A BAYESIAN APPROACH FOR CHRIS-JERRY DISTRIBUTION USING VARIOUS LOSS FUNCTIONS

Dr. G Meenakshi[1] and Balachandar.B[2]

•

(1). Professor, Department of Statistics, Annamalai University
(2). Research Scholar, Department of Statistics, Annamalai University
Email: progmau@gamil.com , balachandarb03061998@gmail.com

Abstract

The paper introduces a Bayesian approach for estimating parameters of the Chris-Jerry distribution, focusing on the use of a conjugate prior, specifically the gamma prior. The Bayesian estimation method is developed with a various loss function, offering a robust framework for parameter estimation. symmetric loss function and Linex loss functions are commonly used in Bayesian statistics to balance the trade-off between bias and variance. The central idea is to derive the Bayes estimate of the distribution parameter by leveraging the properties of the conjugate gamma prior. Conjugate priors simplify the Bayesian analysis by ensuring that the posterior distribution belongs to the same family as the prior, facilitating analytical calculations. The proposed methodology is implemented and validated through numerical illustrations using. This involves applying the developed Bayesian estimation framework to real-world data or simulated scenarios, demonstrating its effectiveness and practical applicability. The numerical and simulation studies are done by using r software

.

Keywords: prior, posterior distribution, posterior mean loss function, linex loss function and symmetric loss function

I. Introduction

In the realm of statistical inference, the Bayesian approach stands as a formidable paradigm, offering a unique perspective that seamlessly integrates prior knowledge with observed data to yield more robust and nuanced estimates. At the heart of Bayesian estimation lies the elegant concept of conjugate priors, a fact that not only simplifies the computational complexity but also enriches the analytical insights.

This article aims to unravel the significance of conjugate priors and their pivotal role in streamlining the inference process. From their foundational principles to practical applications, this paper will explore how these priors provide a harmonious bridge between prior beliefs and empirical evidence, creating a coherent framework for making informed decisions. The Bayesian paradigm, with its emphasis on updating beliefs in light of new information, has found extensive applications across various fields, from finance and engineering to medicine and machine learning. Within this framework, the choice of prior distributions can profoundly impact the outcome of Bayesian analyses. Conjugate priors, by virtue of their mathematical properties, offer an elegant solution, simplifying the computations involved in posterior distribution calculations. This article will delve into the conceptual underpinnings of Bayesian estimation, shedding light on the fundamental principles that distinguish it from frequentist approaches. Then transition to the concept of conjugate priors, explaining how these specially chosen prior distributions yield posterior distributions of the same family, facilitating analytical tractability.

Moreover, the research showcases real-world examples where Bayesian estimation with conjugate priors has proven to be a powerful tool, enhancing decision-making processes in situations ranging from medical diagnostics to quality control in manufacturing. By illustrating the versatility and efficiency of this methodology, we aim to empower readers to harness the full potential of Bayesian analysis in their own pursuits.

The Linex loss function, short for Linear Exponential loss function, is a variant of the asymmetric loss functions commonly used in regression analysis. Unlike traditional symmetric loss functions like Mean Squared Error (MSE) or Mean Absolute Error (MAE), Linex loss asymmetrically penalizes overestimation and underestimation differently. It is particularly useful when the cost of underestimation is not the same as the cost of overestimation, making it suitable for scenarios where errors in one direction are more critical than errors in the other. The Linex loss function is defined as

The LINEX loss function you provided is:

$$L(\theta, \hat{\theta}) = a.e^{b(\theta-\hat{\theta})} - b(\theta - \hat{\theta}) - 1$$

where:

- $L(\theta, \hat{\theta})$ is the LINEX loss function.
- θ is the true parameter value.
- $\hat{\theta}$ is the estimated parameter value.
- a and b are parameters that control the shape of the loss function.

This type of LINEX loss function is sometimes used in Bayesian estimation, and Zellner is indeed associated with Bayesian methods. In Bayesian statistics, the choice of a loss function is crucial in constructing a suitable posterior distribution. The LINEX loss function, as you've written it, is a combination of linear and exponential terms, and the parameters *a* and *b* determine the weight given to these terms.

A symmetric loss function is a mathematical function used to measure the discrepancy or error between predicted and actual values in a regression problem. Unlike asymmetric loss functions, which penalize overestimation and underestimation differently, symmetric loss functions treat overestimation and underestimation equally

The symmetric loss function $L(\theta, d) = C(d - \theta)^{2f}$ penalizes the deviation between the decision *d* and the unknown parameter θ. Here, *C* is a scaling constant, and *f* is a parameter that controls the sensitivity of the loss function to deviations. This loss function is symmetric because it penalizes deviations equally on both sides of the decision *d*. The exponent 2 f controls the curvature of the loss function around *d*. larger values of f make the loss function more sensitive to deviations from *d*, leading to sharper penalties.

In the context of Bayesian estimation, our exploration centers on the symmetric loss function, elegantly expressed as

$$L(\theta, d) = C(d - \theta)^{2f}$$

with C serving as a constant. The transformation into the quadratic loss function (QLF) occurs when f assumes the value of 1, resulting in the concise form

$$L(\theta, d) = C(d - \theta)^2$$

By streamlining the equation through the abstraction of C to 1, we seamlessly transition to the squared error loss function (SELF). Introducing an alternative, the absolute loss function takes the form L (θ, d) = ∣ (d − θ) ∣. Notably, the squared error loss function (SELF)

$$L(\theta, d) = (d - \theta)^2$$

The goal in Bayesian estimation is to find the posterior distribution of the parameter θ given the observed data. This involves combining the likelihood function with a prior distribution and the loss function. The posterior distribution is then obtained by maximizing the posterior expected loss (also known as the Bayes risk) with respect to θ or using other Bayesian decision theoretic criteria.

II. Review of literature

Box, G. E. P., Tiao, G. C., and Jenkins, G. M [6] done a foundational work in the field of Bayesian statistics. Zellner introduces the concept of Bayesian estimation and prediction with asymmetric loss functions and computational methods for Bayesian estimation and prediction using asymmetric loss functions [20]. Parsian introduces the concept of Bayes estimation using a LINEX loss function and explained its uses in Bayesian estimation and its advantages in decision-making under uncertainty [18]. Feroze, N. and Aslam, M. discusses [12] Bayesian analysis of the error function distribution using various loss functions and examine how different loss functions impact Bayesian estimation in the context of the error function distribution. Zaka, A. and Akhter, A. S. compares [19] various methods for estimating parameters of the power function distribution and done a simulation study and real-world. Chrisogonus K. Onyekwere and Okechukwu J. Obulezi [8] have proposed a new one-parameter distribution named Chris-Jerry is suggested from a two-component mixture of Exponential (θ) distribution and Gamma (3, θ) distribution with mixing proportion p = θ/ θ+2 having a flexibility advantage in modeling lifetime data. In this paper the posterior mean of Chris-jerry distribution is derived with various loss function.

III. A Bayesian approach for Chris-jerry distribution

The probability distribution function of Chris-jerry distribution is given by

$$f(x) = \frac{\theta^2}{\theta+2}.(1 + \theta x^2).e^{-\theta x} \qquad x, \theta > 0 \tag{1}$$

In this section the posterior distribution of Chris-jerry distribution is obtained. Let $X_1, X_2, ...$ be a sequence of random variables from Chris-jerry distribution, then the likelihood function is given by

$$\pi(x_i) = \prod_{i=1}^{n} \frac{\theta^2}{\theta+2}.(1 + \theta x_i^2).e^{-\theta x_i} \tag{2}$$

The prior is gamma prior (conjugate prior)

$$p(\theta) = \frac{e^{-\theta}.\theta^{r-1}}{\gamma r} \qquad r > 0, \theta > 0 \tag{3}$$

The posterior distribution is given by

$$p(\theta/x) = \frac{1}{k} * \frac{\theta^{2n}}{(\theta+2)^n} * \prod_{i=1}^{n}(1+\theta x_i) \quad * \left(e^{-\theta\sum_1^n x_i}\right) * \frac{e^{-\theta}.\theta^{r-1}}{\Upsilon r} \tag{4}$$

Were

$$k = \int_0^\infty \frac{\theta^{2n}}{(\theta+2)^n}(1+\theta x_i)^n.\left(e^{-\theta\sum_1^n x_i}\right) * \frac{e^{-\theta}.\theta^{r-1}}{\Upsilon r} d\theta$$

$$K = \sum_{j=0}^{\infty}(-1)^j c(n+j-1,j)\sum_{l=0}^{\infty} c(n,i)(x)^l \; \frac{\{\Upsilon(2n+l+j+r)\}}{2^{n+j}[\Upsilon r][\sum_1^n x_i + 1]^{2n+l+j+r}}$$

The posterior mean is given by

$$E[\theta] = \int_0^\infty \theta p(\theta/x) d\theta$$

$$= \int_0^\infty \theta \frac{1}{k} * \frac{\theta^{2n}}{(\theta+2)^n} * \prod_{i=1}^{n}(1+\theta x_i) \quad * \left(e^{-\theta\sum_1^n x_i}\right) * \frac{e^{-\theta}.\theta^{r-1}}{\Upsilon r} d\theta$$

Case: I

Bayesian estimation of θ under linex loss function by Zellner [**Zellner, A. (1986)**.]

$$L(\theta,\hat{\theta}) = a.e^{b(\theta-\hat{\theta})} - b(\theta-\hat{\theta}) - 1 \tag{5}$$

Where a>0, b≠0; a is scale of loss function and b determines its shape. Without loss of generality, we assume a= 1 and obtain bayes estimate of θ

In Zeller's linex loss function, $\hat{\theta}$ represents the reference value or the target value that the parameter θ is compared to $\hat{\theta}$. It can be thought of as the "ideal" or "desired" value of θ. The linex loss function measures the deviation of θ from $\hat{\theta}$

For example, if you are estimating a parameter θ and you have a prior belief or expectation about its value, you might set $\hat{\theta}$to that prior belief. Then, the linex loss function would measure how much the estimated value of θ deviates from that prior belief.

Here

$$E[L(\theta,\hat{\theta})] = \frac{1}{k\Gamma r}\int_0^\infty \left[1.e^{b(\theta-\hat{\theta})} - b(\theta-\hat{\theta}) - 1\right] * p(\theta/x)\, d\theta$$

$$= \int_0^\infty \left[1.e^{b(\theta-\hat{\theta})} - b(\theta-\hat{\theta}) - 1\right] * \frac{\theta^{2n+r-1}}{\Upsilon r(\theta+2)^n}(1+\theta x_i)^n.\left(e^{-\theta\sum_1^n x_i - \theta}\right)d\theta$$

$$= \frac{1}{\Upsilon r}\int_0^\infty \left[1.e^{b(\theta-\hat{\theta})} - b(\theta-\hat{\theta}) - 1\right] * \frac{\theta^{2n+r-1}}{(\theta+2)^n}(1+\theta x_i)^n.\left(e^{-\theta\sum_1^n x_i - \theta}\right)d\theta$$

$$\sum_{j=0}^{\infty}(-1)^j c(n+j-1\,,j) * \left(\frac{1}{2}\right)^j = d$$

$$\sum_{l=0}^{\infty} c(n,i)(x)^{l} = f$$

$$\sum_{i=1}^{n} x_i + 1 = h$$

$$E\left[L\left(\theta,\hat{\theta}\right)\right] = \frac{1df}{\Upsilon r}\left[\left(\left\{\frac{e^{-b\hat{\theta}}.\Upsilon(2n+r+j+l)}{[h-b]^{2n+r+j+l}}\right\}\right) - \left\{\frac{b(\Upsilon 2n+r+j-l)}{h^{2n+r+j-l}}\right\} + \hat{\theta}b\left\{\frac{\Upsilon 2n+r+j+l}{h^{2n+r+j+l}}\right\} - \left\{\frac{\Upsilon 2n+r+j+l}{h^{2n+r+j+l}}\right\}\right]$$

Case: II (Bayesian estimation of θ under Symmetric loss function)
Symmetric loss function for the decision d for the unknown parameter θ is defined by

$$L(\theta, d) = C(d - \theta)^{2f} \tag{6}$$

f= 1, 2…..

$$E[L(\theta,d)] = \frac{1}{k\Gamma r}\int_{0}^{\infty} [C(d - \theta)^{2f}] * \frac{\theta^{2n+r-1}}{(\theta+2)^{n}}.\left(e^{-\theta(\sum_{1}^{n} x_i+1)}\right)\prod_{i=1}^{n}(1+\theta x_i) \quad d\theta$$

where C is a constant. When f = 1 reduces to quadratic loss function (QLF) given by

$$L(\theta, d) = C(d - \theta)^{2} \tag{7}$$

For some constant C. The value of the constant C makes no difference to a decision, and can be ignored by setting it equal to 1 and reduced to the SELF. Absolute loss function is another symmetric loss function given by

$$L(\theta, d) = |(d - \theta)| \tag{8}$$

The squared error loss function (SELF) is widely used in decision theory problems and is defined as

$$L(\theta, d) = (d - \theta)^{2} \tag{9}$$

IV. Simulation Study

Posterior Distribution: The posterior distribution represents our updated beliefs about the parameter θ after observing the data. It is proportional to the product of the prior distribution and the likelihood function. Sampling: in this research Markov Chain Monte Carlo (MCMC) methods is used, implemented in the r software, to sample from the posterior distribution and estimate the parameters of interest.

The following algorithm outlines the steps involved in Bayesian inference using the specified model and data. It involves defining the model, computing the likelihood, performing posterior inference, analyzing the results, and outputting the estimates and diagnostics.

Stan Code Explanation

- Input Data:

Obtain input data:
n: Number of data points.
x: Observed data points.
r: Parameter influencing the shape of the gamma prior distribution for theta.
C: Constant used in the loss function.
f: Exponent used in the loss function.
d: Target value used in the loss function.

- Initialize Model:
  Define the prior distribution:
  θ ~ Gamma (shape = r, rate = 1).
- Likelihood Calculation:
  Compute the likelihood using a custom loss function:
  Calculate the log-likelihood contribution for each data point:
  Compute terms related to the observed data x, the parameter theta, and the loss function parameters C, f, and d.
  Accumulate the log-likelihood contributions.
- Posterior Inference:
  Combine the prior distribution and likelihood to obtain the posterior distribution:
  Posterior ∝ Prior × Likelihood.
- Bayesian Inference:
  Use Bayesian inference techniques (such as Markov Chain Monte Carlo) to sample from the posterior distribution:
  Obtain posterior samples for the parameter theta using Stan's sampling algorithm.
  Specify the number of chains and iterations for sampling.
- Analysis:
  Analyze the posterior samples to estimate the posterior distribution of theta:
  Compute summary statistics (e.g., mean, median, quantiles) of the posterior samples.
  Visualize the posterior distribution if necessary.
- Output:
  Output the results of the analysis, such as posterior mean estimates, credible intervals, and diagnostic information about the inference procedure.

Table 1: *Comparison of Posterior Means for Different Loss Functions and Sample Sizes (simulated data)*

n/ $E[\theta]$	50	100	200
Without loss	0.0861636	0.04134565	0.02038608
Symmetric loss function	3.458547	3.045888	2.563339
Quadratic loss function (QLF)	0.08745913	0.04159895	0.02039235
Squared error loss function	0.08671715	0.04156098	0.02036867
Linex loss function	0.08599036	0.04149021	0.0203717

From the above table show that the posterior mean increase when d and f values increase. Which mean the larger values of f make the loss function more sensitive to deviations from *d*, leading to sharper penalties.

V. Real life data

Table: 2 *shows the life of fatigue fracture of Kevlar 373/epoxy subjected to constant pressure at 90 % stress level until all had failed. Source:(8)*

0.0251	0.6748	0.912	1.3503	1.7746	2.0408	2.4951	4.8073
0.0886	0.6751	0.9836	1.3551	1.8475	2.0903	2.526	5.4005
0.0891	0.6753	1.0483	1.4595	1.8375	2.1093	2.9911	5.4435
0.2501	0.7696	1.0596	1.488	1.8503	2.133	3.0256	5.5295
0.3113	0.8375	1.0773	1.5728	1.8808	2.21	3.2678	6.5541
0.3451	0.8391	1.1733	1.5733	1.8878	2.246	3.4045	9.096
0.4763	0.8425	1.257	1.7083	1.8881	2.2878	3.4846	
0.565	0.8645	1.2766	1.7263	1.9316	2.3203	3.7433	
0.5671	0.8851	1.295	1.746	1.9558	2.347	3.7455	
0.6566	0.9113	1.3211	1.763	2.0048	2.3513	3.9143	

.

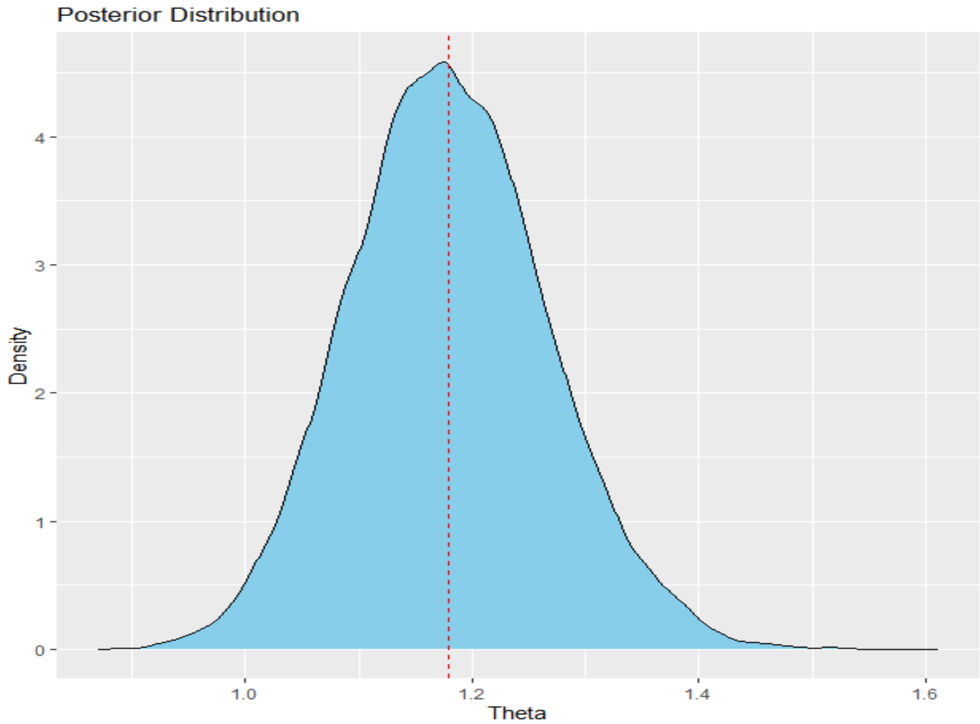

Figure: 1 *Graph of posterior distribution of Chris-jerry distribution*

Table: 3 *Comparison of Posterior Means for Different Loss Functions (real-life data)*

Loss functions	Posterior Mean
Without any loss function	1.179317
Symmetric loss function	1.220796
quadratic loss function	1.200863
Squared error loss function	1.201302
Linex loss function	1.197762

The above presents a comparative analysis of posterior means under different loss functions, revealing notable disparities in estimation outcomes. Without any loss function, the posterior mean is observed to be substantially higher, suggesting potential bias in the estimation process. Conversely, employing loss functions leads to a decrease in the posterior mean, highlighting the influence of the chosen loss function. Specifically, the symmetric loss function yields the highest posterior mean, while the Linex loss function results in the lowest. These findings underscore the significance of selecting an appropriate loss function. Notably, the posterior mean of the provided data has been calculated, emphasizing the practical relevance of these results.

VI. Conclusion

The Bayesian estimation of parameters for the Chris-Jerry distribution with a gamma prior, considering various loss functions, reveals subtle differences in posterior mean estimates. The absence of a specific loss function yields a posterior mean estimate of 1.179317, while employing a symmetric loss function slightly increases the estimate to 1.220796. Conversely, quadratic and squared error loss functions result in slightly lower estimates of 1.200863 and 1.201302, respectively. The use of a Linex loss function produces a posterior mean estimate of 1.197762. These findings underscore the importance of the choice of loss function in Bayesian estimation. While variations are observed in the posterior mean estimates across different loss functions, the differences remain relatively subtle, indicating robustness in the estimation process. However, it is essential to note that these results are contingent upon the provided dataset and may vary with alternative datasets or priors. In conclusion, this study contributes to the understanding of Bayesian estimation methods for the Chris-Jerry distribution with a gamma prior. Future research could delve deeper into exploring additional loss functions and their implications for parameter estimation in Bayesian frameworks, thereby enhancing the applicability of these methods in diverse statistical analyses.

References

[1] Ahmed, A., Ahmad, S., & Reshi, J. (2013). Bayesian analysis of Rayleigh distribution. International Journal of Scientific and Research Publications, 3(10), 1–9.

[2] Ahmed, M. A. (2020). On the alpha power Kumaraswamy distribution: Properties, simulation and application. Revista Colombiana de Estadística, 43(2), 285–313. https://doi.org/10.15446/rce.v43n2.83598

[3] Ahmed, Z., Nofal, Z. M., Abd El Hadi, N. E., et al. (2015). Exponentiated transmuted generalized Rayleigh distribution: A new four-parameter Rayleigh distribution. Pakistan Journal of Statistics and Operation Research, 11(1), 115–134. https://doi.org/10.18187/pjsor.v11i1.873

[4] Andrews, D. F., & Herzberg, A. M. (1985). Stress-rupture life of Kevlar 49/epoxy spherical pressure vessels. Data. Springer, 181–186.

[5] Berger, J. O. (1985). Statistical Decision Theory and Bayesian Analysis. Springer.

[6] Box, G. E. P., Tiao, G. C., & Jenkins, G. M. (1970). Bayesian Statistical Inference. John Wiley & Sons.

[7] Carlin, B. P., & Louis, T. A. (2009). Bayesian Methods for Data Analysis. CRC Press.

[8] Onyekwere, C. K., & Obulezi, O. J. (2022). Chris-Jerry Distribution and Its Applications. Asian Journal of Probability and Statistics, 20(1), 16–30. ISSN: 2582-0230

[9] Dey, D. K., & Liu, P.-S. L. (1992). On comparison of estimators in a generalized life model. Microelectronics Reliability, 32(1-2), 207–221.

[10] Dey, D. K., Ghosh, M., & Srinivasan, C. (1987). Simultaneous estimation of parameters under entropy loss. Journal of Statistical Planning and Inference, 15, 347–363.

[11] Dey, S. (2012). Bayesian estimation of the parameter and reliability function of an Inverse Rayleigh distribution. Malaysian Journal of Mathematical Sciences, 6(1), 113–124.

[12] Feroze, N., & Aslam, M. (2012). A note on Bayesian analysis of error function distribution under different loss functions. International Journal of Probability and Statistics, 1(5), 153–159.

[13] Meenakshi, G. (2013). A Bayesian method to Predetermination of Viral replication in the HIV Dynamics. Fast East Journal of Theoretical Statistics, 42(2), 71–81.

[14] Meenakshi, G., et al. (2019). Prediction of HIV replication in the human immune system using multinomial distribution by Bayesian methodology. International Journal of Scientific Research in Mathematical and Statistical Sciences, 6(1), 46–52.

[15] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehari, A., & Rubin, D. B. (2013). Bayesian Data Analysis. Chapman and Hall/CRC.

[16] Hobbs, N. T., & Hooten, M. B. (2015). Bayesian Models: A Statistical Primer for Ecologists. Princeton University Press.

[17] Lindley, D. V., & Smith, A. F. (1972). Bayes estimates for the linear model. Journal of the Royal Statistical Society. Series B (Methodological), 34(1), 1–41.

[18] Parsian, A. (1990). Bayes estimation using a LINEX loss function. Risk, 1(4), 305–307.

[19] Zaka, A., & Akhter, A. S. (2013). Methods for estimating the parameters of the power function distribution. Pakistan Journal of Statistics and Operational Research, 9(2), 213–224.

[20] Zellner, A. (1986). Bayesian Estimation and Prediction Using Asymmetric Loss Functions. Journal of the American Statistical Association, 81, 446–450.