Chilean Journal of Statistics
Volume 16, Number 1, June 2025, Pages 22–35
DOI: 10.32372/ChJS.16-01-02

Statistical Inference and Bayesian Estimation
Research Article

Closed-form solutions for parameter estimation in exponential families based on maximum a posteriori equations

Roberto Vila[1], Helton Saulo[1,2,*], and Eduardo Nakano[1]

[1]Department of Statistics, University of Brasilia, Brasilia, Brazil
[2]Department of Economics, Federal University of Pelotas, Pelotas, Brazil

(Received: 29 May 2025 · Accepted in final form: 29 June 2025)

Abstract

In this article, we derive closed-form estimators for the parameters of certain exponential family distributions using maximum a posteriori (MAP) equations. A Monte Carlo simulation study is conducted to assess the performance of the proposed estimators. The results indicate that their accuracy improves with increasing sample size. Moreover, the proposed estimators exhibit similar performance to that of traditional MAP and maximum likelihood (ML) estimators. A notable advantage of the proposed approach is its computational simplicity, as it avoids the numerical optimization required by MAP and ML estimation. To illustrate the methodology, we analyze a real dataset of 2023 South American gross domestic product per capita values.

Keywords: Bayesian estimation · Income data · Monte Carlo simulation · Point estimation · R software

Mathematics Subject Classification: Primary 62F10 · Secondary 62C10.

1. Introduction

Closed-form estimators are generally computationally efficient, avoiding the convergence issues and high computational costs commonly found in iterative optimization methods. In this article, we clarify that a "closed-form" expression may involve quadratic roots. Several authors have proposed closed-form estimators derived from likelihood-based methods. For instance, in [1], analytical estimators were obtained for the gamma distribution by considering the generalized gamma distribution. Similar methodologies have been employed for the Nakagami distribution [2, 3], the weighted Lindley distribution [4, 5], and for distributions within the exponential family framework [6, 7], which include the previously mentioned cases. For other distributions, some works have used simulation settings and the modeling of complex data structures [8, 9, 10]. Recently, a novel approach was proposed in [11] for deriving closed-form estimators by extending the Box-Cox transformation.

*Corresponding author. Email: heltonsaulo@gmail.com (H. Saulo)

ISSN: 0718-7912 (print)/ISSN 0718-7920 (online)

http://www.soche.cl/chjs

Following this line, we propose closed-form estimators for the parameters of a family of probability distributions belonging to the exponential class. The proposed estimators are based on maximum a posteriori (MAP) equations and use the same conceptual framework introduced in [1, 12] to develop analytical estimators for certain distributions within the exponential family.

The MAP method is a widely used Bayesian approach for deriving point estimates of distribution parameters. Nevertheless, it is uncommon for multi-parameter distributions to yield MAP estimators in closed form. Therefore, the proposed method may serve as an alternative to traditional MAP estimators. In fact, our Monte Carlo results show that the proposed estimators improve with larger samples, with both bias and MSE decreasing, and they match the performance of traditional MAP and maximum likelihood (ML) estimators in the gamma case.

The article is structured as follows. Section 2 presents the proposed estimation methodology, along with relevant theoretical developments and examples of closed-form estimators. In Section 3, we report a Monte Carlo simulation study designed to evaluate the performance of the proposed estimators. In Section 4, the proposed approach is illustrated through an application to South American gross domestic product (GDP) per capita data. Section 5 concludes the article with final considerations.

2. The new estimators

The family of probability distributions belonging to the exponential class, considered in this work, has a probability density function (PDF) defined as

$$f(x;\boldsymbol{\vartheta}) = \frac{(\mu\sigma)^{\mu}}{\Gamma(\mu)}|T'(x)|T^{\mu-1}(x)\exp(-\mu\sigma T(x)), \tag{1}$$

where $x \in (0,\infty)$, $\boldsymbol{\vartheta} = (\mu,\sigma)^{\top}$, with $\mu,\sigma > 0$, and $T\colon (0,\infty) \to (0,\infty)$ is a real, strictly monotonic and at least twice continuously differentiable function, referred to as the generator. Here, $T'(x)$ denotes the derivative of $T(x)$ with respect to x. Note that $T(x)$ may involve other known parameters; see Table A1 in Appendix A.

By applying the transformation $Y = X^{1/p}$ with $p > 0$ and assuming that X follows the PDF stated in (1), the corresponding PDF of Y takes the form given by

$$f(y;\boldsymbol{\vartheta},p) = p\frac{(\mu\sigma)^{\mu}}{\Gamma(\mu)}y^{p-1}|T'(y^p)|T^{\mu-1}(y^p)\exp(-\mu\sigma T(y^p)), \quad y \in (0,\infty), \tag{2}$$

where $\boldsymbol{\vartheta} = (\mu,\sigma)^{\top}$ and $\mu,\sigma,p > 0$. Table A1, presented in Appendix A, includes examples of generator functions $T(x)$ that can be used in the expressions formulated in (1) and (2).

Let $\{Y_i; i = 1,\ldots,n\}$ be a random sample of size n from Y having the PDF defined in (2) and y_i be its observed value. The likelihood function for $(\boldsymbol{\vartheta},p)$ is defined as

$$L(\boldsymbol{\vartheta},p|\boldsymbol{y}) = p^n\frac{(\mu\sigma)^{n\mu}}{\Gamma^n(\mu)}\prod_{i=1}^{n}y_i^{p-1}|T'(y_i^p)|T^{\mu-1}(y_i^p)\exp\left(-\mu\sigma\sum_{i=1}^{n}T(y_i^p)\right). \tag{3}$$

Let $\pi(\boldsymbol{\vartheta})$ be the joint PDF of $\boldsymbol{\vartheta} = (\mu,\sigma)^{\top}$, and let $p \sim \text{Gamma}(\alpha_3,\beta_3)$, for known hyperparameters α_3 and β_3, with PDF $\pi(p)$. Furthermore, suppose that $\boldsymbol{\vartheta}$ is independent of p. Note that, in this notation, we are committing a notational abuse by using μ, σ, and p to denote both the random variables and their values. This notational abuse is common in Bayesian statistics, as it avoids overloading the notation.

The posterior PDF is given by

$$\pi(\boldsymbol{\vartheta}, p|\boldsymbol{y}) = \frac{L(\boldsymbol{\vartheta}, p|\boldsymbol{y})\pi(\boldsymbol{\vartheta})\pi(p)}{\pi(\boldsymbol{y})}, \tag{4}$$

where $\pi(\boldsymbol{y}) = \int_{(0,\infty)^3} L(\boldsymbol{\vartheta}, p|\boldsymbol{y})\pi(\boldsymbol{\vartheta})\pi(p)\mathrm{d}\boldsymbol{\vartheta}\mathrm{d}p$ is the finite predictive distribution. The MAP method estimates $(\boldsymbol{\vartheta}, p)^\top$, with $\boldsymbol{\vartheta} = (\mu, \sigma)^\top$, as the mode of the posterior PDF $\pi(\boldsymbol{\vartheta}, p \mid \boldsymbol{y})$, that is,

$$(\widehat{\boldsymbol{\vartheta}}, \widehat{p})_{\mathrm{MAP}} = \arg\max_{(\boldsymbol{\vartheta}, p)} \pi(\boldsymbol{\vartheta}, p \mid \boldsymbol{y}).$$

Because $\log(\pi(\boldsymbol{\vartheta}, p \mid \boldsymbol{y}))$ is differentiable on the parameter space, a necessary condition for a maximum is obtained as

$$\nabla \log(\pi(\boldsymbol{\vartheta}, p \mid \boldsymbol{y})) = \mathbf{0}, \tag{5}$$

where $\nabla = (\partial/\partial\mu, \partial/\partial\sigma, \partial/\partial p)^\top$ is the gradient vector, and $\mathbf{0}$ denotes the zero vector in $\mathbb{R}^3$. The system stated in (5) is known as the MAP equations, and we now specify the prior distributions required to compute them.

For simplicity we assume independent gamma priors: $\mu \sim \text{Gamma}(\alpha_1, \beta_1)$, $\sigma \sim \text{Gamma}(\alpha_2, \beta_2)$, and $p \sim \text{Gamma}(\alpha_3, \beta_3)$ with $\alpha_3 > 1$.

To streamline the expressions, we revert the transformation $Y_i = X_i^{1/p}$ so that $X_i = Y_i^p$. All derivatives are now written in terms of X_i.

Taking logarithm in the expressions given in (4) and substituting the formulation stated in (3) yield

$$\begin{aligned}
\log(\pi(\boldsymbol{\vartheta}, p \mid \boldsymbol{y})) &= \log(L(\boldsymbol{\vartheta}, p \mid \boldsymbol{y})) + \log(\pi(\mu)) + \log(\pi(\sigma) = + \log(\pi(p)) - \log(\pi(\boldsymbol{y})) \\
&= n\log(p) + n\mu(\log(\mu) + \log(\sigma)) - n\log(\Gamma(\mu)) - \log(\pi(\boldsymbol{y})) \\
&\quad + \sum_{i=1}^{n} \log(|T'(y_i^p)|) + (p-1)\sum_{i=1}^{n}\log(y_i) - \mu\sigma\sum_{i=1}^{n} T(y_i^p) + (\alpha_1 - 1)\log(\mu) \\
&\quad + (\mu - 1)\sum_{i=1}^{n}\log(T(y_i^p)) - \beta_1\mu + (\alpha_2 - 1)\log(\sigma) - \beta_2\sigma + (\alpha_3 - 1)\log(p) \\
&\quad - \beta_3 p + \alpha_1\log(\beta_1) - \log(\Gamma(\alpha_1)) + \alpha_2\log(\beta_2) - \log(\Gamma(\alpha_2)) + \alpha_3\log(\beta_3) \\
&\quad - \log(\Gamma(\alpha_3)).
\end{aligned}$$

To obtain the MAP equations defined in (5), we now compute the gradient of $\log(\pi(\boldsymbol{\vartheta}, p \mid \boldsymbol{y}))$ with respect to (μ, σ, p), after re-parameterizing $Y_i = X_i^{1/p}$, obtaining that

$$\begin{aligned}
\frac{\partial \log(\pi(\boldsymbol{\vartheta}, p \mid \boldsymbol{y}))}{\partial\mu} &= n\log(\mu) + n\log(\sigma) + n - n\psi^{(0)}(\mu) - \sigma\sum_{i=1}^{n} T(X_i) + \sum_{i=1}^{n}\log(T(X_i)) \\
&\quad + \frac{\alpha_1 - 1}{\mu} - \beta_1 = 0,
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial \log(\pi(\boldsymbol{\vartheta}, p \mid \boldsymbol{y}))}{\partial \sigma} &= \frac{n\mu}{\sigma} - \mu \sum_{i=1}^{n} T(X_i) + \frac{\alpha_2 - 1}{\sigma} - \beta_2 = 0, \\
\frac{\partial \log(\pi(\boldsymbol{\vartheta}, p \mid \boldsymbol{y}))}{\partial p} &= \Bigg(n + \sum_{i=1}^{n} \frac{T''(X_i)}{T'(X_i)} X_i \log(X_i) + \sum_{i=1}^{n} \log(X_i) \\
&\quad - \mu\sigma \sum_{i=1}^{n} T'(X_i) X_i \log(X_i) + (\mu - 1) \sum_{i=1}^{n} \frac{T'(X_i)}{T(X_i)} X_i \log(X_i) + \alpha_3 - 1 \Bigg) \frac{1}{p} \\
&\quad - \beta_3 = 0,
\end{aligned}
$$

Since $Y_i = X_i^{1/p}$ for every $i = 1, \ldots, n$, the derivatives can be written in terms of X_i.

Now, suppose $p \sim \text{Gamma}(\alpha_3, \beta_3)$. Then, we have $\mathbb{E}(1/p) = \beta_3/(\alpha_3 - 1)$, for $\alpha_3 > 1$. Taking the expectation in the score equation with respect to p yields

$$
\begin{aligned}
&\Bigg(n + \sum_{i=1}^{n} \frac{T''(X_i)}{T'(X_i)} X_i \log(X_i) + \sum_{i=1}^{n} \log(X_i) \\
&\quad - \mu\sigma \sum_{i=1}^{n} T'(X_i) X_i \log(X_i) + (\mu - 1) \sum_{i=1}^{n} \frac{T'(X_i)}{T(X_i)} X_i \log(X_i) + \alpha_3 - 1 \Bigg) \frac{\beta_3}{\alpha_3 - 1} - \beta_3 = 0.
\end{aligned}
$$

Hence, the bracketed term equals $\alpha_3 - 1$, that is,

$$
\begin{aligned}
&n + \sum_{i=1}^{n} \frac{T''(X_i)}{T'(X_i)} X_i \log(X_i) + \sum_{i=1}^{n} \log(X_i) - \mu\sigma \sum_{i=1}^{n} T'(X_i) X_i \log(X)_i \\
&+ (\mu - 1) \sum_{i=1}^{n} \frac{T'(X_i)}{T(X_i)} X_i \log(X)_i = 0.
\end{aligned}
$$

Now, set $\alpha_3 = \beta_3 = k$ and let $k \to \infty$. By dividing the last equation by n yields

$$
\begin{aligned}
&1 + \frac{1}{n} \sum_{i=1}^{n} \frac{T''(X_i)}{T'(X_i)} X_i \log(X_i) + \frac{1}{n} \sum_{i=1}^{n} \log(X_i) \\
&- \mu\sigma \frac{1}{n} \sum_{i=1}^{n} T'(X_i) X_i \log(X_i) + (\mu - 1) \frac{1}{n} \sum_{i=1}^{n} \frac{T'(X_i)}{T(X_i)} X_i \log(X_i) = 0.
\end{aligned}
$$

We can express μ as a function of σ as

$$
\mu(\sigma) = \frac{1 + \overline{X}_2}{\sigma \overline{X}_4 - \overline{X}_3}, \tag{6}
$$

with

$$
\begin{aligned}
\overline{X}_2 &= \frac{1}{n} \sum_{i=1}^{n} h_2(X_i), \qquad & h_2(x) &= \log(x) + \left(\tfrac{T''(x)}{T'(x)} - \tfrac{T'(x)}{T(x)} \right) x \log(x), \\
\overline{X}_3 &= \frac{1}{n} \sum_{i=1}^{n} h_3(X_i), \qquad & h_3(x) &= \frac{T'(x)}{T(x)}\, x \log(x),
\end{aligned}
$$

$$\overline{X}_4 = \frac{1}{n}\sum_{i=1}^{n} h_4(X_i), \qquad\qquad h_4(x) = T'(x)\, x \log(x). \tag{7}$$

Such formulations for h_2, h_3, and h_4 are used to obtain a closed-form estimator for σ.

Substituting the expression given in (6) into the score equation $\partial \log \pi / \partial \sigma = 0$ yields the quadratic equation $A\sigma^2 + B\sigma + C = 0$, whose positive root gives the closed-form estimator defined as

$$\widehat{\sigma} = \frac{\frac{1}{n}\overline{X}_5 - (1+\overline{X}_2)\,\overline{X}_1 + \sqrt{\left(\frac{1}{n}\overline{X}_5 - (1+\overline{X}_2)\,\overline{X}_1\right)^2 - 4(\beta_2/n)\,\overline{X}_4\left(\left(\frac{\alpha_2-1}{n}\right)\overline{X}_3 - (1+\overline{X}_2)\right)}}{2(\beta_2/n)\,\overline{X}_4}, \tag{8}$$

provided that the discriminant inside the square root is non-negative and $\overline{X}_4 \neq 0$.

In the formulation stated in (8), we get

$$\overline{X}_1 = \frac{1}{n}\sum_{i=1}^{n} h_1(X_i), \qquad\qquad h_1(x) = T(x),$$

$$\overline{X}_5 = \beta_2\,\overline{X}_3 + (\alpha_2-1)\,\overline{X}_4 = \frac{1}{n}\sum_{i=1}^{n} h_5(X_i), \qquad h_5(x) = \beta_2\, h_3(x) + (\alpha_2-1)\, h_4(x), \tag{9}$$

with $\overline{X}_3, \overline{X}_4$ being defined in (7).

Now, substituting the expression presented in (8) back into the formulation given in (6) gives

$$\widehat{\mu} = \frac{1+\overline{X}_2}{\widehat{\sigma}\,\overline{X}_4 - \overline{X}_3}. \tag{10}$$

Remark 1 The estimator $\widehat{\sigma}$ depends only on the hyper-parameters α_2, β_2, whereas $\widehat{\mu}$ depends on α_2, β_2 only indirectly through $\widehat{\sigma}$. Both estimators are free of α_1, β_1.

Remark 2 Let $p \sim \text{Gamma}(\alpha_3, \beta_3)$ with $\alpha_3 = \beta_3 = k > 1$. Its cumulative distribution function is

$$F_p(v) = \int_0^v \frac{k^k}{\Gamma(k)} u^{k-1} \exp(-ku)\mathrm{d}u, \quad v > 0.$$

For every fixed $v > 0$ we have, as $k \to \infty$,

$$F_p(v) \;\longrightarrow\; \mathbb{1}_{[1,\infty)}(v),$$

that is, the distribution of p converges weakly to a Dirac mass at 1. Hence, we have that

$$p \xrightarrow[k\to\infty]{\mathbb{D}} 1 \quad\Longrightarrow\quad p \xrightarrow[k\to\infty]{\mathbb{P}} 1,$$

where $\xrightarrow{\mathbb{D}}$ and $\xrightarrow{\mathbb{P}}$ denote convergence in distribution and in probability, respectively. We slightly abuse notation by writing "1" for the degenerate random variable that equals 1 with probability one.

A constructive way to see the same limit is to express p as the average of k independent and identically distributed (IID) exponential variables with mean 1 is given by

$$p = \frac{1}{k}\sum_{i=1}^{k} E_i, \quad E_i \overset{\text{IID}}{\sim} \text{Exponential}(1).$$

By the strong law of large numbers,

$$\frac{1}{k}\sum_{i=1}^{k} E_i \xrightarrow[k\to\infty]{\text{almost surely}} \mathbb{E}(E_1) = 1,$$

with $\overset{\text{almost surely}}{\longrightarrow}$ denoting almost sure convergence. Thus, $p \to 1$ almost surely, which implies the previous modes of convergence. Moreover, the event $\{\lim_{k\to\infty}(p, X)^\top = (1, X)^\top\}$ occurs with probability one. Hence, we have

$$(p, X)^\top \xrightarrow[k\to\infty]{\text{almost surely}} (1, X)^\top.$$

Almost-sure convergence is preserved under continuous functions; hence, for any bivariate continuous function g on $(0,\infty)^2$,

$$g(p, X) \xrightarrow[k\to\infty]{\text{almost surely}} g(1, X).$$

Choosing $g(u, v) = v^{1/u}(u, v > 0)$ gives

$$Y = X^{1/p} \xrightarrow[k\to\infty]{\text{almost surely}} X.$$

Consequently, as $\alpha_3 = \beta_3 = k \to \infty$ we have $p \approx 1$ and $Y = X^{1/p} \approx X$ almost surely. In simple terms, collapsing to $p = 1$ is justified because the extended model $Y = X^{1/p}$ reduces to the original model X. Estimators obtained from the MAP equations with $p \approx 1$ are therefore approximations to the full MAP solution for X.

Proposition 1 If $T(x) = x^{-s}$, for $s \neq 0$, then, the closed-form estimators for σ and μ are given by

$$\widehat{\sigma} = \frac{\frac{1}{n}\overline{X}_5 - \overline{X}_1 + \sqrt{\left(\frac{1}{n}\overline{X}_5 - \overline{X}_1\right)^2 - 4\frac{\beta_2}{n}\overline{X}_4\left(\left(\frac{\alpha_2 - 1}{n}\right)\overline{X}_3 - 1\right)}}{2\frac{\beta_2}{n}\overline{X}_4}, \tag{11}$$

provided

$$\left(\frac{1}{n}\overline{X}_5 - \overline{X}_1\right)^2 - 4\frac{\beta_2}{n}\overline{X}_4\left(\left(\frac{\alpha_2 - 1}{n}\right)\overline{X}_3 - 1\right) \geqslant 0, \tag{12}$$

and

$$\widehat{\mu} = \frac{1}{\widehat{\sigma}\overline{X}_4 - \overline{X}_3}, \tag{13}$$

respectively, where

$$\overline{X}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i^{-s}, \quad \overline{X}_3 = \frac{1}{n}\sum_{i=1}^{n} \log(X_i^{-s}), \quad \overline{X}_4 = \frac{1}{n}\sum_{i=1}^{n} X_i^{-s}\log(X_i^{-s}),$$

$$\overline{X}_5 = \beta_2\left(\frac{1}{n}\sum_{i=1}^{n}\log(X_i^{-s})\right) + (\alpha_2 - 1)\left(\frac{1}{n}\sum_{i=1}^{n} X_i^{-s}\log(X_i^{-s})\right).$$

The formulas stated in (13) and (11) are valid provided that $\overline{X}_4 \neq 0$, that $\widehat{\sigma}\,\overline{X}_4 - \overline{X}_3 \neq 0$, and that the discriminant condition stated in (12) holds. For continuous data these requirements are satisfied with probability one.

Proof [Proposition 1] If $T(x) = x^{-s}$ with $s \neq 0$, then $\overline{X}_2 = 0$. Substituting this into the expressions given in (8) and (10) reduces them exactly to the formulations given in (13) and (11), respectively. ■

Remark 3 When $\mu = \mu_0$ is constant, the MAP equations yield

$$\widehat{\sigma} = \frac{\frac{1}{\mu_0}\left(1 + \overline{X}_2\right) + \overline{X}_3}{\overline{X}_4}.$$

See Table A1 (Appendix A) for distributions that satisfy $\mu = \mu_0$, por example, Maxwell–Boltzmann, Rayleigh, Weibull, inverse Weibull (Fréchet), Gompertz, traditional Weibull, flexible Weibull, Burr XII (Singh–Maddala), and Dagum (Mielke beta-kappa).

Likewise, when $\sigma = \sigma_0$ is constant,

$$\widehat{\mu} = \frac{1 + \overline{X}_2}{\sigma_0\overline{X}_4 - \overline{X}_3}.$$

The only distribution in Table A1 that meets $\sigma = \sigma_0$ is the Modified Weibull extension.

Remark 4 When σ is a function of μ, that is, $\sigma = g(\mu)$ for a given function g, insert $g(\mu)$ into the relationship presented as

$$\mu = \frac{1 + \overline{X}_2}{\sigma\overline{X}_4 - \overline{X}_3}.$$

This gives the nonlinear equation stated as

$$\mu g(\mu)\overline{X}_4 - \mu\overline{X}_3 = 1 + \overline{X}_2 \quad \Longleftrightarrow \quad \mu g(\mu) = \frac{1 + \overline{X}_2 + \mu\overline{X}_3}{\overline{X}_4}.$$

Solving this equation for μ yields $\widehat{\mu}$, after which $\widehat{\sigma} = g(\widehat{\mu})$. Examples in Table A1 that satisfy $\sigma = g(\mu)$ are the δ-gamma distribution $\big(g(x) = 1/(\delta x), \delta \text{ known}\big)$ and the chi-squared distribution $\big(g(x) = 1/(2x)\big)$.

3. Illustrative Monte Carlo simulation study

We perform two Monte Carlo simulation studies. The first assesses the performance of the proposed closed-form estimators for μ and σ under the gamma, inverse gamma, Weibull, and inverse Weibull distributions. The second focuses on the gamma distribution, comparing the proposed estimators with traditional MAP estimators.

To evaluate the performance of the proposed and traditional MAP estimators, we compute the mean absolute relative error (MARE) —also known as the absolute (unsigned) relative bias— and the mean squared error (MSE), defined respectively as

$$\text{MARE}(\widehat{\theta}) = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\widehat{\theta}^{(i)} - \theta}{\theta}\right|, \quad \text{MSE}(\widehat{\theta}) = \frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\theta}^{(i)} - \theta\right)^2, \tag{14}$$

where $\theta \in \{\mu, \sigma\}$ denotes the true parameter value and $\widehat{\theta}^{(i)} \in \{\widehat{\mu}^{(i)}, \widehat{\sigma}^{(i)}\}$ is the ith Monte Carlo estimate. The number of replications is $N = 10{,}000$.

Throughout all simulations, we use noninformative priors by setting the hyperparameters $\alpha_1 = \beta_1 = \alpha_2 = \beta_2 = 1/100$, which correspond to gamma priors with large variances. We consider the following sample sizes: $n \in \{15, 30, 60, 120, 240, 480, 760\}$. We set $\mu = 1$, $\sigma = 2$, and $\delta = 1.5$. All computational analyses were performed in R (www.R-project.org), and the complete source code are publicly available on GitHub at github.com/heltonsaulo/mapEstimators. We carried out simulations to assess the performance of the proposed closed-form estimators of σ and μ under the gamma, inverse gamma, Weibull, and inverse Weibull distributions.

Note that, when the condition stated in (12) is satisfied, the estimators of σ and μ based on these distributions are given by the expressions presented in Proposition 1 and formulated as

$$\widehat{\sigma} = \frac{\frac{1}{n}\overline{X}_5 - \overline{X}_1 + \sqrt{\left(\frac{1}{n}\overline{X}_5 - \overline{X}_1\right)^2 - 4\frac{\beta_2}{n}\overline{X}_4\left(\left(\frac{\alpha_2 - 1}{n}\right)\overline{X}_3 - 1\right)}}{\frac{2\beta_2}{n}\overline{X}_4} \tag{15}$$

and

$$\widehat{\mu} = \frac{1}{\widehat{\sigma}\overline{X}_4 - \overline{X}_3}, \tag{16}$$

respectively, with

$$\begin{aligned} &\overline{X}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i^{-s}, \quad \overline{X}_3 = \frac{1}{n}\sum_{i=1}^{n}\log(X_i^{-s}), \quad \overline{X}_4 = \frac{1}{n}\sum_{i=1}^{n} X_i^{-s}\log(X_i^{-s}), \\ &\overline{X}_5 = \beta_2\left(\frac{1}{n}\sum_{i=1}^{n}\log(X_i^{-s})\right) + (\alpha_2 - 1)\left(\frac{1}{n}\sum_{i=1}^{n} X_i^{-s}\log(X_i^{-s})\right), \end{aligned} \tag{17}$$

where the choice of s dictates the appropriate estimator for the parameters of the four distributions (see Table 1). Moreover, in the simulations conducted, the condition stated in (12) was satisfied for all Monte Carlo samples under the selected parameter values.

Figure 1 shows the Monte Carlo results computed according to the expressions given in (14). From this figure, we observe that, as expected, both the relative bias and MSE decrease as the sample size increases, for all distributions and parameters. In particular, the relative biases for σ remain close to zero even for small samples.

Table 1: Some values of s to be used in the expressions given in (15), (16), and (17).

Distribution	Gamma	Inverse gamma	Weibull	Inverse Weibull
Value of s	-1	1	$-\delta$	δ

We compare the proposed closed-form estimators with the traditional MAP and ML estimators in the gamma case. Figure 2 shows the relative bias and MSE of all three estimators for μ and σ across different sample sizes. As expected, the performance of all estimators improves as n increases. The MAP estimators exhibit slightly lower bias for μ, whereas the ML estimator for σ shows noticeably higher bias compared to both the proposed and MAP estimators. In terms of MSE, all three approaches yield similar performance.

Overall, the methods provide comparable results, with the proposed estimators having the advantage of computational simplicity, as they avoid numerical optimization.

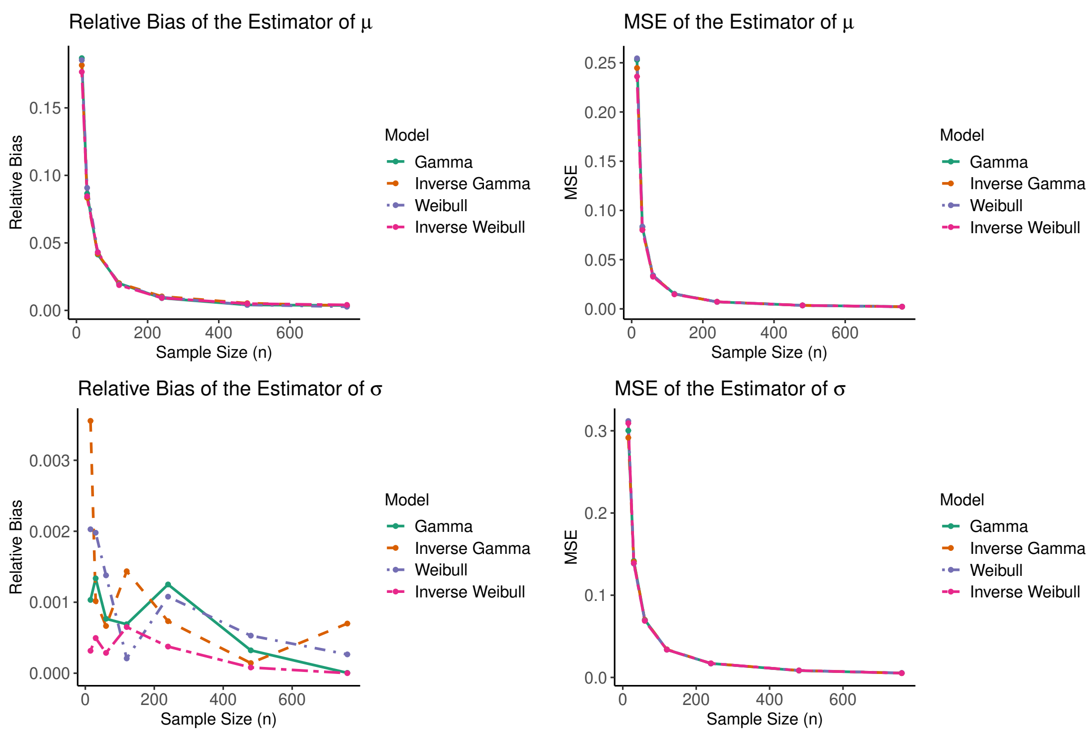

Figure 1: Monte Carlo results for the estimators $\widehat{\mu}$ and $\widehat{\sigma}$ for the indicated models.

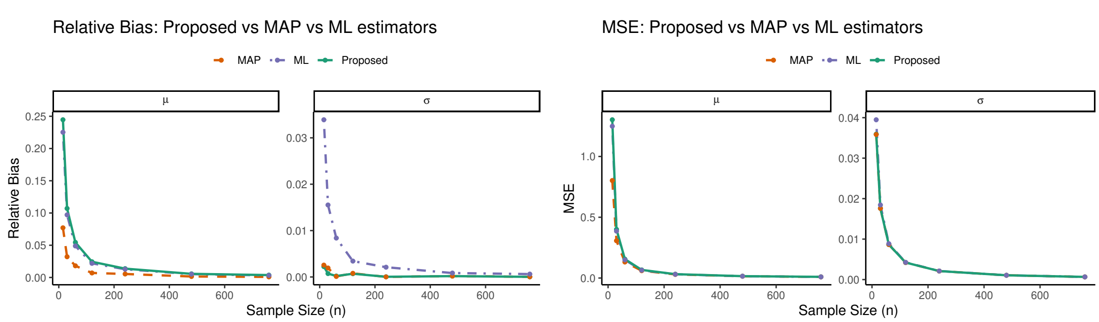

Figure 2: Monte Carlo results comparing the bias and MSE of the proposed, MAP, and ML estimators for the gamma distribution parameters.

4. Application to real data

In this section, we illustrate the proposed estimation method using a dataset on South American GDP per capita for 2023, expressed in international dollars at 2021 prices. All monetary values reported below are expressed in thousands of international dollars. The data, available at ourworldindata.org/grapher/gdp-per-capita-worldbank, exhibit substantial variation across countries. In 2023, GDP per capita ranged from \$9.844 in Bolivia to \$49.315 in Guyana. The countries, listed in descending order of GDP per capita, are: Guyana (\$49.315), Uruguay (\$31.019), Chile (\$29.463), Argentina (\$27.105), Suriname (\$19.044), Brazil (\$19.018), Colombia (\$18.692), Paraguay (\$15.783), Peru (\$15.294), Ecuador (\$14.472), and Bolivia (\$9.844).

The gamma distribution is commonly used to model positive individual-level quantities such as income; see, for example, [13]. Hence, we assume a gamma model for the South American GDP data and compare the proposed closed-form estimators with traditional MAP and ML estimators.

Table 2 reports the estimates of μ and σ. Figure 3 displays goodness-of-fit diagnostics (empirical quantile versus theoretical quantile (QQ) plot and empirical versus fitted CDF) for the three estimators. All of them yield similar and satisfactory fits.

Table 2: Estimates of μ and σ for the gamma distribution obtained with different methods.

Method	$\hat{\mu}$	$\hat{\sigma}$
Proposed	6.7217	0.0436
MAP	4.4965	0.0433
ML	5.4710	0.0442

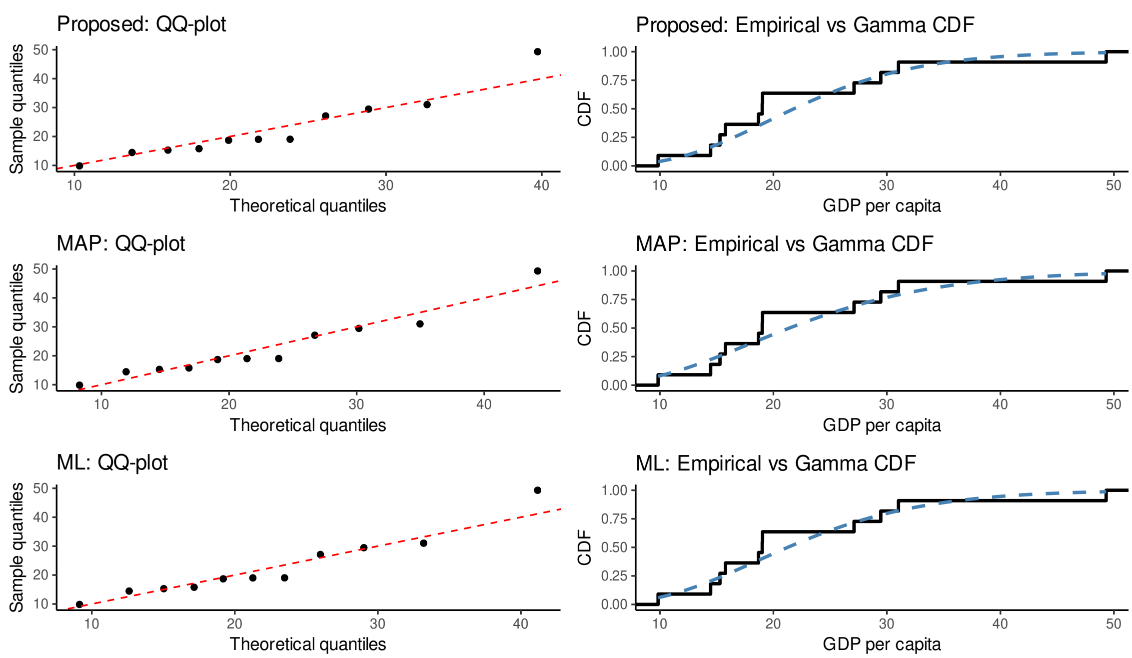

Figure 3: Goodness-of-fit diagnostics —QQ plot and CDF— for the gamma fits based on the proposed, MAP, and ML estimators.

We now apply the Kolmogorov–Smirnov (KS) and Cramér–von Mises (CvM) goodness-of-fit tests; Table 3 presents the results. Because the usual critical values of these tests tend to be liberal for small samples, we report bootstrap p-values.

From Table 3 we observe that the bootstrap p-values for both tests are well above the conventional 5% threshold for all three estimators, providing no evidence against the gamma assumption. The ML fit attains the largest bootstrap p-values (KS = 0.234, p_{boot} = 0.655; CvM = 0.067, p_{boot} = 0.916), indicating the closest agreement with the theoretical CDF.

Overall, the proposed, MAP, and ML approaches all provide an adequate gamma fit to the 2023 South American GDP per capita data.

Table 3: Goodness-of-fit statistics and bootstrap p-values for the gamma fits.

Method	KS statistic	KS p_{boot}	CvM statistic	CvM p_{boot}
Proposed	0.267	0.553	0.095	0.861
MAP	0.223	0.575	0.069	0.853
ML	0.234	0.655	0.067	0.916

5. Conclusions

In this article, we introduced closed-form estimators for a flexible exponential family derived from maximum a posteriori equations. The main advantage of the proposed methodology lies in the elimination of numerical optimization. This feature makes the estimators particularly attractive in applications where computational simplicity is essential.

Monte Carlo simulations revealed that, as expected, the performance of the proposed estimators improves with increasing sample size. In the gamma case, the proposed estimators demonstrated performance comparable to that of the traditional maximum a posteriori and maximum likelihood estimators. In addition, the real-data application to South American GDP per capita in 2023 showed that the proposed estimators provide an adequate fit.

Future research should focus on exploring the proposed methodology to multivariate distributions and exploring additional distribution families. Furthermore, bias-reduction techniques can be considered to reduce the bias of the proposed estimators [14, 15]. These investigations are currently underway, and we expect to report the results in future work.

Appendix A. Examples of generators T

Table A1: Some forms of generators $T(x)$ to be used in the expression given in (1).

Distribution	μ	σ	$T(x)$
Burr type XII (Singh-Maddala) [16]	1	k	$\log(x^c+1)$
Chi-squared [17]	$\frac{\nu}{2}$	$\frac{1}{\nu}$	x
Dagum (Mielke beta-kappa) [18]	1	k	$\log\left(\frac{1}{x^c}+1\right)$
Flexible Weibull [19]	1	a	$\exp\left(bx-\frac{c}{x}\right)$
Gamma [20]	α	$\frac{1}{\alpha\beta}$	x
Generalized gamma [20]	$\frac{\alpha}{\delta}$	$\frac{\delta}{\alpha\beta^\delta}$	x^δ
Generalized inverse gamma [21]	α	$\frac{\beta^\delta}{\alpha}$	$\frac{1}{x^\delta}$
Gompertz [22]	1	α	$\exp(\delta x)-1$
Inverse gamma [23]	α	$\frac{\beta}{\alpha}$	$\frac{1}{x}$
Inverse Weibull (Fréchet) [24]	1	β^δ	$\frac{1}{x^\delta}$

Continued on next page

Table A1 – Continued from previous page

Distribution	μ	σ	$T(x)$
Maxwell-Boltzmann [25]	$\frac{3}{2}$	$\frac{1}{3\beta^2}$	x^2
Modified Weibull extension [26]	$\lambda\alpha$	1	$\exp\left(\left(\frac{x}{\alpha}\right)^{\beta}\right)-1$
Nakagami [27]	m	$\frac{1}{\Omega}$	x^2
New exponentiated generalized gamma	$\frac{\alpha}{\delta}$	$\frac{\delta}{\alpha\beta^\delta}$	$\log^{\delta}(x+1)$
New exponentiated generalized inverse gamma	$\frac{\alpha}{\delta}$	$\frac{\delta}{\alpha\beta^\delta}$	$\log^{\delta}\left(\frac{1}{x}+1\right)$
New extended log-generalized gamma	$\frac{\alpha}{\delta}$	$\frac{\delta}{\alpha\beta^\delta}$	$x^{\delta}(\exp(x)-1)^{\delta}$
New log-generalized gamma	$\frac{\alpha}{\delta}$	$\frac{\delta}{\alpha\beta^\delta}$	$(\exp(x)-1)^{\delta}$
New log-generalized inverse gamma	$\frac{\alpha}{\delta}$	$\frac{\delta}{\alpha\beta^\delta}$	$(\exp(\frac{1}{x})-1)^{\delta}$
New modified log-generalized gamma	$\frac{\alpha}{\delta}$	$\frac{\delta}{\alpha\beta^\delta}$	$\exp^{\delta}\left(x-\frac{1}{x}\right)$
Rayleigh [28]	1	$\frac{1}{2\beta^2}$	x^2
Scaled inverse chi-squared [29]	$\frac{\nu}{2}$	τ^2	$\frac{1}{x}$
Traditional Weibull [30]	1	a	$x^b(\exp(cx^d)-1)$
Weibull [17]	1	$\frac{1}{\beta^\delta}$	x^{δ}
δ-gamma [31]	$\frac{\beta}{\delta}$	$\frac{1}{\beta}$	x^{δ}

In Table A1, we are assuming that $m \geqslant 1/2$ and $\Omega, a, b, c, d, k, \alpha, \delta, \lambda, \beta, \nu, \tau^2 > 0$.

Statements

Acknowledgements

The authors would like to thank the Editors and the two anonymous reviewers for their valuable comments and suggestions, which improved the quality of this article.

Author contributions

Conceptualization: R. Vila, H. Saulo, E. Nakano; data curation: R. Vila, H. Saulo, E. Nakano; formal analysis: R. Vila, H. Saulo, E. Nakano; investigation: R. Vila, H. Saulo, E. Nakano; methodology: R. Vila, H. Saulo, E. Nakano; writing original draft: R. Vila, H. Saulo, E. Nakano; writing review and editing: R. Vila, H. Saulo, E. Nakano. All authors have read and agreed to the published version of the article.

Conflicts of interest

The authors declare no conflict of interest.

Data and code availability

The data and complete source code is publicly available on GitHub at github.com/heltonsaulo/mapEstimators.

Declaration on the use of artificial intelligence (AI) technologies

The authors declare that no generative AI was used in the preparation of this article.

Funding

This study was partly funded by the Brazilian National Council for Scientific and Technological Development (CNPq; grants No. 309674/2020-4 and No. 304716/2023-5 to H. Saulo).

References

[1] Ye, Z.S. and Chen, N., 2017. Closed-form estimators for the gamma distribution derived from likelihood equations. The American Statistician, 71, 177–181. DOI: 10.1080/00031305.2016.1209129

[2] Ramos, P.L., Louzada, F., and Ramos, E., 2016. An efficient, closed-form MAP estimator for Nakagami-m fading parameter. IEEE Communications Letters, 20, 2328–2331. DOI: 10.1109/LCOMM.2016.2603530

[3] Zhao, J., Kim, S., and Kim, H.M., 2021. Closed-form estimators and bias-corrected estimators for the Nakagami distribution. Mathematics and Computers in Simulation, 185, 308–324. DOI: 10.1016/j.matcom.2020.12.026

[4] Kim, H.M. and Jang, Y.H., 2020. New closed-form estimators for weighted Lindley distribution. Journal of the Korean Statistical Society, 50, 580–606. DOI: 10.1007/s42952-020-00097-y

[5] Mazucheli, J., Coelho-Barros, E.A., and Louzada, F., 2016. On the hypothesis testing for the weighted Lindley distribution. Chilean Journal of Statistics, 7, 17–27. URL: soche.cl/chjs/volumes/07/02/Mazucheli_etal(2016).pdf

[6] Vila, R., Nakano, E., and Saulo, H., 2024a. Closed-form estimators for an exponential family derived from likelihood equations. Preprint on arXiv. DOI: 10.48550/arXiv.2405.14509

[7] Vila, R., Nakano, E., and Saulo, H., 2024b. Novel closed-form point estimators for a weighted exponential family derived from likelihood equations. Stat, 13, 1–10. DOI: 10.1002/sta4.723

[8] Barros, M., Galea, M., Gonzalez, M., and Leiva, V., 2010. Influence diagnostics in the tobit censored response model. Statistical Methods and Applications, 19, 379–397. DOI: 10.1007/s10260-010-0135-y

[9] Vila, R., Saulo, H., Marchant, C., Leiva, V., and Castro, C., 2025. Bivariate extended skew-elliptical Heckman models: Mathematical characterization and an application in economic sciences. Computational and Applied Mathematics, 44, 236. DOI: 10.1007/s40314-025-03169-z

[10] Leiva, V., Saulo, H., Leao, J., and Marchant, C., 2014. A family of autoregressive conditional duration models applied to financial data. Computational Statistics and Data Analysis, 79, 175–191. DOI: 10.1016/j.csda.2014.05.016

[11] Kim, H.M., Kim, S., Jang, Y.H., and Zhao, J., 2022. New closed-form estimator and its properties. Journal of the Korean Statistical Society, 51, 47–64. DOI: 10.1007/s42952-021-00118-4

[12] Cheng, J. and Beaulieu, N.C., 2002. Generalized moment estimators for the Nakagami fading parameter. IEEE Communications Letters, 6, 144–146. DOI: 10.1109/4234.996038

[13] Baydil, B., de la Peña, V.H., Zou, H., and Yao, H., 2025. Unbiased estimation of the Gini coefficient. Statistics and Probability Letters, 222, 110376. DOI: 10.1016/j.spl.2025.110376

[14] Leao, J., Balakrishnan, N., Saulo, H., and Tomazella, V., 2017. Estimation for a family of skew scale-mixture distributions. Chilean Journal of Statistics, 8, 45–64. URL: soche.cl/chjs/volumes/08/02/Leao_etal(2017).pdf

[15] Saulo, H., Bourguignon, M., Zhu, X., and Balakrishnan, N., 2019. Some simple estimators for the two-parameter gamma distribution. Communications in Statistics: Simulation and Computation, 48, 2425–2437. DOI: 10.1080/03610918.2018.1457693

[16] Burr, I.W., 1942. Cumulative frequency functions. Annals of Mathematical Statistics, 13, 215–232. URL: www.jstor.org/stable/2235756

[17] Johnson, N.L., Kotz, S., and Balakrishnan, N., 1994. Continuous Univariate Distributions. Wiley, New York, NY, US.

[18] Dagum, C., 1975. A model of income distribution and the conditions of existence of moments of finite order. Bulletin of the International Statistical Institute, 46, 199–205.

[19] Bebbington, M., Lai, C.D., and Zitikis, R., 2007. A flexible Weibull extension. Reliability Engineering and System Safety, 92, 719–726. DOI: 10.1016/j.ress.2006.03.004

[20] Stacy, E.W., 1962. A generalization of the gamma distribution. Annals of Mathematical Statistics, 33, 1187–1192. DOI: 10.1214/aoms/1177704481

[21] Lee, M. and Gross, A., 1991. Lifetime distributions under unknown environment. Journal of Statistical Planning and Inference, 29, 137–143. DOI: 10.1016/0378-3758(92)90128-F

[22] Gompertz, B., 1825. On the nature of the function expressive of the law of human mortality and on the new model of determining the value of life contingencies. Philosophical Transactions of the Royal Society of London, 115, 513–585. DOI: 10.1098/rspl.1815.0271

[23] Cook, J.D., 2008. Inverse gamma distribution. John D. Cook Technical Report. URL: www.johndcook.com/inverse_gamma.pdf

[24] Khan, M.S., Pasha, G.R., and Pasha, A.H., 2008. Theoretical analysis of inverse Weibull distribution. WSEAS Transactions on Mathematics, 7, 30–38. URL: www.wseas.us/e-library/transactions/mathematics/2008/theoretical.pdf

[25] Dunbar, R.C., 1982. Deriving the Maxwell distribution. Journal of Chemical Education, 59, 22–23. DOI: 10.1021/ed059p22

[26] Xie, M., Tang, Y., and Goh, T.N., 2002. A modified Weibull extension with bathtub-shaped failure rate function. Reliability Engineering and System Safety, 76, 279–285. DOI: 10.1016/S0951-8320(02)00022-4

[27] Laurenson, D., 1994. Indoor radio channel propagation modeling by ray tracing techniques. Proceedings of the International Conference on Wireless Communications, pp. 65–68. URL: hdl.handle.net/1842/12397

[28] Rayleigh, L., 1880. XII. On the resultant of a large number of vibrations of the same pitch and of arbitrary phase. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 10, 73–78. DOI: 10.1080/14786448008626893

[29] Bernardo, J.M. and Smith, A.F.M., 1993. Bayesian Theory. Wiley, New York, NY, US.

[30] Nadarajah, S. and Kotz, S., 2005. On some recent modifications of Weibull distribution. IEEE Transactions on Reliability, 54, 561–562. DOI: 10.1109/TR.2005.858811

[31] Rahman, G., Mubeen, S., Rehman, A., and Naz, M., 2014. On k-gamma and k-beta distributions and moment generating functions. Journal of Probability and Statistics, 2014, 982013. DOI: 10.1155/2014/982013