2018), p. 2277-2293.
- [7] S. Fletcher, K. Ponnambalam, "Estimation of reservoir yield and storage distribution using moments analysis", *J. Hydrol.* **182** (1996), no. 1-4, p. 259-275.
- [8] C. Fox, "The G and H functions as symmetrical Fourier kernels", *Trans. Am. Math. Soc.* **98** (1961), p. 395-429.
- [9] P. Kumaraswamy, "A generalized probability density function for double-bounded random processes", *J. Hydrol.* **46** (1980), no. 1-2, p. 79-88.
- [10] Y. L. Luke, *The Special Functions and Their Approximations*, Academic Press Inc., 1979.
- [11] P. J. Smith, S. C. Choi, E. Gunel, "Bayesian Analysis of a 2×2 Contingency Table with Both Completely and Partially Cross-Classified Data", *J. Educ. Stat.* **10** (1985), no. 1, p. 31-43.