Revista Colombiana de Estadística
Special Issue: Colombian and Brazil Graduate Programs
December 2025, volume 48, issue 3, pp. 397–431
http://doi.org/10.15446/rce.v48n3.123477

A Bayesian Approach for Unit-Gamma Regression through Reparameterization in the Mean and Quantiles

Un enfoque bayesiano para la regresión gamma unitaria mediante reparametrización en la media y los cuantiles

Éric O. Rocha[2,a], Juvêncio S. Nobre[2,b], Caio L. N. Azevedo[1,c], Rafael B. A. Farias[2,d], Manoel Santos-Neto[2,e]

[1]Departamento de Estatística, Instituto de Matemática, Estatística e Computação Científica, Universidade de Campinas, Campinas, Brasil

[2]Departamento de Estatística e Matemática Aplicada, Centro de Ciências, Universidade Federal do Ceará, Fortaleza, Brasil

Abstract

In recent years, considerable attention has been devoted to statistical models designed for continuous doubly bounded dependent variables, defined on a finite interval (a, b) with known limits satisfying $-\infty < a < b < \infty$. A predominant focus, however, has been on the special case in which the response is restricted to the standard unit interval, namely $(0, 1)$. In this context, we introduce a new class of quantile regression models based on the unit-gamma distribution, along with a mean-based model. For inference, we adopt a Bayesian framework, employing Markov Chain Monte Carlo (MCMC) stochastic simulation algorithms. We conducted extensive simulations to verify the computational implementation and assess the properties of the Bayesian estimators. All analyses were carried out using the `nimble` package in R, which provides a flexible environment for building and fitting Bayesian hierarchical models. Our results suggest that the proposed Bayesian approach provides unbiased and consistent estimators for all model parameters. Additionally, we perform a detailed model fit assessment, comparing the proposed models with several established alternatives, and conduct an influence analysis to identify potential outliers or influential

[a]Master's degree. E-mail: eric.rocha@alu.ufc.br
[b]Ph.D. E-mail: juvencio@ufc.br
[c]Ph.D. E-mail: cnaber@unicamp.br
[d]Ph.D. E-mail: rafa@ufc.br
[e]Ph.D. E-mail: santosneto@dema.ufc.br

data points. Finally, we apply the proposed models to a real-world dataset, demonstrating their practical utility and systematically comparing their performance with that of existing models commonly used in the literature.

Keywords: Bayesian inference; Bayesian unit-gamma; Bounded data; Quantile regression.

Resumen

En los últimos años, se ha dedicado una atención considerable al desarrollo de modelos estadísticos diseñados para variables dependientes continuas doblemente acotadas, definidas en un intervalo finito (a, b) con límites conocidos que satisfacen $-\infty < a < b < \infty$. Sin embargo, el enfoque predominante ha sido el caso especial en el cual la variable respuesta está restringida al intervalo unitario estándar $(0, 1)$. En este contexto, presentamos una nueva clase de modelos de regresión cuantílica basados en la distribución gamma unitaria, junto con un modelo basado en la media. Para la inferencia, adoptamos un enfoque bayesiano, empleando algoritmos de simulación estocástica de tipo *Markov Chain Monte Carlo* (MCMC). Se realizaron extensas simulaciones para verificar la implementación computacional y evaluar las propiedades de los estimadores bayesianos. Todos los análisis se llevaron a cabo utilizando el paquete `nimble` en R, que ofrece un entorno flexible para la construcción y el ajuste de modelos jerárquicos bayesianos. Nuestros resultados indican que el enfoque bayesiano propuesto proporciona estimadores insesgados y consistentes para todos los parámetros del modelo. Además, se realizó una evaluación detallada del ajuste del modelo, comparando las propuestas con varias alternativas establecidas, y un análisis de influencia para identificar posibles valores atípicos o puntos de datos influyentes. Finalmente, aplicamos los modelos propuestos a un conjunto de datos reales, demostrando su utilidad práctica y comparando sistemáticamente su desempeño con el de modelos existentes comúnmente utilizados en la literatura.

Palabras clave: Datos acotados; Gamma unitaria bayesiana; Inferencia bayesiana; Regresión cuantílica.

1. Introduction

In the literature, regression models for variables with support in the interval $(0, 1)$ have been extensively studied under the frequentist paradigm. Early contributions include beta regression models developed by Ferrari & Cribari-Neto (2004), simplex regression models introduced by Song & Tan (2000), and the beta rectangular distribution proposed by Hahn (2008), which was later reparameterized in terms of the mean by Bayes et al. (2012). Other notable developments include the Kumaraswamy regression models formulated by Mitnik & Baek (2013) and the unit-gamma (UG) regression models designed by Mousa et al. (2016). Furthermore, extensions to handle longitudinal bounded data have been discussed by Venezuela et al. (2007), Ribeiro et al. (2021), and Freitas et al. (2023). More recently, Chesneau (2025) proposed a novel, simple, and intuitive two-parameter unit distribution derived from the gamma distribution, offering a complementary alternative to the existing UG distribution.

An alternative to mean-based modeling is quantile regression, introduced by Koenker & Bassett Jr (1978), which relates the conditional quantiles of the response variable to the covariates. This approach has inspired several bounded quantile regression models, such as the Kumaraswamy model (Hamedi-Shahraki et al., 2021), the unit-Weibull model (Mazucheli et al., 2020), the unit-Burr-XII model (Korkmaz & Chesneau, 2021), and the unit-Chen model (Korkmaz et al., 2022). For a comprehensive overview of bounded quantile regression models, see Mazucheli et al. (2022).

Within the Bayesian framework, various studies have focused on modeling the mean of bounded response variables. Branscum et al. (2007) investigated a hierarchical beta regression model, while Bayes et al. (2012) developed a robust version based on the beta distribution. The Bayesian approach was also applied to the simplex model by López (2013), and a beta regression model with mixed effects was proposed by Figueroa-Zúñiga et al. (2013) to incorporate hierarchical data structures. Later, Migliorati et al. (2018) introduced a flexible class of beta models derived from a mixture of two beta distributions.

More recent efforts have moved beyond mean-based modeling. Bayes et al. (2017) developed quantile regression models for bounded responses using the beta distribution, while Barrientos et al. (2017) proposed a fully nonparametric Bayesian framework to estimate the entire conditional distribution of the response. Mazucheli et al. (2019) and Guerra et al. (2021) contributed to the frequentist literature by introducing and exploring the unit-Weibull and unit extended Weibull distributions as flexible tools for modeling bounded data. Zhou et al. (2020) advanced a mode-based beta regression model, linking the mode of the response to covariates via a link function. In the Bayesian context, Oliveira et al. (2022) proposed quantile regression models tailored for bounded variables with heavy tails, employing the No-U-Turn sampler to improve sampling efficiency. More recently, Castro et al. (2024) developed the rectangular Kumaraswamy model, and Rocha et al. (2024) introduced the quantile unit-gamma model, expanding the repertoire of Bayesian tools for analyzing bounded data. Additionally, El-Awady & Ramadan (2025) introduced a novel one-parameter bounded distribution, termed the unit Rayleigh Half-Normal distribution, designed for modeling data on the unit interval $(0, 1)$, which frequently arises in fields such as economics, actuarial science, and medicine. Furthermore, Kim et al. (2025) developed the Scale-Location-Truncated beta (SLTB) regression model to address the limitations of standard beta regression in handling boundary values at 0 and 1. The SLTB model extends the beta distribution's domain to the $[0, 1]$ interval by employing scale–location transformation and truncation, which allows positive finite mass at the boundary values. This approach offers a flexible framework for modeling data with outcomes at 0 and 1.

In this paper, we propose a class of quantile UG regression models, in addition to considering the mean-based model under the Bayesian approach. Quantile regression allows for the estimation of the relationship between covariates and any quantile of the response variable, rather than just the mean, thus providing a more comprehensive modeling of the data. We discuss and present the entire estimation process using MCMC algorithms, model comparison, influence analysis, and application to real-life dataset.

The remainder of this paper is organized as follows. Section 2 introduces a reparameterization of the unit-gamma (UG) distribution, providing the foundation for its use in regression modeling. In Section 3, we present the UG regression model reparameterized in terms of the mean and quantiles, highlighting the advantages of this formulation, particularly in settings where the interest lies in specific distributional features of the response variable. Section 4 describes the implementation details of the proposed model, including the specification of prior distributions and the adopted Bayesian inference strategy. Section 5 discusses model fit assessment and comparison criteria, allowing for an evaluation of the relative performance across different model specifications. Section 6 presents simulation studies conducted to assess parameter recovery of the model under various scenarios. Section 7 illustrates the application of the proposed model to a real dataset. Finally, Section 8 presents concluding remarks and outlines potential directions for future research.

2. Reparametrization of the UG Distribution

A parameterization of the UG distribution in terms of the mean is presented in Mousa et al. (2016). Under this parameterization, the probability density function (pdf) of the UG with mean parameter $\mu \in (0,1)$ and precision $\phi \in \mathbb{R}_+^*$ is given by

$$f(y|\mu,\phi) = \left[\frac{\mu^{1/\phi}}{1-\mu^{1/\phi}}\right]^{\phi} \frac{1}{\Gamma(\phi)}\, y^{\frac{\mu^{1/\phi}}{1-\mu^{1/\phi}}-1}\, [-\log(y)]^{\phi-1}\, \mathbb{1}_{(0,1)}(y), \tag{1}$$

thus, we say that $Y|\mu,\phi \sim \mathrm{UG}_{\mu}(\mu,\phi)$, where $\mathbb{1}_{\mathcal{B}}(y)$ denotes the indicator function of the set $\mathcal{B}$. Furthermore, the cumulative distribution function (cdf) is given by

$$F(y|\mu,\phi) = \begin{cases} 0, & \text{if } y < 0, \\ 1 - H\left[\phi, -\log(y)\frac{\mu^{1/\phi}}{1-\mu^{1/\phi}}\right], & \text{if } y \in [0,1), \\ 1, & \text{if } y \geq 1, \end{cases}$$

in which $H(a,x) := \frac{1}{\Gamma(a)} \int_0^x t^{a-1}\exp(-t)dt$ is the regularized lower incomplete gamma function. The expected value and variance of a UG_{μ} distribution are given, respectively, by $\mathbb{E}(Y) = \mu$ and $\mathrm{Var}(Y) = \mu\{[1/(2-\mu^{1/\phi})^{\phi}] - \mu\}$.

Figure 1(a) shows the pdf for different values of μ with ϕ fixed, while Figure 1(b) shows the cdf of UG_{μ} distribution. In both cases in Figure 1(a), different patterns can be observed: symmetric, right-skewed, and left-skewed.

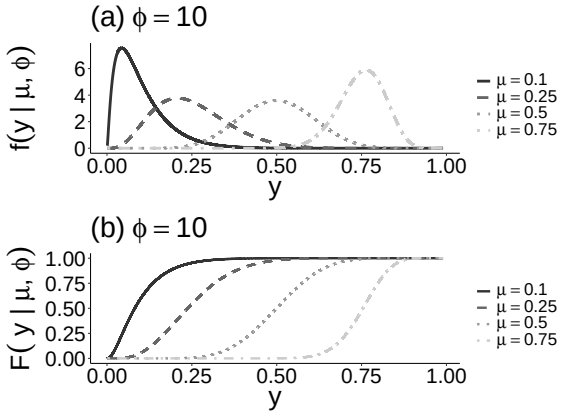


FIGURE 1: The pdf and cdf of the UG_μ distribution for different values of μ, with ϕ fixed.

The UG distribution, reparameterized in terms of quantiles as proposed by Rocha et al. (2024), with parameters $\sigma \in \mathbb{R}_+^*$ (shape parameter), $\theta_\tau \in (0,1)$ (location parameter), and $\tau \in (0,1)$, has a pdf given by

$$f(y|\theta_\tau,\sigma) = \left[\frac{H^{-1}(\sigma,1-\tau)}{-\log(\theta_\tau)}\right]^\sigma \frac{1}{\Gamma(\sigma)}\, y^{\frac{H^{-1}(\sigma,1-\tau)}{-\log(\theta_\tau)}-1}\, [-\log(y)]^{\sigma-1}\, \mathbb{1}_{(0,1)}(y), \quad (2)$$

where $H^{-1}(\sigma,\cdot)$ denotes the inverse (with respect to the second argument) of the regularized lower incomplete gamma function.[1]. So we say that $Y|\theta_\tau,\sigma \sim \text{UG}_\tau(\theta_\tau,\sigma)$. The cdf is given by

$$F(y|\theta_\tau,\sigma) = \begin{cases} 0, & \text{if } y < 0, \\ 1 - H\left[\sigma, -\log(y)\frac{H^{-1}(\sigma,1-\tau)}{-\log(\theta_\tau)}\right], & \text{if } y \in [0,1), \\ 1, & \text{if } y \geq 1, \end{cases}$$

the expected value and variance of a UG_τ distribution are given, respectively, by

$$\mathbb{E}(Y) = \left(\frac{H^{-1}(\sigma,1-\tau)}{H^{-1}(\sigma,1-\tau)-\log(\theta_\tau)}\right)^\sigma,$$

and

$$\text{Var}(Y) = \left(\frac{H^{-1}(\sigma,1-\tau)}{H^{-1}(\sigma,1-\tau)-2\log(\theta_\tau)}\right)^\sigma - \left(\frac{H^{-1}(\sigma,1-\tau)}{H^{-1}(\sigma,1-\tau)-\log(\theta_\tau)}\right)^{2\sigma}.$$

[1]For $\sigma > 0$ and $0 \leq p \leq 1$, the regularized lower incomplete gamma function is given by $H(\sigma,x) = \frac{\gamma(\sigma,x)}{\Gamma(\sigma)}$, where $\gamma(\sigma,x) := \int_0^x t^{\sigma-1}e^{-t}\,dt$. The inverse $H^{-1}(\sigma,p)$ satisfies $H(\sigma,H^{-1}(\sigma,p)) = p$.

Figures 2 (a) and 3 (a) show the pdfs for different values of μ and τ (extreme values, i.e., 0.10 and 0.90), with σ held fixed, while Figures 2 (b) and 3 (b) show the cdfs of the UG_τ distribution. In both cases in Figure 1 (a), different patterns can be observed: symmetric, right-skewed, and left-skewed.

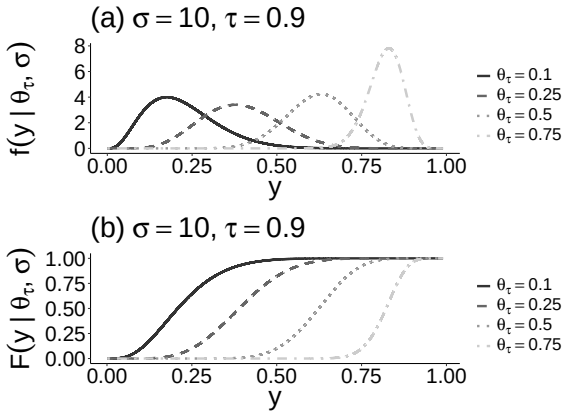


FIGURE 2: The pdf and cdf of the UG_τ distribution for different values of θ_τ and $\tau = 0.10$, with σ fixed.

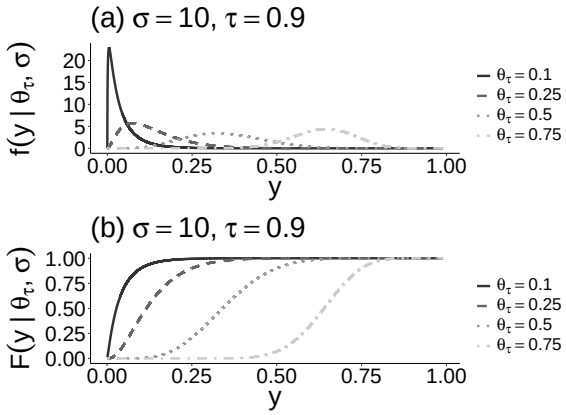


FIGURE 3: The pdf and cdf of the UG_τ distribution for different values of θ_τ and $\tau = 0.90$, with σ fixed.

3. The Reparameterized UG Regression Model in Terms of the Mean and Quantiles

Let the response variable y_i have a UG_μ with mean $\mu_i \in (0, 1)$ and precision parameter $\phi_i \in \mathbb{R}_+^*$, whose pdf is given in Equation (1). We write $y_i|\mu_i, \phi_i \sim \text{UG}_\mu(\mu_i, \phi_i)$, $i = 1, \ldots, n$, and assume that $y_1, \ldots, y_n$ are independent.

Similarly, let the response variable y_i follow a UG_τ distribution with quantile $\theta_{\tau(i)} \in (0,1)$ and shape parameter $\sigma_i \in \mathbb{R}_+^*$ whose pdf is given in Equation (2). We denote $y_i|\theta_{\tau(i)}, \sigma_i \sim \mathrm{UG}_\tau(\theta_{\tau(i)}, \sigma_i)$, $i = 1, \ldots, n$, and again assume that $y_1, \ldots, y_n$ are independent.

We assume that the regression structures for the mean μ_i, the precision parameter ϕ_i, the quantile $\theta_{\tau(i)}$ and the shape parameter σ_i are, respectively, as follows:

$$\underset{\text{(mean submodel)}}{g_\mu(\mu_i) = \mathbf{X}_i^\top \boldsymbol{\beta}}, \qquad \underset{\text{(precision submodel)}}{g_\phi(\phi_i) = \mathbf{Z}_i^\top \boldsymbol{\gamma}},$$

and

$$\underset{\text{(quantile submodel)}}{g_{\theta_\tau}(\theta_{\tau,i}) = \mathbf{W}_i^\top \boldsymbol{\lambda}}, \qquad \underset{\text{(shape submodel)}}{g_\sigma(\sigma_i) = \mathbf{V}_i^\top \boldsymbol{\kappa}},$$

where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_{p_1})^\top \in \mathbb{R}^{p_1}$, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_{p_2})^\top \in \mathbb{R}^{p_2}$, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_{p_3})^\top \in \mathbb{R}^{p_3}$ and $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_{p_4})^\top \in \mathbb{R}^{p_4}$ are vectors of unknown regression parameters; $\mathbf{X}_i = (x_{i1}, \ldots, x_{ip_1})^\top$, $\mathbf{Z}_i = (z_{i1}, \ldots, z_{ip_2})^\top$, $\mathbf{W}_i = (w_{i1}, \ldots, w_{ip_3})^\top$, and $\mathbf{V}_i = (v_{i1}, \ldots, v_{ip_4})^\top$ are vectors of covariates of lengths p_1, p_2, p_3, and p_4, respectively, with $p_1+p_2+p_3+p_4 = p < n$, and are assumed fixed and known. In addition, the link functions $g_\mu(\cdot) : (0,1) \to \mathbb{R}$, $g_\phi(\cdot) : (0,\infty) \to \mathbb{R}$, $g_{\theta\tau}(\cdot) : (0,1) \to \mathbb{R}$, and $g_\sigma(\cdot) : (0,\infty) \to \mathbb{R}$ are strictly monotonic and twice differentiable. Examples of link functions include the logit function, $h(t) = \log(t/(1-t))$; the probit function, $h(t) = \Phi^{-1}(t)$, where $\Phi^{-1}(t)$ denotes the quantile function of the standard normal distribution; the complementary log-log function, $h(t) = \log\{-\log(1-t)\}$; the cauchit function, $h(t) = \tan\{\pi(t-1/2)\}$; and the Aranda-Ordaz function, $h(t) = \log\left\{[(1-t)^{-\lambda} - 1]/\lambda\right\}$, with $\lambda > 0$. In this paper, we adopt the logit link for both models because it provides a straightforward interpretation of the regression parameters in terms of odds ratios, while acknowledging that alternative link functions may offer distinct properties and potential advantages. A more comprehensive description can be found in the following references: Atkinson (1985), McCullagh & Nelder (1989), Smithson & Verkuilen (2006) and Bourguignon & Gallardo (2025).

The regression coefficients associated with the central tendency parameter μ_i can be interpreted through the odds ratio (OR), defined as the ratio between μ_i and $1-\mu_i$. In this model, the odds take the form of an exponential function of the linear predictor:

$$\mathrm{odds}(\mathbf{X}_i) = \frac{\mu_i}{1-\mu_i} = \exp(\mathbf{X}_i^\top \boldsymbol{\beta}).$$

This expression allows for a straightforward interpretation of the model coefficients in terms of multiplicative effects on the odds. Specifically, $\exp(\beta_0)$ represents the odds when all covariates are set to zero. To assess the impact of a specific covariate, consider two individuals whose covariate vectors differ only in the jth component: $\mathbf{X}_i = (x_{i1}, \ldots, x_{ij}, \ldots, x_{ip_1})$ and $\mathbf{x}_i^* = (x_{i1}, \ldots, x_{ij}+1, \ldots, x_{ip_1})$. The odds ratio comparing these two profiles is given by:

$$\mathrm{OR}(\mathbf{X}_i^*, \mathbf{X}_i) = \frac{\mathrm{odds}(\mathbf{X}_i^*)}{\mathrm{odds}(\mathbf{X}_i)} = \exp(\beta_j),$$

indicating that a one-unit increase in the jth covariate results in a multiplicative change of $\exp(\beta_j)$ in the odds, holding all other covariates constant.

Similarly, if the model is parameterized in terms of a conditional quantile $\theta_{\tau(i)}$, the regression coefficients can also be interpreted via the odds ratio defined as:

$$\text{odds}_\tau(\mathbf{W}_i) = \frac{\theta_{\tau(i)}}{1-\theta_{\tau(i)}} = \exp(\mathbf{W}_i^\top \boldsymbol{\lambda}),$$

where $\boldsymbol{\delta}$ denotes the vector of regression coefficients associated with the τ-th quantile. In this context, $\exp(\delta_0)$ corresponds to the odds associated with the τ-th quantile when all covariates are zero. Analogously, the odds ratio for a one-unit increase in the jth covariate is given by:

$$\text{OR}_\tau(\mathbf{W}_i^*, \mathbf{W}_i) = \frac{\text{odds}_\tau(\mathbf{W}_i^*)}{\text{odds}_\tau(\mathbf{W}_i)} = \exp(\lambda_j),$$

indicating that δ_j has the same multiplicative interpretation as in the mean-based model, but now in terms of the τ-th quantile of the response distribution.

Moreover, the use of the logarithmic link function for the parameters ϕ and σ ensures that these parameters remain strictly positive, thereby respecting the natural constraints of the model. Furthermore, the log link allows the effects of covariates on dispersion to be modeled multiplicatively, which facilitates the interpretation of regression coefficients. Canterle & Bayer (2019) investigate beta regression models with regression structures for both the mean and the dispersion, emphasizing that the use of the logarithmic link for the dispersion contributes to greater numerical stability and more transparent interpretation of covariate effects. Additionally, Smithson & Verkuilen (2006) employ the logarithmic link with a negative sign for ϕ to enhance the interpretability of the coefficients associated with the precision submodel.

4. Bayesian Approach

4.1. Hierarchical Structure and Likelihood Function

The Bayesian approach for the UG_μ model can be developed similarly to the UG_τ model by assigning suitable prior distributions to the regression coefficient vectors $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\lambda}$, and $\boldsymbol{\kappa}$. Consider the following hierarchical structure for both Bayesian models

$$\begin{gathered} y_i \mid \boldsymbol{\beta}, \boldsymbol{\gamma} \sim \text{UG}_\mu(g_\mu^{-1}(\mathbf{X}_i^\top \boldsymbol{\beta}), g_\phi^{-1}(\mathbf{Z}_i^\top \boldsymbol{\gamma})), \\ (\boldsymbol{\beta}, \boldsymbol{\gamma})^\top \sim \pi(\boldsymbol{\beta}, \boldsymbol{\gamma}) \end{gathered}$$

and

$$\begin{gathered} y_i \mid \boldsymbol{\lambda}, \boldsymbol{\kappa} \sim \text{UG}_\tau(g_{\theta_\tau}^{-1}(\mathbf{W}_i^\top \boldsymbol{\lambda}), g_\sigma^{-1}(\mathbf{V}_i^\top \boldsymbol{\kappa})), \quad i = 1, \ldots, n. \\ (\boldsymbol{\lambda}, \boldsymbol{\kappa})^\top \sim \pi(\boldsymbol{\lambda}, \boldsymbol{\kappa}) \end{gathered}$$

Moreover, the likelihood function for the UG_{μ} and UG_{τ} regression models is given by

$$L(\boldsymbol{\beta},\boldsymbol{\gamma}\mid \mathbf{y}) \;=\; \prod_{i=1}^{n}\left[\frac{\mu_i^{1/\phi_i}}{1-\mu_i^{1/\phi_i}}\right]^{\phi_i}\frac{1}{\Gamma(\phi_i)}\, y_i^{\frac{\mu_i^{1/\phi_i}}{1-\mu_i^{1/\phi_i}}-1}\,[-\log(y_i)]^{\phi_i-1},$$

and

$$L(\boldsymbol{\lambda},\boldsymbol{\kappa}\mid \mathbf{y}) \;=\; \prod_{i=1}^{n}\left[\frac{H^{-1}(\sigma_i,1-\tau)}{-\log(\theta_{\tau(i)})}\right]^{\sigma_i}\frac{1}{\Gamma(\sigma_i)}\, y_i^{\frac{H^{-1}(\sigma_i,1-\tau)}{-\log(\theta_{\tau(i)})}-1}\,[-\log(y_i)]^{\sigma_i-1},$$

where $\mu_i = g_{\mu}^{-1}(\mathbf{X}_i^{\top}\boldsymbol{\beta})$, $\phi_i = g_{\phi}^{-1}(\mathbf{Z}_i^{\top}\boldsymbol{\gamma})$, $\theta_{\tau(i)} = g_{\theta_\tau}^{-1}(\mathbf{W}_i^{\top}\boldsymbol{\lambda})$ and $\sigma_i = g_{\sigma}^{-1}(\mathbf{V}_i^{\top}\boldsymbol{\kappa})$. Since no prior information is available regarding the dependence among $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\lambda}$, and $\boldsymbol{\kappa}$, we assume conditional independence such that

$$\pi(\boldsymbol{\beta},\boldsymbol{\gamma}) = \pi(\boldsymbol{\beta})\pi(\boldsymbol{\gamma}) \quad \text{and} \quad \pi(\boldsymbol{\lambda},\boldsymbol{\kappa}) = \pi(\boldsymbol{\lambda})\pi(\boldsymbol{\kappa}).$$

This assumption simplifies the prior specification and computation. Furthermore, because the link functions map $(0,1)$ and $(0,\infty)$ into $\mathbb{R}$, all regression coefficient vectors $(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\lambda},\boldsymbol{\kappa})$ naturally live in $\mathbb{R}^{p_j}$. Positivity constraints apply to the parameters (ϕ_i,σ_i) via the links, not to the coefficients themselves. We therefore place priors on $\mathbb{R}^{p_1}$, $\mathbb{R}^{p_2}$, $\mathbb{R}^{p_3}$, and $\mathbb{R}^{p_4}$ for $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\lambda}$, and $\boldsymbol{\kappa}$, respectively. In particular, we adopt multivariate Student-t priors $\boldsymbol{\beta} \sim t_{p_1}\big(\nu_\beta,\boldsymbol{\mu}_\beta,\boldsymbol{\Sigma}_\beta\big)$, $\boldsymbol{\gamma} \sim t_{p_2}\big(\nu_\gamma,\boldsymbol{\mu}_\gamma,\boldsymbol{\Sigma}_\gamma\big)$, $\boldsymbol{\lambda} \sim t_{p_3}\big(\nu_\lambda,\boldsymbol{\mu}_\lambda,\boldsymbol{\Sigma}_\lambda\big)$, and $\boldsymbol{\kappa} \sim t_{p_4}\big(\nu_\kappa,\boldsymbol{\mu}_\kappa,\boldsymbol{\Sigma}_\kappa\big)$, with $\boldsymbol{\mu}_\beta = \boldsymbol{\mu}_\gamma = \boldsymbol{\mu}_\lambda = \boldsymbol{\mu}_\kappa = \mathbf{0}$ and $\nu_\beta = \nu_\gamma = \nu_\lambda = \nu_\kappa = 4$ so that the prior variance is finite. We set $\boldsymbol{\Sigma}_\beta = \boldsymbol{\Sigma}_\gamma = \boldsymbol{\Sigma}_\lambda = \boldsymbol{\Sigma}_\kappa = 50\,\mathbf{I}$, which implies $\mathrm{Var}(\boldsymbol{\beta}) = \mathrm{Var}(\boldsymbol{\gamma}) = \mathrm{Var}(\boldsymbol{\lambda}) = \mathrm{Var}(\boldsymbol{\kappa}) = 100\,\mathbf{I}$ because, for a t_ν prior with $\nu > 2$, $\mathrm{Var}(\cdot) = \frac{\nu}{\nu-2}\,\boldsymbol{\Sigma}$. These heavy-tailed priors serve as weakly informative, robust regularizers that stabilize estimation under separation and collinearity while avoiding excessive shrinkage.

In Bayesian inference, several summary statistics can be used as point estimates of model parameters based on their posterior distributions. The posterior mean is widely adopted due to its desirable decision theoretic property of minimizing the expected squared error loss (Gelman et al., 2013). Alternatively, the posterior median is a robust estimator that minimizes the expected absolute deviation and is particularly useful in the presence of skewed or heavy-tailed posterior distributions (Robert, 2007). The posterior mode, often referred to as the maximum a posteriori estimate, corresponds to the mode of the posterior distribution and may be preferred in settings where the posterior is unimodal and computational efficiency is critical (Bernardo & Smith, 1994). Each of these estimators has advantages and limitations depending on the context, such as symmetry, multimodality, or robustness to outliers. When reporting results, it is common practice to provide not only point estimates but also credible intervals to reflect posterior uncertainty (McElreath, 2020). The choice among these summaries should be guided by the modeling goals and the inferential context. In this paper, we consider three posterior point estimates: the posterior mean (PM), the posterior median (Med), and the maximum a posterior (MAP).

4.2. Model Implementation Using NIMBLE

We used `NIMBLE` (de Valpine et al., 2017, 2024) in R (R Core Team, 2024) to obtain posterior samples for our Bayesian hierarchical model, for which the full conditional distributions are not available in closed form. `NIMBLE` allows model specification in the flexible BUGS dialect directly within R, automatically compiles the model and custom MCMC samplers into `C++` for efficient execution, and offers fine-grained control over sampler configuration (e.g., blocking, tuning, user-written samplers). Comparative benchmarks (Beraha et al., 2021) show that `NIMBLE` can outperforms JAGS and Stan in effective sample size and runtime for mixture models and conjugate structures. Moreover, its extensibility supports state-of-the-art sampling methods beyond MCMC (e.g. particle filters implemented in `nimbleSMC` and MCEM), making it well suited to complex, computationally intensive Bayesian inference.

The complete compilation workflow followed by `NIMBLE` is detailed in Algorithm 1, illustrating how the source model and algorithm are transformed into compiled objects ready for execution.

Algorithm 1 Model and Algorithm Compilation in `NIMBLE`

Input: Model, Algorithm
Process:
  Send Model and Algorithm to the **NIMBLE Compiler**
  The NIMBLE Compiler generates C++ code
  Send C++ code to the **C++ Compiler**
  Compile to generate:
    Compiled Model
    Compiled Algorithm
Output: Compiled Model, Compiled Algorithm

5. Model Fit Assessment and Model Comparison

5.1. Model Adequacy

To assess the adequacy of the proposed model, we employ the randomized quantile residuals introduced by Dunn & Smyth (1996). They are defined as

$$r_i^q = \Phi^{-1}\left(F(y_i; \widehat{\boldsymbol{\Omega}})\right),$$

where $\widehat{\boldsymbol{\Omega}}$ denotes an appropriate Bayesian estimate of the parameter vector $\boldsymbol{\Omega}$, such as the posterior mean considered in this study. In this expression, $F(\cdot; \widehat{\boldsymbol{\Omega}})$ represents cdf of the UG_μ or UG_τ, and $\Phi(\cdot)$ denotes the standard normal cdf.

If the fitted model is correctly specified, the randomized quantile residuals $\{r_i^q\}$ are approximately independent draws from a standard normal distribution. Therefore, diagnostic plots such as histograms or Q–Q plots of r_i^q against the quantiles

of the standard normal distribution can be used to visually check the adequacy of the model fit. Significant deviations from normality may indicate misspecification of the model, such as an incorrect link function or omitted covariates.

As a complementary diagnostic tool, we also employ the Cox–Snell residuals to evaluate whether the fitted regression model is appropriate, we consider the Cox–Snell residuals (Cox & Snell, 1968), defined as

$$r_i = -\log S(y_i \mid \widehat{\boldsymbol{\Omega}}), \quad i = 1, \ldots, n, \tag{3}$$

where $S(\cdot \mid \widehat{\boldsymbol{\Omega}}) = 1 - F(\cdot \mid \widehat{\boldsymbol{\Omega}})$ denotes the survival function estimated from the fitted model.

These residuals have the important property that, if the model fits the data correctly, the values r_i follow a standard exponential distribution. Therefore, if we order the residuals and plot $r_{(i)}$ against $-\log \hat{S}(r_i)$, where $\hat{S}(r_i)$ is the Kaplan–Meier estimator for the residuals, the points are expected to fall approximately along a straight line with intercept zero and slope one (Lee & Wang, 2003).

A practical way to assess this assumption is through the half-normal plot with simulated envelope proposed by Atkinson (1985). The procedure can be summarized as follows:

(i) Fit the model to the data and generate a sample of size n from the estimated parameters of the fitted model;

(ii) Refit the model to this generated sample, compute the absolute values of the residuals, and order them in increasing order;

(iii) Repeat steps (i) and (ii) B times to obtain B sets of ordered residuals;

(iv) For each ordered position i, compute the quantiles $\gamma/2$, the median, and the quantile $1-\gamma/2$ across the B simulated sets;

(v) Plot the ordered residuals from the original sample against the expected order statistics of a half-normal distribution, which can be approximated by

$$\Phi^{-1}\left(\frac{i+n-0.125}{2n+0.5}\right),$$

where Φ^{-1} denotes the quantile function of the standard normal distribution.

According to Atkinson (1985), if the model is correctly specified, no more than $\gamma \times 100\%$ of the observations should lie outside the envelope bands. If a substantial proportion of the data points exceed these limits, this provides evidence against the adequacy of the fitted model.

5.2. Model Comparison Via WAIC

A variety of criteria can be employed to evaluate competing models fitted to a dataset and identify the one that most adequately represents the underlying

data-generating process. In this work, we adopt the Watanabe Akaike Information Criterion (WAIC) (Watanabe, 2010), a fully Bayesian alternative to classical information criteria, which approximates leave-one-out cross-validation (Gelman et al., 2014). The WAIC is computed as

$$\text{WAIC} = -2\left(\text{lppd} - p_{\text{WAIC}}\right),$$

where lppd denotes the log pointwise predictive density, given by

$$\text{lppd} = \sum_{i=1}^{n} \log \int p(y_i \mid \boldsymbol{\Omega})\, p(\boldsymbol{\Omega} \mid \mathbf{y})\, d\boldsymbol{\Omega},$$

with $\boldsymbol{\Omega}$ representing the full vector of model parameters. The penalty term p_{WAIC} accounts for the effective number of parameters and is computed as

$$p_{\text{WAIC}} = \sum_{i=1}^{n} \text{Var}_{\text{post}}\left[\log p(y_i \mid \boldsymbol{\Omega})\right],$$

where Var_{post} denotes the posterior variance of the log-likelihood contribution for each observation, estimated from MCMC draws. This penalization reflects model complexity and serves to prevent overfitting.

The WAIC is advantageous because it averages over the posterior and can be straightforwardly estimated using samples from MCMC draws. Among a set of candidate models, the one with the smallest WAIC value is selected as the one with superior expected predictive accuracy.

In our analysis, WAIC values were computed using the `loo` package in R (Vehtari et al., 2017). Following the recommendations of Merkle et al. (2019), we emphasize that the choice between conditional and marginal log-likelihood formulations depends on the inferential goals.

5.3. Model Evaluation via Conditional Predictive Ordinate

In Bayesian model fitting, the specification of a prior distribution is essential. In the absence of substantive prior knowledge, an alternative approach is to evaluate and compare models through their predictive performance. We employ the Conditional Predictive Ordinate (CPO) (Geisser & Eddy, 1979), an observation-wise leave-one-out predictive density that quantifies the model's ability to predict each data point based on the posterior built from the remaining data.

Let $\mathcal{D}$ denote the complete dataset, and let $\mathcal{D}^{-i}$ represent the dataset with the ith observation removed. The CPO for observation y_i is defined as the predictive density of y_i under the model estimated from $\mathcal{D}^{-i}$, that is,

$$\text{CPO}_i = \pi(y_i \mid \mathcal{D}^{-i}) = \left\{ \int \frac{\pi(\boldsymbol{\Omega} \mid \mathcal{D})}{\pi(y_i \mid \boldsymbol{\Omega})}\, d\boldsymbol{\Omega} \right\}^{-1}.$$

Following Christensen et al. (2011), the CPO can be approximated using MCMC samples $\boldsymbol{\Omega}^{(1)}, \ldots, \boldsymbol{\Omega}^{(M)}$ from the posterior distribution $\pi(\boldsymbol{\Omega} \mid \mathcal{D})$ as

$$\widehat{\mathrm{CPO}}_i = \left\{ \frac{1}{M} \sum_{k=1}^{M} \frac{1}{f_i(y_i \mid \boldsymbol{\Omega}^{(k)}, \mathcal{D})} \right\}^{-1},$$

where $f_i(y_i \mid \boldsymbol{\Omega}^{(k)}, \mathcal{D})$ denotes the likelihood contribution of y_i evaluated at the kth MCMC draw $\boldsymbol{\Omega}^{(k)}$. A higher CPO value indicates greater predictive accuracy, i.e., the model better explains the data point in question.

To assess global model performance, the log pseudo-marginal likelihood (LPML) is computed as the sum of the log-CPO values across all observations:

$$\mathrm{LPML} = \sum_{i=1}^{n} \log(\widehat{\mathrm{CPO}}_i).$$

Among competing models, the one with the highest LPML is considered to have superior overall predictive adequacy.

5.4. Local Influence via Kullback–Leibler Divergence

To assess the presence of influential observations within a Bayesian regression framework, we employ the Kullback–Leibler (K–L) divergence as a diagnostic criterion, following the methodology proposed by Cho et al. (2009). The K–L divergence quantifies the discrepancy between the posterior distributions with and without a given observation and is formally defined as

$$K(\mathbb{P}, \mathbb{P}^{-i}) = \int \pi(\boldsymbol{\Omega} \mid \mathcal{D}) \log \left(\frac{\pi(\boldsymbol{\Omega} \mid \mathcal{D})}{\pi(\boldsymbol{\Omega} \mid \mathcal{D}^{-i})} \right) d\boldsymbol{\Omega},$$

where $\boldsymbol{\Omega}$ denotes the vector of model parameters, $\mathbb{P}$ is the posterior distribution based on the full dataset $\mathcal{D}$, and $\mathbb{P}^{-i}$ is the posterior distribution obtained by excluding the ith observation.

According to Cho et al. (2009), the K–L divergence may also be expressed as

$$K(\mathbb{P}, \mathbb{P}^{-i}) = -\log(\mathrm{CPO}_i) + \mathbb{E}_{\boldsymbol{\Omega}|\mathcal{D}} \left[\log f(y_i \mid \boldsymbol{\Omega}) \right],$$

where CPO_i is the conditional predictive ordinate for the ith observation and $f(y_i \mid \boldsymbol{\Omega})$ is the likelihood contribution. The expectation is taken with respect to the posterior distribution of $\boldsymbol{\Omega}$ under the full dataset $\mathcal{D}$. This representation enables the practical computation of the K–L divergence using posterior samples obtained via MCMC methods.

Following the approach of Vidal & Castro (2010), an observation is considered influential when $K(\mathbb{P}, \mathbb{P}^{-i}) > 0.14$. Although this threshold has intuitive appeal, it may be overly conservative. Thus, alternative rules have been proposed, such as comparing the K–L divergence of each observation to twice the mean of all $K(\mathbb{P}, \mathbb{P}^{-i})$ values, providing a more flexible and adaptive criterion.

Furthermore, McCulloch (1989) introduced a calibration method for interpreting K–L divergences through comparison with divergences between Bernoulli distributions. Letting $K(Q_1, Q_2) = k$, where Q_1 and Q_2 are two posterior distributions, one may define a calibration value $q(k)$ such that

$$K(\text{Ber}\{0.5\}, \text{Ber}\{q(k)\}) = k,$$

where $\text{Ber}\{p\}$ denotes a Bernoulli distribution with parameter p. If $q(k) \approx 0.5$, the distributions Q_1 and Q_2 are considered similar; deviations from 0.5 indicate growing dissimilarity. When p_i is defined as the probability associated with the ith observation, the K–L divergence may be approximated as

$$K(\mathbb{P}, \mathbb{P}^{-i}) = K(\mathbb{B}\{0.5\}, \mathbb{B}\{p_i\}) = -\frac{1}{2}\log\left(4p_i(1-p_i)\right).$$

This calibration provides an interpretable scale for evaluating the relative influence of observations in Bayesian models.

Moreover, the K–L divergence can be transformed into a probability scale to aid interpretation. Specifically, following Cho et al. (2009), the influence measure p_i associated with the ith observation is defined as

$$p_i = \frac{1 + \sqrt{1 - \exp\left\{-2K(\mathbb{P}, \mathbb{P}^{-i})\right\}}}{2}, \quad \text{with} \quad p_i \in [0.5, 1].$$

This transformation provides an interpretable metric: values of p_i closer to 0.5 suggest potential influence, whereas values near 1 indicate better agreement between the full and leave-one-out posterior distributions. Consequently, p_i can be employed to assess how well the model explains each observation individually, where deviations from 0.5 are indicative of local sensitivity and potential leverage.

5.5. Posterior Predictive Checking

Posterior predictive checking evaluates whether data replicated from the fitted model resemble the observed data. Following Rubin (1984), and as further developed by Meng (1994) and Gelman et al. (1996), the idea is that, under a well-specified model, replicated data generated from the posterior predictive distribution should exhibit behavior similar to the observed data. Formally, the posterior predictive distribution is given by

$$p(y^{\text{rep}} \mid y) = \int p(y^{\text{rep}} \mid \boldsymbol{\Omega})\, p(\boldsymbol{\Omega} \mid y)\, d\boldsymbol{\Omega},$$

where y^{rep} denotes replicated data generated under the model and $\boldsymbol{\Omega}$ is the parameter vector.

The comparison between observed and replicated data is performed via test quantities $T(y, \boldsymbol{\Omega})$, which may depend on both the data and the parameters. The discrepancy between $T(y, \boldsymbol{\Omega})$ and $T(y^{\text{rep}}, \boldsymbol{\Omega})$ is summarized by the posterior predictive p-value, defined as

$$\begin{aligned} p_B &= \int\int \mathbb{1}\{T(y^{\text{rep}}, \boldsymbol{\Omega}) \geq T(y, \boldsymbol{\Omega})\} p(y^{\text{rep}} \mid \boldsymbol{\Omega}) p(\boldsymbol{\Omega} \mid y)\, dy^{\text{rep}} d\boldsymbol{\Omega} \\ &= \mathbb{P}(T(y^{\text{rep}}, \boldsymbol{\Omega}) \geq T(y, \boldsymbol{\Omega}) \mid y). \end{aligned}$$

In this expression, we use the conditional independence property of the posterior predictive distribution, namely $p(y^{\text{rep}} \mid \boldsymbol{\Omega}, y) = p(y^{\text{rep}} \mid \boldsymbol{\Omega})$; see Gelman et al. (1995) for details. The resulting posterior predictive p-value, interpreted as a posterior probability, provides evidence against the model when it is close to 0 or 1.

Extreme p-values (e.g., less than 0.01 or greater than 0.99) indicate substantial model inadequacy with respect to the specific aspect being tested, and may justify model expansion or revision. Mild discrepancies may suggest minor model refinements or may be disregarded if they do not affect the primary inferential goals.

The choice of T plays a central role in determining the sensitivity and diagnostic value of the posterior predictive check. Below, we delineate the principal classes of discrepancy functions employed in the literature, each tailored to detect specific forms of model misspecification.

Among the various types of discrepancy functions used in posterior predictive checks, several categories stand out due to their practical utility and interpretability. For example, the class includes classical summary statistics, such as the sample mean $T(y) = \bar{y}$, $T(y) = \tilde{y}$ the median, $T(y) = \text{Var}(y)$, and the extreme values $T(y) = \min_i y_i$ and $T(y) = \max_i y_i$. These functions are particularly useful for detecting global deviations in location, scale, and tail behavior between the model and the data.

Finally, residual-based discrepancy functions are often used when the model provides conditional expectations. In such cases, the sum of squared residuals, $T(y, \boldsymbol{\Omega}) = \sum_{i=1}^{n} \big[y_i - \mathbb{E}(y_i \mid \boldsymbol{\Omega})\big]^2$, or the proportion of large standardized (or studentized) residuals can reveal local lack of fit, particularly under heteroscedasticity, outliers, or misspecification of the conditional variance.

6. Simulation Studies

6.1. Computational Aspects

All simulations and computational analyses presented in this paper were performed using the Virtual Laboratory of the Institute of Mathematics, Statistics and Scientific Computing at the University of Campinas (IMECC/Unicamp). The computational environment consists of a cluster of high-performance workstations, each equipped with an Intel® Core™ i9-13900K processor (up to 5.8 GHz, 6 cores and 12 threads) and 128 GB of RAM, running the Debian Testing operating system.

To assess estimator performance in the Monte Carlo study, we considered the bias, mean squared error (MSE), mean absolute error (MAE), absolute value of relative bias (AVRB), and coverage rate (CR). We generated a total of 85000 MCMC samples across two parallel chains, discarding the first 5000 iterations as burn-in and applying a thinning interval of 10, resulting in 8000 posterior samples per chain (16000 in total). Specifically, for each parameter Ψ (the true value) and replication $r = 1, \ldots, 500$ with estimate $\widehat{\Psi}_r$, we computed: the MSE,

$\text{MSE}(\widehat{\Psi}) = \frac{1}{500}\sum_{r=1}^{500}(\widehat{\Psi}_r - \Psi)^2$, which accounts for both systematic error and estimator variability; the bias, $\text{Bias}(\widehat{\Psi}) = \frac{1}{500}\sum_{r=1}^{500}(\widehat{\Psi}_r - \Psi)$, which captures the systematic tendency to over or underestimate the true value; the MAE, $\text{MAE}(\widehat{\Psi}) = \frac{1}{500}\sum_{r=1}^{500}|\widehat{\Psi}_r - \Psi|$, which is more robust to extreme estimates than the MSE; the AVRB, $\text{AVRB}(\widehat{\Psi}) = \frac{1}{500}\sum_{r=1}^{500}\left|\widehat{\Psi}_r/\Psi - 1\right|$, which provides a relative-error measure with respect to the magnitude of Ψ. Finally, we computed the coverage rate (CR), defined as the proportion of replications in which the true value Ψ was contained within the corresponding 95% credible interval, i.e., $\text{CR}(\widehat{\Psi}) = \frac{1}{500}\sum_{r=1}^{500}\mathbb{1}\left(\Psi \in \text{HPD}_{r,0.95}\right)$, where $\mathbb{1}(\cdot)$ is the indicator function and $\text{HPD}_{r,0.95}$ denotes the 95% HPD interval for $\widehat{\Psi}_r$.

6.2. Simulation Study for the Ug Regression Model Reparameterized in Terms of the Mean and Quantiles

To evaluate and compare the performance of the UG regression models parameterized by the mean and by quantiles, we conducted a Monte Carlo simulation study. Both formulations were assessed over multiple sample sizes.

UG_μ model: The mean parameterized UG regression model was examined considering the true regression coefficients $(\beta_0, \beta_1) = (-0.5, 1)$ for the mean submodel and $(\gamma_0, \gamma_1) = (1, 1)$ for the precision submodel. Three sample sizes were considered: $n \in \{50, 100, 500\}$. For each sample size, 500 independent datasets were simulated. The covariate matrices $\mathbf{X}$ and $\mathbf{Z}$ associated with the mean and precision predictors, respectively were generated once for each value of n and kept fixed across all replications. Each matrix included an intercept (a column of ones) and a single continuous covariate. For the mean submodel, the covariate was independently sampled from a uniform distribution $\mathcal{U}(0, 1)$, while for the precision submodel, it was drawn from a Bernoulli distribution with success probability 0.5, Ber(0.5). The response variable was then generated from the UG_μ distribution according to the true model parameters and fixed covariates.

UG_τ model: In parallel, we examined the UG model parameterized at specific quantiles, using the same true values for the regression coefficients: $(\lambda_0, \lambda_1) = (-0.5, 1)$ and $(\kappa_0, \kappa_1) = (1, 1)$. Instead of a single level of the response distribution, this model was evaluated across three quantiles of interest: $\tau \in \{0.1,\ 0.5,\ 0.9\}$, providing insight into model behavior at the lower, central, and upper parts of the distribution. As in the mean based case, we used the same sample sizes $(n = 50, 100, 500)$ and replications per n (500). The covariate matrices $\mathbf{W}$ and $\mathbf{V}$, corresponding to the linear predictors of the quantile-based parameters μ_i and ϕ_i, were also fixed for each n, and constructed analogously to $\mathbf{X}$ and $\mathbf{Z}$. Each matrix contained an intercept and a single covariate, drawn from $\mathcal{U}(0, 1)$ for $\mathbf{W}$ and from Ber(0.5) for $\mathbf{V}$, respectively. The response variable was simulated from the UG_τ distribution using appropriate reparameterization for the chosen quantile levels.

For each simulated data set and model specification, posterior sampling was implemented in NIMBLE. We routinely monitored chain behavior by visual inspection of trace plots and autocorrelation functions, complemented by the Gelman–Rubin

potential scale reduction factor and effective sample size summaries (Gelman & Rubin, 1992). Across all scenarios, these diagnostics indicated satisfactory mixing and no evidence of lack of convergence; consequently, the retained post–burn-in draws were used to compute the summaries reported below.

This dual framework design enables a systematic comparison of model behavior under identical conditions, isolating the effect of parameterization choice (mean vs. quantile) on inferential performance.

The results of the simulation study, presented in Figures 4–7, allow us to evaluate the performance of the posterior estimates across different scenarios and sample sizes. In general, the mean absolute error (MAE) and the mean squared error (MSE) decrease systematically as the sample size increases, indicating greater precision and consistency, while the bias remains close to zero in almost all settings. In Figure 4 reports results for β_0, β_1, γ_0, and γ_1, showing satisfactory performance even for small sample sizes. By contrast, Figures 5–7, which summarize λ_0, λ_1, κ_0, and κ_1, show greater sensitivity in small samples, particularly in the scenario of Figure 5, where a more pronounced bias and higher initial variability are observed. Comparatively, the scenarios in Figures 6 and 7 illustrate more consistent performance, even with moderate sample sizes, suggesting improved estimation accuracy. Overall, the findings indicate good parameter estimation recovery, though in some cases larger sample sizes may be required to ensure the desired level of precision.

The results of the 95% coverage intervals for the UG_{μ} and UG_{τ} models are presented in Figures 8–11. From these figures, we observe that for the sample size $n = 50$, all evaluated parameters exhibited coverage values well below the nominal 95% level. In the UG_{μ} model, the parameters associated with the mean showed the lowest coverage values when compared with the precision parameters. In contrast, in the UG_{τ} model, the parameters related to the quantile presented higher coverage values than those associated with the shape.

When increasing the sample size to $n = 100$, the same pattern persists, although with an overall improvement in the coverage values. For $n = 500$, the coverage intervals become substantially closer to the nominal 95% level. Therefore, among the sample sizes considered, $n = 500$ yielded the coverage estimates closest to the nominal value for both models.

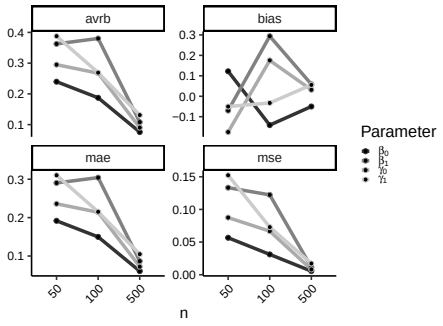


FIGURE 4: AVRB, Bias, MAE and MSE of $\beta_0, \beta_1, \gamma_0$ and γ_1 after 500 replications of the UG_μ model.

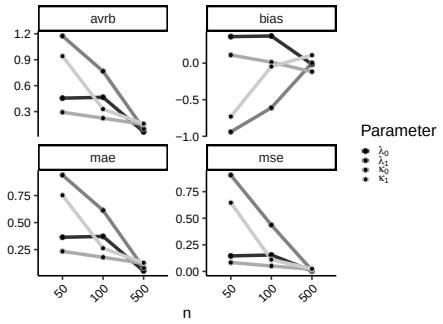


FIGURE 5: AVRB, Bias, MAE and MSE of $\lambda_0, \lambda_1, \kappa_0$ and κ_1 after 500 replications of the UG_τ model with $\tau = 0.10$.

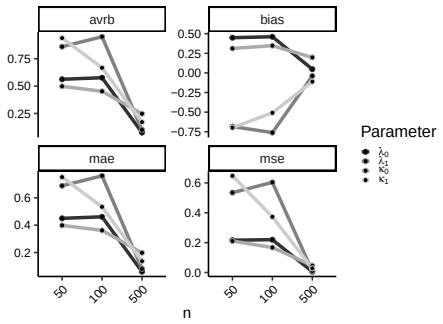


FIGURE 6: AVRB, Bias, MAE and MSE of $\lambda_0, \lambda_1, \kappa_0$ and κ_1 after 500 replications of the UG_τ model with $\tau = 0.50$.

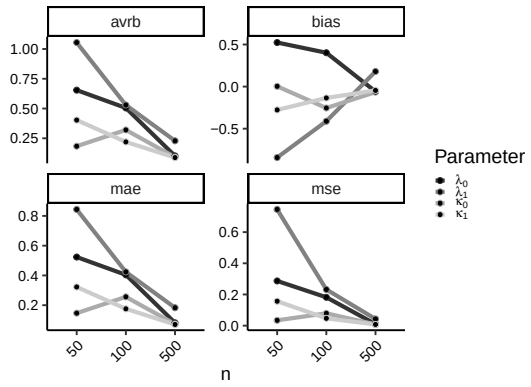


FIGURE 7: AVRB, Bias, MAE and MSE of $\lambda_0, \lambda_1, \kappa_0$ and κ_1 after 500 replications of the UG_τ model with $\tau = 0.90$.

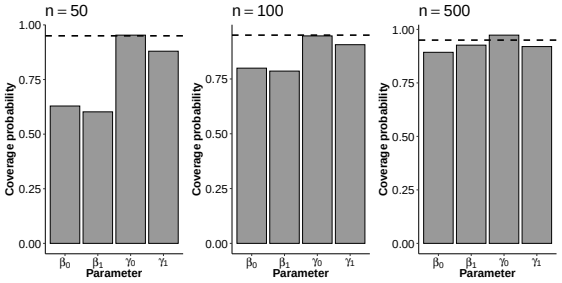


FIGURE 8: 95% CP for $\beta_0, \beta_1, \gamma_0$ and γ_1 after 500 replications of the UG_μ model with $n = 50, 100$ and 500.

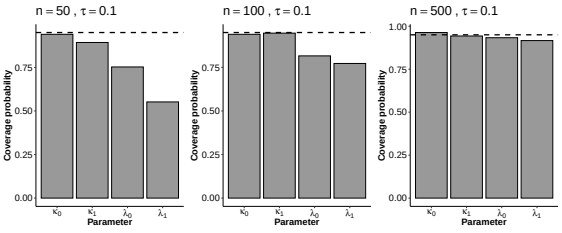


FIGURE 9: 95% CP for $\lambda_0, \lambda_1, \kappa_0$ and κ_1 after 500 replications of the UG_τ ($\tau = 0.1$) model with $n = 50, 100$ and 500.

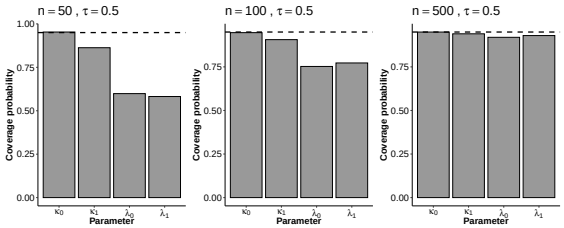


FIGURE 10: 95% CP for $\lambda_0, \lambda_1, \kappa_0$ and κ_1 after 500 replications of the UG_τ ($\tau = 0.5$) model with $n = 50, 100$ and 500.

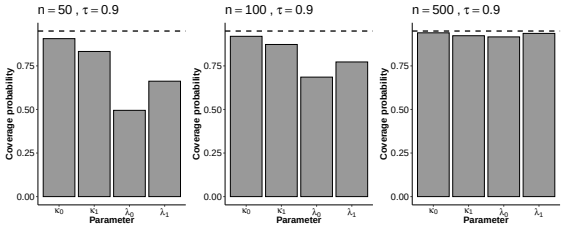


FIGURE 11: 95% CP for $\lambda_0, \lambda_1, \kappa_0$ and κ_1 after 500 replications of the UG_τ ($\tau = 0.9$) model with $n = 50, 100$ and 500.

7. Real Data Application

In this section, we present a real data analysis to illustrate the potential applications of the proposed UG_τ regression model. For comparison, we also consider the unit-Weibull (UW) quantile regression model introduced by Mazucheli et al. (2020).

7.1. Educational Attainment Data Set

The Better Life Index (BLI) data from OECD countries are used in this study. The dataset contains information on the educational attainment of OECD member countries and is available at `https://stats.oecd.org/index.aspx?DataSetCode=BLI`. The aim of the analysis is to examine the relationship between the educational attainment values (y) and two explanatory variables: labor market insecurity (LMI) (x_1) and homicide rate (HR) (x_2). This dataset has previously been examined by Altun (2021) and Mazucheli et al. (2023).

The distribution of the educational attainment is illustrated in Figure 12 through violin plots (a) and boxplot (b), providing a joint visualization of the density of the observed values. The results indicate that the variable predominantly takes on high values, mostly between 0.75 and 0.88, with a median of 0.8150 and a slightly lower mean (0.7724). The interquartile range is narrow (0.1225), suggesting low dispersion and homogeneity in the sample. The shape of the empirical distribution

reveals a negative skew, with higher density around 0.8 and a tail toward lower values near 0.4. Although the minimum recorded value is 0.37, such cases are rare and do not characterize the presence of extreme outliers.

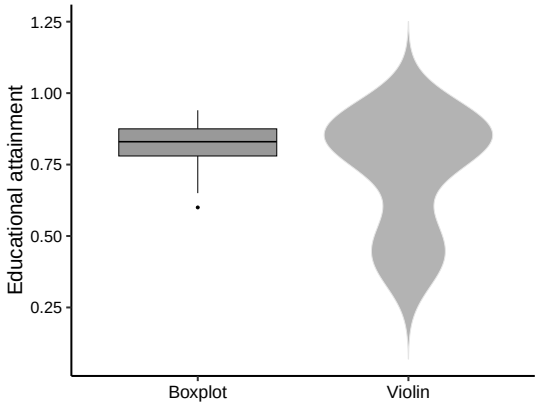


FIGURE 12: Violin plot (a) and boxplot (b) for educational attainment.

Given the strong concentration of educational attainment values in the upper range and the relatively small variability around the median, it seems more appropriate to focus on the 0.9 quantile rather than the 0.5 quantile. While the median (0.8150) summarizes central tendency, it does not fully capture the behavior of the highest-performing observations that are of primary interest in this context. In contrast, the 0.9 quantile better reflects the distributional tail where values approach the maximum, providing a more informative assessment of the factors associated with superior outcomes. We fitted a UG_τ regression model at $\tau = 0.9$ and, for comparison, a UW regression model also at $\tau = 0.9$. The fitted models are $y_i \overset{ind.}{\sim} \mathrm{UG}_\tau(\theta_i, \sigma)$, $y_i \overset{ind.}{\sim} \mathrm{UW}(\theta_i, \sigma)$ with $\mathrm{logit}(\theta_i) = \lambda_0 + \lambda_1 \mathrm{LMI}_i + \lambda_2 \mathrm{HR}_i$, $\log(\sigma) = -\kappa_0$ for $i = 1, \ldots, 38$ with θ_i corresponding to the quantile 0.9 and σ being the shape parameter. The prior distributions are set as $\boldsymbol{\lambda} \sim t_3(4, \mathbf{0}, 100\, \mathbf{I}_3)$ and $\kappa_0 \sim t(4, 0, 100)$. For the Bayesian inference, we ran two MCMC chains with 16000 iterations each. The first 1000 iterations of each chain were discarded as burn-in, and a thinning interval of 10 was applied, resulting in 3000 posterior samples (1500 per chain) used for inference. Convergence was evaluated through visual inspection of the trace plots and the Gelman–Rubin diagnostic ($\hat{R}$) (Gelman & Rubin, 1992), with results indicating satisfactory mixing and stability (plots not shown). Further convergence details are provided in the Appendix (Appendix C).

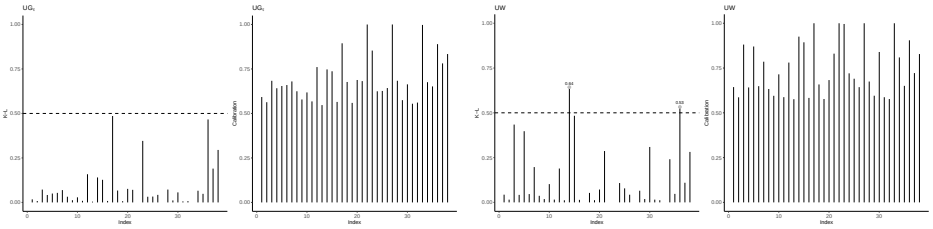
Table 1 reports the posterior estimates of the parameters for the UG_τ and UW models, including mean, median, mode, standard deviation (SD), 95% HPD credibility intervals, as well as model comparison statistics. Regarding the parameters, the location estimates (λ_0) are similar across models, with posterior means of 3.0246 (UG_τ) and 3.2074 (UW). The 95% HPD intervals for λ_0 are relatively narrow, indicating good precision: (2.5975, 3.5477) for UG_τ and (2.7619, 3.7413) for UW. For the coefficients λ_1 and λ_2, both models yield negative values, with

HPD intervals excluding zero, thus providing strong evidence of a negative effect of these covariates on the response. Finally, the shape parameter κ_0 shows larger magnitude (in absolute terms) under the UG_τ model (-1.5146, -0.6192) compared to the UW model (-0.8362, -0.3296), reflecting structural differences between the parameterizations. With respect to model fit, the UG_τ model achieves a lower WAIC value (-63.5043 vs. -61.2317), higher lppd (36.0856 vs. 35.2879), and higher LPML (20.2840 vs. -46.2722) compared to the UW model. These results indicate superior performance of the UG_τ model in terms of both fit and predictive ability. Moreover, the Bayesian p-values (0.4917 for UG_τ and 0.4855 for UW) are close to 0.5, suggesting an overall good adequacy of both models. In summary, while both models provide consistent and adequate estimates, the information criteria (WAIC, lppd, and LPML) favor the UG_τ model, indicating that it offers a better fit to the data along with improved predictive performance when compared to the UW model.

TABLE 1: Posterior parameter estimates (mean, median, and mode), standard deviations (SD), HPD intervals (for posterior mean), effective sample size (ESS), $\hat{R}$ statistics, and model comparison measures for the UG_τ and UW models with $\tau = 0.9$.

UG_τ Model							
Par	Mean	Median	Mode	SD	HPD(95%)	ESS	$\hat{R}$
λ_0	3.0246	3.0057	2.9563	0.2415	(2.5975, 3.5477)	750.9969	1.00
λ_1	-0.0974	-0.0955	-0.0911	0.0274	(-0.1566, -0.0487)	1391.8583	1.01
λ_2	-0.0529	-0.0521	-0.0534	0.0191	(-0.0935, -0.0190)	2670.6666	1.00
κ_0	-1.0875	-1.0987	-1.1305	0.2310	(-1.5146, -0.6192)	844.9697	1.00
WAIC				**-63.5043**			
lppd				**36.0856**			
LPML				**20.2840**			
Bayesian p-value				**0.4917**			
UW Model							
Par	Mean	Median	Mode	SD	HPD(95%)	ESS	$\hat{R}$
λ_0	3.2074	3.1842	3.1454	0.2507	(2.7619, 3.7413)	664.3171	1.00
λ_1	-0.1083	-0.1064	-0.1007	0.0306	(-0.1725, -0.0536)	1207.6732	1.00
λ_2	-0.0571	-0.0554	-0.0545	0.0192	(-0.1016, -0.0251)	2875.8278	1.00
κ_0	-0.5920	-0.5984	-0.6021	0.1280	(-0.8362, -0.3296)	749.3182	1.00
WAIC				-61.2317			
lppd				35.2879			
LPML				-46.2722			
Bayesian p-value				0.4855			

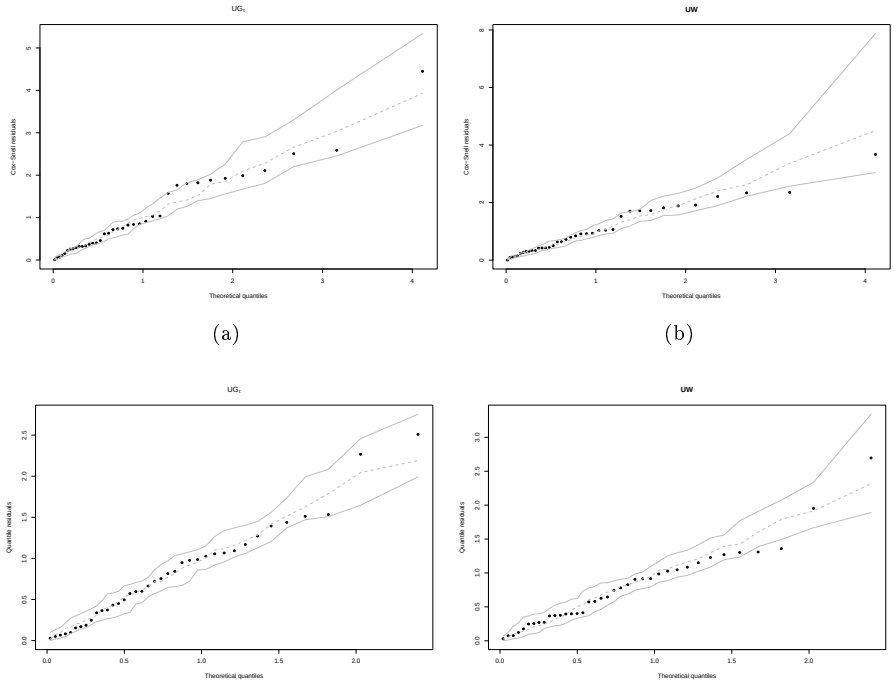
Figure 13 displays the Kullback-Leibler (KL) divergence and calibration plots for the UG_τ and UW models, respectively (see Section 5.4). For the UG_τ model, the KL divergence values remain below 0.5, indicating that the fitted model provides a good approximation to the data-generating process. In contrast, the UW model exhibits relatively higher KL divergence values in several cases, with two of them exceeding 0.5. Specifically, observations #14 and #34 have KL divergence values of 0.64 and 0.53, respectively. The calibration plots further support these findings, showing that the UG_τ model yields lower deviations for some observations compared to the UW model. Overall, these results indicate that UG_τ outperforms UW in terms of both KL divergence and calibration.



(a) (b)

FIGURE 13: The Kullback-Leibler divergence measure and calibration for each observation for UG_τ and UW models.

Figure 14 presents half-normal plots of the Cox–Snell and quantile residuals with simulated envelopes for the UG_τ and UW models. Residuals from the UW model show notable deviations from the reference line, especially in the tails, indicating potential model misspecification. In contrast, the UG_τ model exhibits residuals that closely follow the reference line, with reduced dispersion and no systematic patterns, suggesting a superior fit to the data.



(a) (b)

(c) (d)

FIGURE 14: The half-normal plot with simulated envelope for the Cox–Snell and quantile residuals for the UG_τ and UW models.

Figure 15 displays posterior predictive checks (PPCs) for the UG_τ model. Panel (a) presents boxplots of the observed data y compared to replicated datasets y_{rep} generated from the posterior predictive distribution. The similarity between the distributions indicates that the model is capable of reproducing the overall variability and shape of the observed responses. No systematic discrepancy is apparent, suggesting that the predictive distribution is well aligned with the empirical distribution of the data. Panel (b) applies an interval-based PPC, see Gabry et al. (2019), where each observation y_i is compared with the distribution of its corresponding replicates $y_{\mathrm{rep},i}$. The majority of the observed responses fall well within the outer predictive intervals and are generally close to the medians of their respective predictive distributions, indicating that the model adequately captures both the central tendency and variability of the observations, with no evidence of systematic bias along the sample. Panels (c) and (d) summarize posterior predictive distributions of the sample median and mean, respectively. In both cases, the observed statistics fall well within the posterior predictive distributions, indicating that the model successfully reproduces the main location measures of the data.

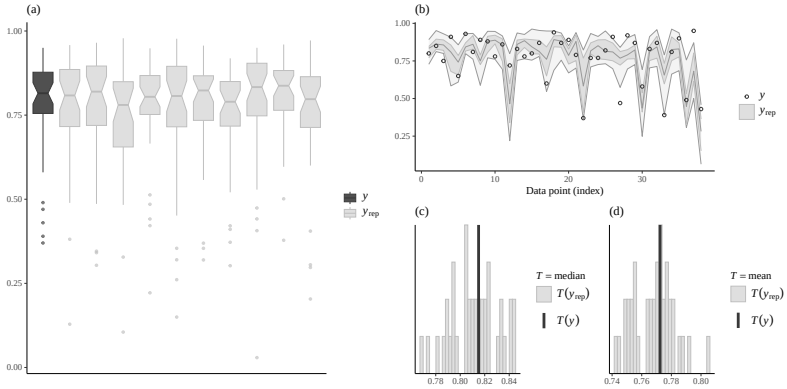


FIGURE 15: Posterior predictive checks (PPCs) for the UG_τ model

8. Concluding Remarks

In this paper we propose a new quantile regression model based on the unit-gamma distribution and additionally consider its parametrization in terms of the mean under a Bayesian framework. Point and interval estimates are obtained from the posterior distributions, and a comprehensive set of influence diagnostics and posterior predictive checks is developed through MCMC sampling. The efficiency of the estimates is evaluated via an extensive simulation study. Furthermore, we analyze data on education and self-reported health from OECD countries to illustrate the flexibility and comparative performance of the UG_τ model. The results indicate that our proposal outperforms the UW model. Thus, our methodology can be a good addition to the tools for statisticians and practitioners interested in the modeling of bounded-support continuous data.

Future research includes deriving a quantile regression framework for proportions with errors-in-variables based on the UG_τ distribution under the recently proposed reparameterization. It would also be of interest to investigate settings with missing covariates. Additional directions include extending the model to accommodate multivariate responses, temporal/spatial structures, as well as developing influence diagnostics within this framework. Furthermore, future work should focus on proposing a zero-and-one-augmented version of the quantile regression model, enabling proper handling of boundary observations (i.e., 0 or 1 outcomes) through additional parameters rather than ad hoc adjustments such as jittering or data transformation. For the mean-reparameterized case, a related development can be found in Silva et al. (2025), who proposed a Shewhart-type control chart based on a UG distribution inflated at zero or one (Silva et al., 2025). These research avenues can be explored under both frequentist and Bayesian paradigms.

Acknowledgements

This paper was developed during the master's research of the first author, carried out under the supervision of the second and third authors, within The Graduation Program in Modeling and Quantitative Methods at the Federal University of Ceará (UFC)- Brazil. We are grateful to the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil (CAPES) – Finance Code 001 for partially supporting the first author. The second and third authors are grateful to the National Council for Scientific and Technological Development (CNPq), Brazil grants 305179/2024-1, 401038/2025-4 and 308058/2022-4, respectively, for partial financial support.

[Received: September 2025 — Accepted: November 2025]

References

Altun, E. (2021), 'The log-Weighted exponential regression model: alternative to the beta regression model', *Communications in Statistics-Theory and Methods* **50**(10), 2306–2321.

Atkinson, A. C. (1985), *Plots, transformation, and regression: an introduction to graphical methods of diagnostic regression analysis*, Clarendon Press, Oxford.

Barrientos, A. F., Jara, A. & Quintana, F. A. (2017), 'Fully nonparametric regression for bounded data using dependent Bernstein polynomials', *Journal of the American Statistical Association* **112**(518), 806–825.

Bayes, C. L., Bazán Guzmán, J. L. & Castro, M. d. (2017), 'A quantile parametric mixed regression model for bounded response variables', *Statistics and its interface* **10**(3), 483–493.

Bayes, C. L., Bazán, J. L. & García, C. (2012), 'A New Robust Regression Model for Proportions', *Bayesian Analysis* **7**(4), 841 – 866.

Beraha, M., Falco, D. & Guglielmi, A. (2021), 'JAGS, NIMBLE, Stan: a detailed comparison among Bayesian MCMC software', arXiv:2107.09357 [stat.CO]. https://arxiv.org/abs/2107.09357

Bernardo, J. M. & Smith, A. F. M. (1994), *Bayesian Theory*, Wiley.

Bourguignon, M. & Gallardo, D. I. (2025), 'A general and unified parameterization of the beta distribution: A flexible and robust beta regression model', *Statistica Neerlandica* **79**(2), 1–19.

Branscum, A. J., Johnson, W. O. & Thurmond, M. C. (2007), 'Bayesian beta regression: applications to household expenditure data and genetic distance between foot-and-mouth disease viruses', *Australian & New Zealand Journal of Statistics* **49**(3), 287–301.

Canterle, D. R. & Bayer, F. M. (2019), 'Variable dispersion beta regressions with parametric link functions', *Statistical Papers* **60**(5), 1541–1567.

Castro, M., Azevedo, C. & Nobre, J. (2024), 'A robust quantile regression for bounded variables based on the Kumaraswamy rectangular distribution', *Statistics and Computing* **34**(2), 74.

Chesneau, C. (2025), 'Introducing a new unit gamma distribution: properties and applications', *European Journal of Statistics* **5**, 1–25.

Cho, H., Ibrahim, J. G., Sinha, D. & Zhu, H. (2009), 'Bayesian case influence diagnostics for survival models', *Biometrics* **65**(1), 116–124.

Christensen, R., Johnson, W., Branscum, A. & Hanson, T. E. (2011), *Bayesian ideas and data analysis: an introduction for scientists and statisticians*, Chapman and Hall, New York.

Cox, D. R. & Snell, E. J. (1968), 'A general definition of residuals', *Journal of the Royal Statistical Society: Series B (Methodological)* **30**, 248–275.

de Valpine, P., Paciorek, C., Turek, D., Michaud, N., Anderson-Bergman, C., Obermeyer, F., Wehrhahn Cortes, C., Rodríguez, A., Temple Lang, D. & Paganin, S. (2024), *NIMBLE user manual*, NIMBLE Development Team. R package manual, version 1.3.0. https://r-nimble.org

de Valpine, P., Turek, D., Paciorek, C., Anderson-Bergman, C., Lang, D. T. & Bodik, R. (2017), 'Programming with models: writing statistical algorithms for general model structures with NIMBLE', *Journal of Computational and Graphical Statistics* **26**, 403–413.

Dunn, P. K. & Smyth, G. K. (1996), 'Randomized quantile residuals', *Journal of Computational and Graphical Statistics* **5**(3), 236–244.

El-Awady, M. M. & Ramadan, A. T. (2025), 'Unit Rayleigh half-normal distribution: Bayesian and Non-Bayesian inference, regression model for bounded response data and application', *Annals of Data Science* pp. 1–34.

Ferrari, S. & Cribari-Neto, F. (2004), 'Beta regression for modelling rates and proportions', *Journal of applied statistics* **31**(7), 799–815.

Figueroa-Zúñiga, J. I., Arellano-Valle, R. B. & Ferrari, S. L. (2013), 'Mixed beta regression: A Bayesian perspective', *Computational Statistics & Data Analysis* **61**, 137–147.

Freitas, J. V. B., Nobre, J. S., Espinheira, P. L. & Rêgo, L. C. (2023), 'Unit gamma regression models for correlated bounded data', *Brazilian Journal of Probability and Statistics* **37**(4), 693–719.

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M. & Gelman, A. (2019), 'Visualization in Bayesian workflow', *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **182**(2), 389–402.

Geisser, S. & Eddy, W. F. (1979), 'A predictive approach to model selection', *Journal of the American Statistical Association* **74**(365), 153–160.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. & Rubin, D. B. (2013), *Bayesian Data Analysis*, 3 edn, CRC Press.

Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B. (1995), *Bayesian data analysis*, Chapman and Hall/CRC.

Gelman, A., Hwang, J. & Vehtari, A. (2014), 'Understanding predictive information criteria for Bayesian models', *Statistics and computing* **24**(6), 997–1016.

Gelman, A., Meng, X.-L. & Stern, H. (1996), 'Posterior predictive assessment of model fitness via realized discrepancies', *Statistica sinica* **6**(4), 733–760.

Gelman, A. & Rubin, D. B. (1992), 'Inference from iterative simulation using multiple sequences', *Statistical Science* **7**(4), 457–472.

Guerra, R. R., Peña-Ramírez, F. A. & Bourguignon, M. (2021), 'The unit extended Weibull families of distributions and its applications', *Journal of Applied Statistics* **48**(16), 3174–3192.

Hahn, E. D. (2008), 'Mixture densities for project management activity times: A robust approach to PERT', *European Journal of operational research* **188**(2), 450–459.

Hamedi-Shahraki, S., Rasekhi, A., Yekaninejad, M. S., Eshraghian, M. R. & Pakpour, A. H. (2021), 'Kumaraswamy regression modeling for bounded outcome scores', *Pakistan Journal of Statistics and Operation Research* **17**(1), 79–88.

Kim, M., Kaplan, B. A., Koffarnus, M. N. & Franck, C. T. (2025), 'Scale-location-truncated beta regression: expanding beta regression to accommodate 0 and 1', *arXiv preprint arXiv:2509.13167* .

Koenker, R. & Bassett Jr, G. (1978), 'Regression quantiles', *Econometrica: journal of the Econometric Society* **46**(1), 33–50.

Korkmaz, M. Ç., Altun, E., Chesneau, C. & Yousof, H. M. (2022), 'On the unit-Chen distribution with associated quantile regression and applications', *Mathematica Slovaca* **72**(3), 765–786.

Korkmaz, M. Ç. & Chesneau, C. (2021), 'On the unit Burr-XII distribution with the quantile regression modeling and applications', *Computational and Applied Mathematics* **40**(1), 29.

Lee, E. T. & Wang, J. W. (2003), *Statistical Methods for Survival Data Analysis*, Wiley Series in Probability and Statistics, 3 edn, Wiley.

López, F. O. (2013), 'A Bayesian approach to parameter estimation in simplex regression model: a comparison with beta regression', *Revista Colombiana de Estadística* **36**(1), 1–21.

Mazucheli, J., Alves, B., Menezes, A. F. & Leiva, V. (2022), 'An overview on parametric quantile regression models and their computational implementation with applications to biomedical problems including COVID-19 data', *Computer Methods and Programs in Biomedicine* **221**, 106816.

Mazucheli, J., Korkmaz, M. Ç., Menezes, A. F. B. & Leiva, V. (2023), 'The unit generalized half-normal quantile regression model: formulation, estimation, diagnostics, and numerical applications', *Soft Computing* **27**(1), 279–295.

Mazucheli, J., Menezes, A., Fernandes, L., De Oliveira, R. & Ghitany, M. (2019), 'The unit-Weibull distribution as an alternative to the Kumaraswamy distribution for the modeling of quantiles conditional on covariates', *Journal of Applied Statistics* **47**(6), 954–974.

Mazucheli, J., Menezes, A., Fernandes, L., De Oliveira, R. & Ghitany, M. (2020), 'The unit-Weibull distribution as an alternative to the Kumaraswamy distribution for the modeling of quantiles conditional on covariates', *Journal of Applied Statistics* **47**(6), 954–974.

McCullagh, P. & Nelder, J. A. (1989), *Generalized Linear Models*, 2 edn, Chapman and Hall, London.

McCulloch, R. E. (1989), 'Local model influence', *Journal of the American Statistical Association* **84**(406), 473–478.

McElreath, R. (2020), *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*, 2 edn, CRC Press.

Meng, X.-L. (1994), 'Posterior predictive p-values', *The Annals of Statistics* **22**(3), 1142–1160.

Merkle, E. C., Furr, D. & Rabe-Hesketh, S. (2019), 'Bayesian comparison of latent variable models: Conditional versus marginal likelihoods', *Psychometrika* **84**(3), 802–829.

Migliorati, S., Di Brisco, A. M. & Ongaro, A. (2018), 'A new regression model for bounded responses', *Bayesian Analysis* **13**(3), 845–872.

Mitnik, P. A. & Baek, S. (2013), 'The Kumaraswamy distribution: median-dispersion re-parameterizations for regression modeling and simulation-based estimation', *Statistical Papers* **54**(1), 177–192.

Mousa, A. M., El-Sheikh, A. A. & Abdel-Fattah, M. A. (2016), 'A gamma regression for bounded continuous variables', *Advances and Applications in Statistics* **49**(4), 305 – 326.

Oliveira, E. S., de Castro, M., Bayes, C. L. & Bazan, J. L. (2022), 'Bayesian quantile regression models for heavy tailed bounded variables using the No-U-Turn sampler', *Computational Statistics* **40**, 3007—-3040.

R Core Team (2024), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ribeiro, V. S., Nobre, J. S., dos Santos, J. R. S. & Azevedo, C. L. (2021), 'Beta rectangular regression models to longitudinal data', *Brazilian Journal of Probability and Statistics* **35**(4), 851–874.

Robert, C. P. (2007), *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*, 2 edn, Springer.

Rocha, E. O., Azevedo, C. L. N., Mota, J. M. A., Batista, M. J. & Nobre, J. S. (2024), 'Bayesian inference for unit gamma distribution', *Caderno Pedagógico* **21**(9), 1–23.

Rubin, D. B. (1984), 'Bayesianly justifiable and relevant frequency calculations for the applied statistician', *The Annals of Statistics* **12**(4), 1151–1172.

Silva, C. R. d., Lima Filho, L. M. d. A., Pereira, T. L. & Duarte Neto, P. J. (2025), 'Shewhart-type control chart based on unit gamma distribution inflated at zero or one', *Journal of Statistical Computation and Simulation* **95**(8), 1887–1908.

Smithson, M. & Verkuilen, J. (2006), 'A better lemon squeezer? maximum-likelihood regression with beta-distributed dependent variables', *Psychological methods* **11**(1), 54–71.

Song, P. X.-K. & Tan, M. (2000), 'Marginal models for longitudinal continuous proportional data', *Biometrics* **56**(2), 496–502.

Vehtari, A., Gelman, A. & Gabry, J. (2017), 'Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC', *Statistics and computing* **27**(5), 1413–1432.

Venezuela, M. K., Aparecida Botter, D. & Carneiro Sandoval, M. (2007), 'Diagnostic techniques in generalized estimating equations', *Journal of Statistical Computation and Simulation* **77**(10), 879–888.

Vidal, I. & Castro, L. M. (2010), 'Influential observations in the independent student-t measurement error model with weak nondifferential error', *Chilean Journal of Statistics* **1**(2), 17–34.

Watanabe, S. (2010), 'Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory', *Journal of machine learning research* **11**(12), 3571–3594.

Zhou, H., Huang, X. & Initiative, A. D. N. (2020), 'Parametric mode regression for bounded responses', *Biometrical Journal* **62**(7), 1791–1809.

Appendix A.

```
 y <- c(80, 85, 75, 91, 65, 93, 81, 89, 88, 78, 86,
       72, 83, 78, 80, 87, 60, 94, 87, 89, 79, 37, 77, 77,
       82, 91, 47, 92, 87, 58, 83, 87, 39, 81, 90, 49, 95,
       43)/100

x1 <- c(4.3, 2.7, 4.8, 3.9, 8.1, 1.8, 2.3, 4, 2.7, 5,
        2, 17.4, 4.8, 2.6, 2.1, 2.6, 8.1, 1.5, 2.4,
        6.8, 3.2, 4.6, 2.1, 4.9, 2.7, 4.3, 6.5, 6.7,
        4, 17.3, 5.7, 1.8, 13, 2.6, 3.8, 4.9, 3.6, 26.5)

x2 <- c(1, 0.4, 1, 1.4, 4.5, 0.8, 0.7, 3.1, 1.4,
        0.6, 0.4, 1, 1.2, 0.9, 0.6, 1.7, 0.8, 0.3, 1.1,
        6.6, 0.6, 17.9, 0.6, 1.3, 0.6, 0.8, 1, 0.8, 0.6,
        0.6, 1, 0.5, 1.7, 0.2, 4.9, 27.6, 11.3, 10)
```

Appendix B.

In this section, we present a reparameterization in terms of quantiles and provide analytical derivations of the moments of UG_τ.

The pdf of the UG distribution with parameters $\alpha > 0$ and $\beta > 0$ is given by

$$f(y \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\, y^{\beta-1}\, [-\log(y)]^{\alpha-1}\, \mathbb{1}_{(0,1)}(y),$$

and its cdf is

$$F(y \mid \alpha, \beta) = 1 - \frac{\gamma(\alpha, \beta(-\log(y)))}{\Gamma(\alpha)},$$

where $\gamma(\alpha, a)$ denotes the lower incomplete gamma function.

We consider the following equation involving the incomplete gamma function, where we set $\theta_\tau = y_\tau$ and $\sigma = \alpha$:

$$1 - \frac{\gamma(\sigma, \beta(-\log(\theta_\tau)))}{\Gamma(\sigma)} = \tau,$$

with $y \in (0,1)$, $\sigma > 0$, and $\tau \in (0,1)$. Here:

- $\gamma(\sigma, x) = \int_0^x t^{\sigma-1} e^{-t}\, dt$ is the lower incomplete gamma function;
- $\Gamma(\sigma)$ is the complete gamma function;
- $H(\sigma, x) = \gamma(\sigma, x)/\Gamma(\sigma)$ is the regularized lower incomplete gamma function.

Rewriting,

$$\frac{\gamma(\sigma, \beta(-\log(\theta_\tau)))}{\Gamma(\sigma)} = 1 - \tau,$$
$$H(\sigma, \beta(-\log(\theta_\tau))) = 1 - \tau.$$

Applying the inverse function H^{-1} yields

$$\beta[-\log(\theta_\tau)] = H^{-1}(\sigma, 1-\tau),$$

and hence

$$\beta = \frac{1}{-\log(\theta_\tau)}\, H^{-1}(\sigma, 1-\tau).$$

Let the random variable Y have pdf

$$f(y \mid \theta_\tau, \sigma) = \frac{\beta_*^\sigma}{\Gamma(\sigma)}\, y^{\beta_*-1}\, [-\log(y)]^{\sigma-1}\, \mathbb{1}_{(0,1)}(y), \qquad \beta_* := \frac{H^{-1}(\sigma, 1-\tau)}{-\log(\theta_\tau)},$$

where $H^{-1}(\sigma, \cdot)$ denotes the inverse (with respect to the second argument) of the regularized lower incomplete gamma function.

For $k > -\beta$, the k-th moment of Y is

$$\begin{aligned}\mathbb{E}[Y^k] &= \int_0^1 y^k f(y \mid \theta_\tau, \sigma)\, dy \\ &= \frac{\beta_*^\sigma}{\Gamma(\sigma)} \int_0^1 y^{\beta_*+k-1}\, [-\log(y)]^{\sigma-1}\, dy.\end{aligned}$$

Using the substitution $t = -\log(y)$ ($y = e^{-t}$, $dy = -e^{-t}dt$), the integral becomes

$$\begin{aligned}\mathbb{E}[Y^k] &= \frac{\beta_*^\sigma}{\Gamma(\sigma)} \int_0^\infty t^{\sigma-1} e^{-(\beta_*+k)t}\, dt \\ &= \left(\frac{\beta_*}{\beta_*+k}\right)^\sigma,\end{aligned}$$

since $\int_0^\infty t^{\sigma-1} e^{-(\beta_*+k)t} dt = \Gamma(\sigma)/(\beta_*+k)^\sigma$. Thus,

$$\mathbb{E}[Y^k] = \left(\frac{\beta_*}{\beta_*+k}\right)^\sigma.$$

For $k = 1$,

$$\mathbb{E}[Y] = \left(\frac{\beta_*}{\beta_*+1}\right)^\sigma = \left(\frac{H^{-1}(\sigma, 1-\tau)}{H^{-1}(\sigma, 1-\tau) - \log(\theta_\tau)}\right)^\sigma.$$

For $k = 2$,

$$\mathbb{E}[Y^2] = \left(\frac{\beta_*}{\beta_* + 2}\right)^{\sigma} = \left(\frac{H^{-1}(\sigma, 1-\tau)}{H^{-1}(\sigma, 1-\tau) - 2\log(\theta_\tau)}\right)^{\sigma}.$$

Finally, the variance is

$$\begin{aligned}\mathrm{Var}(Y) &= \mathbb{E}[Y^2] - \big(\mathbb{E}[Y]\big)^2 \\ &= \left(\frac{\beta_*}{\beta_* + 2}\right)^{\sigma} - \left(\frac{\beta_*}{\beta_* + 1}\right)^{2\sigma} \\ &= \left(\frac{H^{-1}(\sigma, 1-\tau)}{H^{-1}(\sigma, 1-\tau) - 2\log(\theta_\tau)}\right)^{\sigma} - \left(\frac{H^{-1}(\sigma, 1-\tau)}{H^{-1}(\sigma, 1-\tau) - \log(\theta_\tau)}\right)^{2\sigma}.\end{aligned}$$

Remarks

- The existence of the k-th moment requires $H^{-1}(\sigma, 1-\tau) - k\log(\theta_\tau) > 0$.
- In practice, $H^{-1}(\sigma, 1-\tau)$ can be evaluated in R via `qgamma(1 - tau, shape = sigma, scale = 1)`.

Appendix C.

The supplementary materials provide additional results associated with the application of the UG_τ model. These include detailed MCMC diagnostics, such as convergence assessments, autocorrelation structures, and effective sample sizes, ensuring the reliability of the posterior inference. Complementary graphical displays are also presented to facilitate a thorough evaluation of the sampling performance and parameter behavior.

Furthermore, we report the results of the posterior predictive checks, which assess the adequacy of the model in capturing the main features of the observed data. In this analysis, summary statistics computed from the empirical data are compared with their corresponding posterior predictive distributions, thereby highlighting the ability of the UG_τ model to reproduce the salient characteristics of the response variable. Taken together, these supplementary results strengthen the evidence in favor of the proposed model.

Figure A1 presents the Gelman–Rubin convergence statistics. All values of $\widehat{R}$ are below the recommended threshold of 1.05, providing strong evidence of convergence across chains. This result is consistent with the stability observed in the trace plots and reinforces the reliability of the posterior estimates.

Figure A2 presents the trace plots for the parameters κ_0, λ_1, λ_2, and λ_3. The chains display stable oscillations around their respective posterior means with no systematic trends or drifts, indicating convergence to the stationary distribution. The overlap between chains further confirms satisfactory mixing and supports the adequacy of the sampling process.

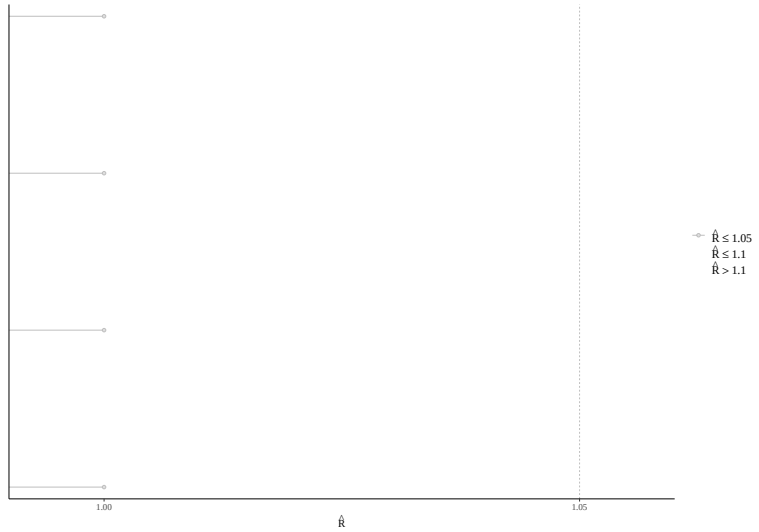


FIGURE A1: Gelman–Rubin convergence diagnostics ($\hat{R}$) for the UG$_\tau$ model parameters.

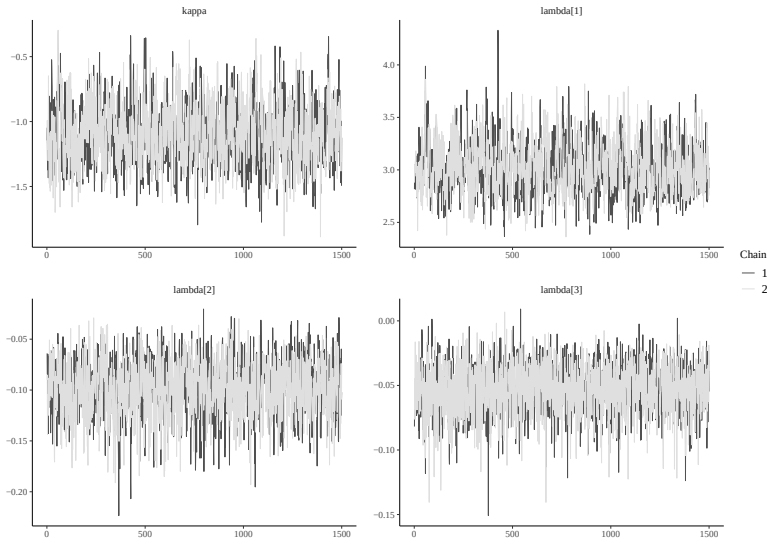


FIGURE A2: Trace plots of the parameters κ_0, λ_1, λ_2, and λ_3 from the UG$_\tau$ model.

Figure A3 depicts the sample autocorrelation functions for the parameters. In all cases, the autocorrelation declines rapidly with increasing lag, suggesting weak serial dependence among posterior draws. This behavior ensures higher effective sample sizes and enhances the overall efficiency of the MCMC procedure.

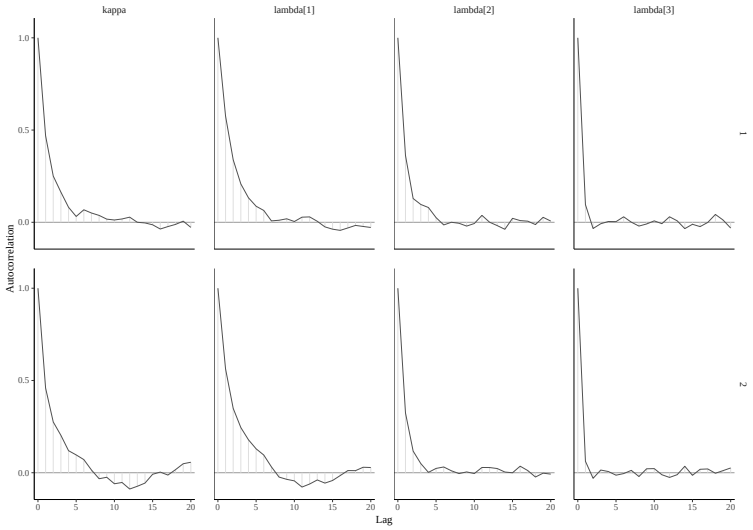


FIGURE A3: Autocorrelation functions of the MCMC samples for the parameters of the UG_τ model.

Figure A4 shows the marginal posterior distributions of the parameters. The histograms reveal unimodal and approximately symmetric distributions, with no evidence of multimodality. These results suggest that the priors and likelihood combined consistently to yield well-behaved posterior summaries.

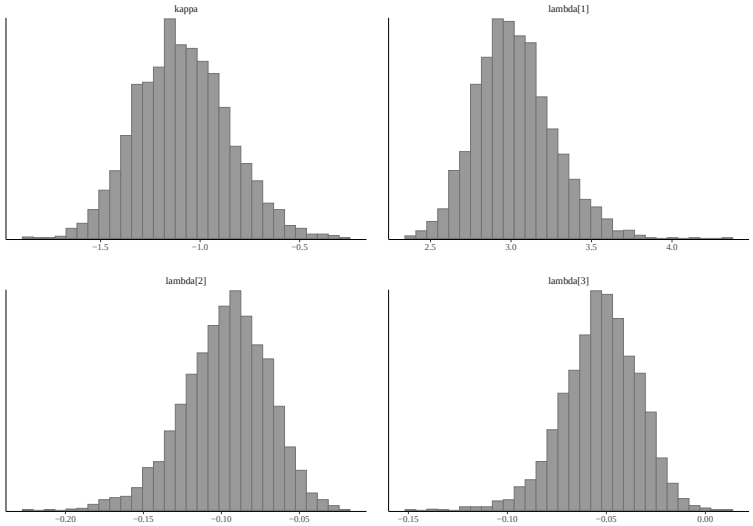


FIGURE A4: Posterior histograms of the parameters of the UG_τ model.

Figure A5 displays the pairwise joint posterior distributions. The scatterplots indicate weak dependencies between parameters, with no signs of strong collinearity problems. The posterior samples appear to explore the parameter space adequately.

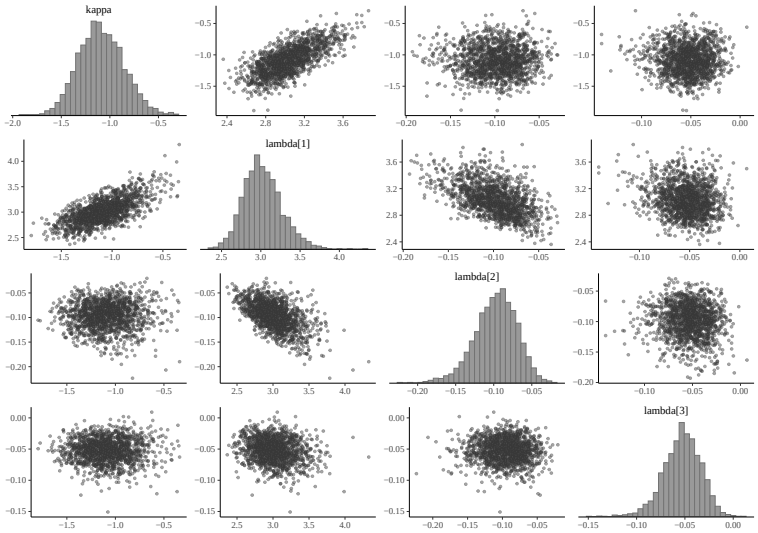


FIGURE A5: Pairs plots of the joint posterior distributions of the UG_{τ} model parameters.