Bayesian Analysis of Cancer Data Using a 4-Component Exponential Mixture Model

Farzana Noor, Saadia Masood, Yumna Sabar, Syed Bilal Hussain Shah, Touqeer Ahmad, Asrin Abdollahi, and Ahthasham Sajid

1Department of Mathematics & Statistics, International Islamic University, Islamabad, Pakistan
2Department of Mathematics and Statistics, PMAS University of Arid Agriculture, Rawalpindi, Pakistan
3Manchester Metropolitan University, UK
4Department of Statistics Sciences University of Padova, Italy
5Department of Electrical Engineering, University of Kurdistan, Sanandaj, Iran
6Department of Computer Science, Faculty of ICT, BUITEMS, Quetta, Baluchistan, Pakistan

Correspondence should be addressed to Asrin Abdollahi; a.abdollahi@eng.uok.ac.ir

Received 6 August 2021; Revised 3 September 2021; Accepted 11 September 2021; Published 12 October 2021

Copyright © 2021 Farzana Noor et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Cancer is among the major public health problems as well as a burden for Pakistan. About 148,000 new patients are diagnosed with cancer each year, and almost 100,000 patients die due to this fatal disease. Lung, breast, liver, cervical, blood/bone marrow, and oral cancers are the most common cancers in Pakistan. Perhaps smoking, physical inactivity, infections, exposure to toxins, and unhealthy diet are the main factors responsible for the spread of cancer. We preferred a novel four-component mixture model under Bayesian estimation to estimate the average number of incidences and death of both genders in different age groups. For this purpose, we considered 28 different kinds of cancers diagnosed in recent years. Data of registered patients all over Pakistan in the year 2012 were taken from GLOBOCAN. All the patients were divided into 4 age groups and also split based on genders to be applied to the proposed mixture model. Bayesian analysis is performed on the data using a four-component exponential mixture model. Estimators for mixture model parameters are derived under Bayesian procedures using three different priors and two loss functions. Simulation study and graphical representation for the estimates are also presented. It is noted from analysis of real data that the Bayes estimates under LINEX loss assuming Jeffreys’ prior is more efficient for the no. of incidences in male and female. As far as no. of deaths are concerned again, LINEX loss assuming Jeffreys’ prior gives better results for the male population, but for the female population, the best loss function is SELF assuming Jeffreys’ prior.

1. Introduction

Most people think that there is no fatal diagnosis other than that of cancer. However, this may be an exaggerated and overgeneralized vision of cancer. But, it is always admitted as a serious life-threatening disease. Cancer is considered a major cause of death among deaths due to the noncontagious diseases in Pakistan. Even then, we do not find any study from Pakistan on cancer-specific incidence and mortality rates based on age groups. In the last 25 years, Pakistan had observed a significant increase in the number of various kinds of cancer cases. In Pakistan, there is no work done in the field of cancer using Bayesian analysis. In this study, we have conducted a Bayesian analysis of data about the survival rate of cancer patients through a mixture model.

Intercellular communication and culture conditions are the main supporter in the formation of cancer cells [1], but no specific cause for cancer can be identified [2]. In Pakistan, cancer is a major health problem. The use of tobacco, the growing and aging population, and the westernization diet are the various factors that tend to increase cancer cases. Studies reveal that every year, nearly 300,000 various new kinds of cancer cases are reported from all over the country. Everyone can be affected by cancer at any age. Early
Some important studies on thyroid cancer include [3

physically active with at least a 30-minute workout daily.
merely by not using tobacco, using healthy diet, and being
third to forty percent of all cancer cases are preventable
very important to stop and prevent the disease. At least one-
detection, diagnosis, treatment, and especially awareness are

detection, diagnosis, treatment, and especially awareness are
very important to stop and prevent the disease. At least one-
detection, diagnosis, treatment, and especially awareness are

Table 1: Simulation results of informative prior, Jeffreys’ prior, and Jeffreys’ gamma prior under different loss functions when $\pi_1 = 1.5, \pi_2 = 2.5, \pi_3 = 1.75, \pi_4 = 2.5, p_1 = 0.35, p_2 = 0.20, p_3 = 0.15, T = 1.05$.

| n   | Parameter | IP | SELF | Jeffreys’ gamma | IP | LINEX | Jeffreys’ gamma |
|-----|-----------|----|------|-----------------|----|-------|----------------|
|     |           | IP | JP   | Jeffreys’ gamma | IP | JP    | Jeffreys’ gamma |
| 100 | $\tilde{\pi}_1$ | 1.674 (0.168) | 1.850 (0.259) | 1.648 (0.163) | 1.653 (0.021) | 1.707 (0.113) | 1.643 (0.001) |
|     | $\tilde{\pi}_2$ | 2.560 (0.461) | 2.625 (0.714) | 2.424 (0.436) | 2.429 (0.054) | 2.304 (0.312) | 2.360 (0.004) |
|     | $\tilde{\pi}_3$ | 1.824 (0.279) | 2.019 (0.751) | 1.686 (0.254) | 1.742 (0.033) | 1.737 (0.312) | 1.673 (0.002) |
|     | $\tilde{\pi}_4$ | 2.543 (0.335) | 2.601 (0.453) | 2.491 (0.326) | 2.472 (0.041) | 2.392 (0.207) | 2.441 (0.003) |
|     | $\hat{p}_1$  | 0.338 (0.002) | 0.326 (0.0028) | 0.332 (0.0027) | 0.332 (0.0013) | 0.318 (0.009) | 0.308 (0.035) |
|     | $\hat{p}_2$  | 0.204 (0.0019) | 0.209 (0.0020) | 0.206 (0.0018) | 0.203 (0.0004) | 0.205 (0.003) | 0.200 (0.009) |
|     | $\hat{p}_3$  | 0.152 (0.0017) | 0.1610 (0.0019) | 0.158 (0.0016) | 0.152 (0.0003) | 0.158 (0.002) | 0.155 (0.005) |
|     | $\tilde{\pi}_1$ | 1.647 (0.097) | 1.712 (0.121) | 1.646 (0.095) | 1.623 (0.011) | 1.654 (0.055) | 1.623 (0.001) |
|     | $\tilde{\pi}_2$ | 2.530 (0.285) | 2.559 (0.365) | 2.472 (0.279) | 2.455 (0.034) | 2.385 (0.169) | 2.434 (0.003) |
|     | $\tilde{\pi}_3$ | 1.791 (0.203) | 1.875 (0.366) | 1.709 (0.191) | 1.755 (0.024) | 1.700 (0.161) | 1.697 (0.002) |
| 200 | $\tilde{\pi}_4$ | 2.521 (0.199) | 2.572 (0.237) | 2.499 (0.195) | 2.466 (0.024) | 2.455 (0.112) | 2.497 (0.002) |
|     | $\hat{p}_1$  | 0.339 (0.001) | 0.336 (0.0015) | 0.336 (0.0015) | 0.336 (0.0011) | 0.326 (0.009) | 0.313 (0.0337) |
|     | $\hat{p}_2$  | 0.203 (0.001) | 0.205 (0.0010) | 0.204 (0.0009) | 0.203 (0.0003) | 0.203 (0.002) | 0.199 (0.0076) |
|     | $\hat{p}_3$  | 0.153 (0.000) | 0.165 (0.00010) | 0.157 (0.0009) | 0.152 (0.0002) | 0.155 (0.0001) | 0.153 (0.0039) |
|     | $\tilde{\pi}_1$ | 1.616 (0.065) | 1.643 (0.078) | 1.618 (0.066) | 1.582 (0.008) | 1.621 (0.037) | 1.609 (0.0007) |
|     | $\tilde{\pi}_2$ | 2.503 (0.205) | 2.505 (0.244) | 2.471 (0.203) | 2.485 (0.025) | 2.399 (0.118) | 2.438 (0.0022) |
|     | $\tilde{\pi}_3$ | 1.795 (0.163) | 1.821 (0.248) | 1.713 (0.154) | 1.747 (0.020) | 1.707 (0.114) | 1.709 (0.0017) |
| 300 | $\tilde{\pi}_4$ | 2.519 (0.142) | 2.563 (0.162) | 2.496 (0.140) | 2.471 (0.017) | 2.450 (0.077) | 2.467 (0.0015) |
|     | $\hat{p}_1$  | 0.341 (0.001) | 0.338 (0.0011) | 0.338 (0.0010) | 0.339 (0.0011) | 0.328 (0.009) | 0.315 (0.0337) |
|     | $\hat{p}_2$  | 0.203 (0.000) | 0.204 (0.0006) | 0.203 (0.0007) | 0.201 (0.0002) | 0.202 (0.0002) | 0.1993 (0.0070) |
|     | $\hat{p}_3$  | 0.152 (0.000) | 0.156 (0.0007) | 0.155 (0.0006) | 0.152 (0.0001) | 0.155 (0.0021) | 0.153 (0.0039) |

Here, $f_j$ and $p_j$ represent component densities and weight factors, respectively. A mixture model generally may be composed of several components which can take the same or a different distributional form. Simplicity is obtained if the mixture model is composed of the same distributions.

Many authors have considered estimation of mixture models in their work such as McCullagh [9] who generates a mixture of linear exponential models using quadratic and exponential models. Abu-Taleb et al. [10] present Bayes estimation for the parameters of the lifetime distribution when both censoring and survival time are exponentially distributed. Noor et al. [11] have analyzed a mixture model by mixing Rayleigh and Burr XII distribution under a Bayesian setup. Abu Zinadah [12] presents maximum likelihood estimation and Bayesian analysis on exponential distribution and exponential pareto under type II censoring. Feroze and Aslam [13] have considered the Bayesian analysis of Burr type X distribution. Noor and Aslam [14] present the Bayesian analysis for the mixture of two inverse Weibull models. Tsutakawa [15] applies the Bayesian technique for assessing death rates of cancer when the recurrence of passing over a predefined era is expected to have Poisson distribution. Lambert et al. [16] consider a study of population-based

\[
f(y|\pi) = \sum_{j=1}^{m} p_j f_j(y|\pi_j) \text{where } p_j > 0, j = 1, 2, \ldots, m \text{ and } \sum_{j=1}^{m} p_j = 1. \tag{1}
\]

death, diagnosis, treatment, and especially awareness are
detection, diagnosis, treatment, and especially awareness are
cancer. They expand the parametric nonmixture cure part and accordingly give estimates of the cure division in population-based cancer. Hamdi [17] considers Bayesian statistical modelling in Ohio State for count data on cancer death.

Censoring is inevitable in experiments related to the life testing of some subjects/objects. A sample is a censored sample whenever it does not contain full information due to some experimental conditions. For example, a lung cancer patient is enrolled for a clinical trial to test the effect of a drug on his survival from his disease. But, he died in a car accident after $T$ years of his disease. His survival with lung cancer is at least $T$ years, but the exact years cannot be known. Though researchers have introduced/used different censoring schemes such as right, left, type I or type II censoring, and interval censoring, but right censoring is mostly used in life testing, see Cohen [18] for details on the censoring.

By introducing a 4-component exponential mixture model, the objective of the study is to contribute to the widest spreading field of mixture models and provide its application to cancer data. The Bayesian technique is opted to analyze the mixture model. Bayesian analysis is performed using different priors and loss functions assuming data is right-censored. Mainly, the paper is designed in the following manner: Materials and Methods contains a four-component exponential mixture model, likelihood, posterior densities using informative prior (IP) and noninformative prior (NIP), Bayes estimators, and posterior risks. In Results and Discussions, simulated and real-life data results are presented. Finally, the conclusion of the study is presented.

### 2. Materials and Methods

#### 2.1. Component Mixture of Exponential Distributions and Likelihood Function

Let a random variable $Y$ be exponentially distributed with parameter $\pi_l$, with probability density function:

$$f_l(y; \pi_l) = \pi_l e^{-\pi_l y}, y \geq 0, \pi_l > 0, l = 1, 2, 3, 4. \tag{2}$$

The parameter $\pi_l$ represents the rate at which an event occurs.

And the c.d.f is given as

$$F_l(y) = 1 - e^{-\pi_l y}. \tag{3}$$

Thus, a mixture model following a 4-component density

| $n$ | Parameter | IP | SELF | IP | \$IP_4$ | Jefferys’ gamma | IP | LINEX | Jefferys’ gamma |
|-----|-----------|----|------|----|---------|-----------------|----|-------|-----------------|
| 100 | $\hat{\pi}_1$ | 1.812 (0.261) | 2.010 (0.402) | 1.719 (0.240) | 1.737 (0.030) | 1.866 (0.187) | 1.660 (0.007) |
|     | $\hat{\pi}_2$ | 2.257 (0.220) | 2.287 (0.278) | 2.240 (0.215) | 2.214 (0.027) | 2.157 (0.130) | 2.217 (0.006) |
|     | $\hat{\pi}_3$ | 1.752 (0.311) | 1.921 (0.138”) | 1.567 (0.277) | 1.679 (0.037) | 1.441 (0.399) | 1.535 (0.008) |
|     | $\hat{\pi}_4$ | 2.553 (0.387) | 2.611 (0.542) | 2.490 (0.378) | 2.452 (0.046) | 2.408 (0.254) | 2.398 (0.011) |

| 200 | $\hat{\pi}_1$ | 0.253 (0.0025) | 0.246 (0.0023) | 0.254 (0.0023) | 0.250 (0.0007) | 0.241 (0.0045) | 0.251 (0.0001) |
|     | $\hat{\pi}_2$ | 0.395 (0.0029) | 0.387 (0.0028) | 0.382 (0.0027) | 0.386 (0.0018) | 0.372 (0.0151) | 0.381 (0.0002) |
|     | $\hat{\pi}_3$ | 0.097 (0.0011) | 0.111 (0.0015) | 0.108 (0.0012) | 0.098 (0.0001) | 0.112 (0.0011) | 0.108 (0.0004) |
|     | $\hat{\pi}_4$ | 1.785 (0.154) | 1.951 (0.208) | 1.758 (0.150) | 1.757 (0.018) | 1.842 (0.097) | 1.741 (0.004) |

| 300 | $\hat{\pi}_1$ | 0.252 (0.0013) | 0.247 (0.0012) | 0.252 (0.0012) | 0.249 (0.0005) | 0.242 (0.0038) | 0.251 (0.0008) |
|     | $\hat{\pi}_2$ | 0.395 (0.0015) | 0.392 (0.0015) | 0.389 (0.0014) | 0.391 (0.0017) | 0.377 (0.0147) | 0.388 (0.0021) |
|     | $\hat{\pi}_3$ | 0.098 (0.0006) | 0.107 (0.0008) | 0.104 (0.0006) | 0.099 (0.0001) | 0.107 (0.0007) | 0.104 (0.0002) |
|     | $\hat{\pi}_4$ | 1.792 (0.114) | 1.861 (0.136) | 1.750 (0.108) | 1.766 (0.013) | 1.803 (0.064) | 1.743 (0.003) |

Finally, the conclusion of the study is presented.
Figure 1: Graphical representation of Bayes estimator and posterior risk of simulated data.
Table 3: Bayes estimates and posterior risk for incidences of male and female under different priors and loss functions.

| Parameter | IP | SELF | Jeffreys’ gamma | IP | LINEX | Jeffreys’ gamma |
|-----------|----|------|-----------------|----|-------|-----------------|
| \( \hat{\pi}_1 \) | 0.0064 | 0.0062 | 0.0064 | 0.0068 | 0.0060 | 0.0060 |
| (8.28 \times 10^{-5}) | (8.10 \times 10^{-5}) | (8.51 \times 10^{-5}) | (1.03 \times 10^{-7}) | (1.01 \times 10^{-7}) | (4.25 \times 10^{-7}) |
| \( \hat{\pi}_2 \) | 0.0051 | 0.0042 | 0.0047 | 0.0045 | 0.0039 | 0.0040 |
| (8.66 \times 10^{-5}) | (1.08 \times 10^{-4}) | (1.22 \times 10^{-3}) | (1.07 \times 10^{-7}) | (1.36 \times 10^{-7}) | (6.12 \times 10^{-7}) |
| \( \hat{\pi}_3 \) | 0.0032 | 0.0028 | 0.0042 | 0.0028 | 0.0030 | 0.0035 |
| (1.32 \times 10^{-4}) | (1.04 \times 10^{-4}) | (1.31 \times 10^{-3}) | (1.65 \times 10^{-7}) | (1.31 \times 10^{-7}) | (6.56 \times 10^{-7}) |
| \( \hat{\pi}_4 \) | 0.0029 | 0.0032 | 0.0028 | 0.0030 | 0.0035 | 0.0030 |
| (7.05 \times 10^{-5}) | (7.09 \times 10^{-5}) | (5.85 \times 10^{-4}) | (8.94 \times 10^{-8}) | (8.88 \times 10^{-8}) | (2.92 \times 10^{-7}) |
| \( \hat{p}_1 \) | 0.2555 (0.0008) | 0.2545 (0.0008) | 0.2547 (0.0008) | 0.2534 (0.0004) | 0.2536 (0.0004) | 0.2508 (0.0038) |
| \( \hat{p}_2 \) | 0.2053 (0.0007) | 0.2058 (0.0008) | 0.2041 (0.0009) | 0.2038 (0.0002) | 0.2025 (0.0003) | 0.2019 (0.0021) |
| \( \hat{p}_3 \) | 0.2139 (0.0012) | 0.2235 (0.0012) | 0.2048 (0.0010) | 0.2153 (0.0004) | 0.2227 (0.0004) | 0.2025 (0.0022) |

Which assumes exponential distributions with unknown mixing proportions \( p_1, p_2, \) and \( p_3 \) may take the form:

\[
f(y; \pi_1, \pi_2, \pi_3, \pi_4, p_1, p_2, p_3) = p_1 \pi_1 e^{-\pi_1 y} + p_2 \pi_2 e^{-\pi_2 y} + p_3 \pi_3 e^{-\pi_3 y} + (1 - p_1 - p_2 - p_3) \pi_4 e^{-\pi_4 y}.
\]

(4)

And the c.d.f of the mixture model is

\[
F(y) = 1 - p_1 e^{-\pi_1 y} - p_2 e^{-\pi_2 y} - p_3 e^{-\pi_3 y} - (1 - p_1 - p_2 - p_3) e^{-\pi_4 y}.
\]

(5)

Let an experiment for testing lifetimes of some objects with \( n \) units is performed for the 4-component mixture model. It is assumed that, for a prespecified time, the experimenter will get \( s \) units failed and the remaining \( n - s \) units are removed from the experiment without knowing their lifetime and population as well. These failed \( s \) units are classified as \( s_1, s_2, s_3, \) and \( s_4 \) that can be assigned to respective subpopulations after knowing the cause of their failure according to Mendenhall and Hader [19] such that \( s = s_1 + s_2 + s_3 + s_4. \) Now, define \( y_{ij} \), \( 0 < y_{ij} < t_i \), as the failure time of \( i^{th} \) unit belonging to \( j^{th} \) subpopulation \( s_j \), \( j = 1, 2, 3, 4 \). Thus, the likelihood function of the 4-component mixture model for the random variable \( y \) is given as

\[
L(p \mid y) \propto \prod_{i=1}^{n} \left\{ p_1 f_1(y_{i1}) \right\} \left\{ p_2 f_2(y_{i2}) \right\} \left\{ p_3 f_3(y_{i4}) \right\} \left\{ 1 - F(t) \right\}^{-n-s},
\]

(6)

2.2. The Posterior Distribution Using IP. Gamma distribution is used as prior for component parameters \( \pi_1, \pi_2, \pi_3, \) and \( \pi_4, \) and bivariate beta distribution is chosen prior for proportion
Table 4: Bayes estimates and posterior risk for deaths of male and female under different priors and loss functions.

| Parameter | IP | SELF | Jefferys’ gamma | IP | LINEX | Jefferys’ gamma |
|-----------|----|------|-----------------|----|-------|-----------------|
| Male      |    |      |                 |    |       |                 |
| $\hat{\pi}_1$ | 0.0119 | 0.0112 | 0.0116 | 0.0125 | 0.0119 | 0.0120 |
|           | (2.72 × 10^{-4}) | (2.61 × 10^{-4}) | (2.72 × 10^{-4}) | (3.40 × 10^{-7}) | (3.27 × 10^{-7}) | (1.36 × 10^{-6}) |
| $\hat{\pi}_2$ | 0.0072 | 0.0064 | 0.0071 | 0.0066 | 0.0056 | 0.0075 |
|           | (1.90 × 10^{-4}) | (1.72 × 10^{-4}) | (1.93 × 10^{-4}) | (2.36 × 10^{-7}) | (2.15 × 10^{-7}) | (9.68 × 10^{-7}) |
| $\hat{\pi}_3$ | 0.0061 | 0.0049 | 0.0062 | 0.0055 | 0.0042 | 0.0056 |
|           | (1.64 × 10^{-4}) | (1.43 × 10^{-4}) | (1.75 × 10^{-4}) | (2.06 × 10^{-7}) | (1.80 × 10^{-7}) | (8.77 × 10^{-7}) |
| $\hat{\pi}_4$ | 0.0032 | 0.0032 | 0.0033 | 0.0027 | 0.0025 | 0.0039 |
|           | (2.41 × 10^{-4}) | (3.05 × 10^{-4}) | (2.74 × 10^{-5}) | (3.01 × 10^{-8}) | (3.82 × 10^{-8}) | (1.37 × 10^{-7}) |
| $\hat{p}_1$ | 0.2521 (0.0008) | 0.2524 (0.0004) | 0.2524 (0.0008) | 0.2500 (0.0004) | 0.2513 (0.0004) | 0.2486 (0.0037) |
|           | 0.2153 (0.0008) | 0.2086 (0.0003) | 0.2069 (0.0007) | 0.2136 (0.0003) | 0.2079 (0.0003) | 0.2047 (0.0022) |
| $\hat{p}_3$ | 0.1879 (0.0007) | 0.2002 (0.0003) | 0.1930 (0.0008) | 0.1891 (0.0002) | 0.1995 (0.0003) | 0.1912 (0.0018) |

| Female    |    |      |                 |    |       |                 |
|-----------|----|------|-----------------|----|-------|-----------------|
| $\hat{\pi}_1$ | 0.0105 | 0.0099 | 0.0102 | 0.0110 | 0.0095 | 0.0112 |
|           | (2.46 × 10^{-4}) | (2.34 × 10^{-4}) | (2.52 × 10^{-3}) | (2.77 × 10^{-6}) | (2.63 × 10^{-6}) | (1.26 × 10^{-6}) |
| $\hat{\pi}_2$ | 0.0073 | 0.0066 | 0.0072 | 0.0070 | 0.0060 | 0.0067 |
|           | (3.78 × 10^{-4}) | (2.76 × 10^{-4}) | (3.96 × 10^{-3}) | (4.45 × 10^{-6}) | (3.11 × 10^{-6}) | (1.98 × 10^{-6}) |
| $\hat{\pi}_3$ | 0.0032 | 0.0028 | 0.0034 | 0.0037 | 0.0033 | 0.0030 |
|           | (1.22 × 10^{-4}) | (5.96 × 10^{-4}) | (1.52 × 10^{-3}) | (1.58 × 10^{-6}) | (6.71 × 10^{-7}) | (7.62 × 10^{-7}) |
| $\hat{\pi}_4$ | 0.0081 | 0.0077 | 0.0081 | 0.0075 | 0.0071 | 0.0089 |
|           | (1.85 × 10^{-4}) | (1.63 × 10^{-3}) | (1.99 × 10^{-3}) | (2.18 × 10^{-6}) | (1.83 × 10^{-6}) | (9.96 × 10^{-7}) |
| $\hat{p}_1$ | 0.2422 (0.0007) | 0.2433 (0.0007) | 0.2436 (0.0007) | 0.2342 (0.0016) | 0.2353 (0.0018) | 0.2402 (0.0033) |
|           | 0.2131 (0.0011) | 0.2031 (0.0009) | 0.2051 (0.0011) | 0.2070 (0.0086) | 0.1981 (0.0073) | 0.2027 (0.0023) |
| $\hat{p}_3$ | 0.2506 (0.0012) | 0.2527 (0.0011) | 0.2498 (0.0014) | 0.2382 (0.0013) | 0.2438 (0.0013) | 0.2457 (0.0040) |

where

\[
p_1 = E_0^{-1} p_2 E_0^{-1} p_3 E_0^{-1} (1 - p_1 - p_2 - p_3) H_0^{-1},
\]

\[
j_{11} = s_1 + c_1, j_{12} = s_2 + c_2, j_{13} = s_3 + c_3, j_{14} = s_4 + c_4, K_{11}
\]

\[
= nt - st - ut + \sum_{l=1}^{s_1} y_{1l} + d_1,
\]

\[
K_{12} = ut - vt + \sum_{l=1}^{s_1} y_{2l} + d_2, K_{13} = vt - wt + \sum_{l=1}^{s_1} y_{3l} + d_3, K_{14}
\]

\[
= wt - \sum_{l=1}^{s_1} y_{4l} + d_4,
\]

\[
E_{01} = n - s - u + s_1 + e_1, F_{01} = u - v + s_2 + e_2, G_{01}
\]

\[
= v - w + s_3 + e_3, H_{01} = w + s_4 + e_4,
\]

\[
Q_1 = \left| \frac{1}{j_{11}} \right| \left| \frac{1}{j_{12}} \right| \left| \frac{1}{j_{13}} \right| \left| \frac{1}{j_{14}} \right| \sum_{u=0}^{n-s} \sum_{v=0}^{u} \sum_{w=0}^{v} \left( \frac{n-s}{u} \right) \left( \frac{u}{v} \right)
\]

\[
\cdot \left( \frac{v}{w} \right) \left( \frac{k_{11}^{-1} K_{12}^{-1} K_{13}^{-1} K_{14}^{-1}}{B(E_{01}, F_{01}, G_{01}, H_{01})} \right).
\]

\[
(10)
\]
2.3. The Posterior Distribution Using NIP. According to Jeffrey [20], Jeffrey’s prior is defined as
\[ p(\pi_m) \propto \sqrt{I(\pi_m)}, \quad m = 1, 2, 3, 4, \]  
(11)
where \( I(\pi_m) = -E[(\partial^2 f(y | \pi)) / \partial \pi^2_m] \) is Fisher’s information matrix.

It is assumed that parameters \( \pi_1, \pi_2, \pi_3, \) and \( \pi_4 \) follow Jeffrey’s prior while mixing proportions \( p_1, p_2, \) and \( p_3 \) assume a uniform prior over an interval (0, 1). Thus, the joint prior distribution of parameters is given by
\[ \delta(p | y) \propto \frac{1}{\pi_1 \pi_2 \pi_3 \pi_4}, \quad \pi_1, \pi_2, \pi_3, \pi_4 > 0, \quad p_1, p_2, p_3 > 0, \quad p_1 + p_2 + p_3 \leq 1. \]  
(12)

So, the posterior distribution of \( \pi_1, \pi_2, \pi_3, \pi_4, p_1, p_2, \) and \( p_3 \) jointly using the JP is
\[ g_3(p | y) = Q_3 \left[ \sum_{u=0}^{c} \sum_{w=0}^{c} \left( \frac{n-s}{u} \right) \left( \frac{u}{v} \right)^{n_1 - n_1 - n_1 - n_1} \exp \left[ (-\pi_1 K_{31}) \exp (-\pi_2 K_{32}) \exp (-\pi_3 K_{33}) \exp (-\pi_4 K_{34}) \right] \right] \]
\[ \cdot \left( 1 - p_1 - p_2 - p_3 \right)^{H_{03}-1}, \]  
(13)
where
\[ J_{21} = s_2, \quad J_{22} = s_2, \quad J_{23} = s_3, \quad J_{24} = s_4, \quad K_{21} = nt - st - ut + \sum_{i=1}^{s_1} y_{i1}, \quad K_{22} = ut - vt + \sum_{i=1}^{s_1} y_{i2}, \quad K_{23} = vt - ut + \sum_{i=1}^{s_1} y_{i3}, \quad K_{24} = wt + \sum_{i=1}^{s_1} y_{i4}, \quad E_01 = n - s - u + s_1 + 1, \quad G_01 = n - w + s_1 + 1, \quad H_{01} = w + s_1 + 1, \quad Q_2 = \left[ (J_{21}) (J_{22}) (J_{23}) (J_{24}) \right], \]
\[ \sum_{u=0}^{c} \sum_{w=0}^{c} \left( \frac{n-s}{u} \right) \left( \frac{u}{v} \right)^{n_1 - n_1 - n_1 - n_1} \exp \left[ (-\pi_1 K_{31}) \exp (-\pi_2 K_{32}) \exp (-\pi_3 K_{33}) \exp (-\pi_4 K_{34}) \right] \]
\[ \left( 1 - p_1 - p_2 - p_3 \right)^{H_{03}-1}, \]

2.4. The Posterior Distribution Using Jeffrey’s Gamma Prior. The joint prior density of \( \pi_1, \pi_2, \pi_3, \) and \( \pi_4 \) under Jeffrey’s prior is defined earlier in (12).

Now, suppose
\[ \pi_i \sim JGamma(\alpha_i, \beta_i), \quad i = 1, 2, 3, 4, \]  
\[ p_i \sim U(0, 1). \]  
(14)

Considering the independence of priors, we get a joint prior as
\[ \delta_4(p | y) \propto \pi_1^{(n_1-1)} e^{-\beta_1 \pi_2} \pi_2^{(n_1-1)} e^{-\beta_1 \pi_2} \pi_3^{(n_1-1)} e^{-\beta_1 \pi_2} \pi_4^{(n_1-1)} e^{-\beta_1 \pi_2} \pi_4^{n_1}, \]  
(15)

So, the joint posterior distribution of parameters \( \pi_1, \pi_2, \pi_3, \pi_4, p_1, p_2, p_3 \) using JGamma prior is
\[ g_4(p | y) = Q_4 \left[ \sum_{u=0}^{c} \sum_{w=0}^{c} \left( \frac{n-s}{u} \right) \left( \frac{u}{v} \right)^{n_1 - n_1 - n_1 - n_1} \exp \left[ (-\pi_1 K_{31}) \exp (-\pi_2 K_{32}) \exp (-\pi_3 K_{33}) \exp (-\pi_4 K_{34}) \right] \right] \]
\[ \cdot \left( 1 - p_1 - p_2 - p_3 \right)^{H_{03}-1}, \]  
(16)

2.5. Bayes Estimators and Posterior Risks Using JP, JP, and JG under SELF and LINEX

2.5.1. Loss Functions. Let \( \hat{d} \) is the Bayes estimator then \( E(L(\pi, \hat{d})) \) is its posterior risk. Our purpose, in this study, is to check out the properties of derived estimators and look for efficient loss functions using different priors. Two different loss functions, namely, SELF and LINEX, are used to obtain the Bayes estimators and their posterior risks. The SELF is defined as \( L(d, \pi) = (d - \pi)^2 \). The LINEX loss function can be defined as \( L(d, \pi) = \exp (c(\pi - d)) - c(\pi - d) - 1, c \neq 0 \).

One can get a Bayes estimator with associated posterior risk under SELF as \( \hat{d} = E(\pi) \) and \( \tau(d) = E(\pi^2) - \{E(\pi)\}^2 \). Similarly, using LINEX, loss Bayes estimators and posterior risks can be obtained by \( \tilde{\pi} = -1/c \log \text{E}_{\text{post}}(e^{-\tau}) \) and Risk \( (\tilde{\pi}_{\text{LINEX}}) = \text{E}(c(\hat{\pi} - \pi)) + cE(\pi) \). The Bayes estimators and posterior risks under JP, JP, and JG for parameters \( \pi_1, \pi_2, \pi_3, \) and \( \pi_4, p_1, p_2, p_3 \) under SELF and LINEX are obtained as
\[ \hat{\pi}_1 = Q_3 \left[ \sum_{u=0}^{c} \sum_{w=0}^{c} \left( \frac{n-s}{u} \right) \left( \frac{u}{v} \right)^{n_1 - n_1 - n_1 - n_1} \exp \left[ (-\pi_1 K_{31}) \exp (-\pi_2 K_{32}) \exp (-\pi_3 K_{33}) \exp (-\pi_4 K_{34}) \right] \right] \]
\[ K_{31}^{-1}K_{32}^{-1}K_{33}^{-1}K_{34}^{-1}B(E_{03}, F_{03}, G_{03}, H_{03}), \]  
(17)

\[ \hat{\pi}_2 = Q_4 \left[ \sum_{u=0}^{c} \sum_{w=0}^{c} \left( \frac{n-s}{u} \right) \left( \frac{u}{v} \right)^{n_1 - n_1 - n_1 - n_1} \exp \left[ (-\pi_1 K_{31}) \exp (-\pi_2 K_{32}) \exp (-\pi_3 K_{33}) \exp (-\pi_4 K_{34}) \right] \right] \]
\[ K_{31}^{-1}K_{32}^{-1}K_{33}^{-1}K_{34}^{-1}B(E_{03}, F_{03}, G_{03}, H_{03}), \]  
(18)

\[ \hat{\pi}_3 = Q_4 \left[ \sum_{u=0}^{c} \sum_{w=0}^{c} \left( \frac{n-s}{u} \right) \left( \frac{u}{v} \right)^{n_1 - n_1 - n_1 - n_1} \exp \left[ (-\pi_1 K_{31}) \exp (-\pi_2 K_{32}) \exp (-\pi_3 K_{33}) \exp (-\pi_4 K_{34}) \right] \right] \]
\[ K_{31}^{-1}K_{32}^{-1}K_{33}^{-1}K_{34}^{-1}B(E_{03}, F_{03}, G_{03}, H_{03}), \]  
(19)
\[
\pi_{ih} = Q_h \sum_{u=0}^{n-s} \sum_{v=0}^{n} \left( \frac{n-s}{u} \right) \left( \frac{u}{v} \right) \left( \frac{1}{w} \right) \left( J_{h1} \right) \left( J_{h2} \right) \left( J_{h3} \right) \left( J_{h4} \right) + 1
\]

(20)

\[
\hat{p}_{ih} = Q_h \sum_{u=0}^{n-s} \sum_{v=0}^{n} \left( \frac{n-s}{u} \right) \left( \frac{u}{v} \right) \left( \frac{1}{w} \right) \left( J_{h1} \right) \left( J_{h2} \right) \left( J_{h3} \right) \left( J_{h4} \right)
\]

(21)

\[
\hat{p}_{ih} = Q_h \sum_{u=0}^{n-s} \sum_{v=0}^{n} \left( \frac{n-s}{u} \right) \left( \frac{u}{v} \right) \left( \frac{1}{w} \right) \left( J_{h1} \right) \left( J_{h2} \right) \left( J_{h3} \right) \left( J_{h4} \right)
\]

(22)

\[
\pi_h(\bar{\pi}_{h1}) = Q_h \sum_{u=0}^{n-s} \sum_{v=0}^{n} \left( \frac{n-s}{u} \right) \left( \frac{u}{v} \right) \left( \frac{1}{w} \right) \left( J_{h1} \right) \left( J_{h2} \right) \left( J_{h3} \right) \left( J_{h4} \right) + 2
\]

(24)

\[
\pi_h(\bar{\pi}_{h2}) = Q_h \sum_{u=0}^{n-s} \sum_{v=0}^{n} \left( \frac{n-s}{u} \right) \left( \frac{u}{v} \right) \left( \frac{1}{w} \right) \left( J_{h1} \right) \left( J_{h2} \right) \left( J_{h3} \right) \left( J_{h4} \right) + 2
\]

(25)

\[
\pi_h(\bar{\pi}_{h3}) = Q_h \sum_{u=0}^{n-s} \sum_{v=0}^{n} \left( \frac{n-s}{u} \right) \left( \frac{u}{v} \right) \left( \frac{1}{w} \right) \left( J_{h1} \right) \left( J_{h2} \right) \left( J_{h3} \right) \left( J_{h4} \right) + 2
\]

(26)

\[
\pi_h(\bar{\pi}_{h4}) = Q_h \sum_{u=0}^{n-s} \sum_{v=0}^{n} \left( \frac{n-s}{u} \right) \left( \frac{u}{v} \right) \left( \frac{1}{w} \right) \left( J_{h1} \right) \left( J_{h2} \right) \left( J_{h3} \right) \left( J_{h4} \right) + 2
\]

(27)

where \( h = 1 \) for the IP, \( h = 2 \) for the JP, and \( h = 3 \) for the JG. The Bayes estimators and posterior risks using IP, JP, and JG prior under LINEX are also derived and are presented in Appendix A.

3. Results and Discussions

3.1. Simulation Study. Simulated results are obtained for first, second, third, and fourth component densities \( f_1(y; \pi_1), f_2(y; \pi_2), f_3(y; \pi_3) \), and \( f_4(y; \pi_4) \) chosen randomly from the sample of sizes \( p_1, n_2, n_3, \) and \( 1 - p_1 - p_2 - p_3 \) from their true value. Mixing proportions \( \pi_1, \pi_2, \pi_3, \) and \( \pi_4 \) are overestimated under self-assuming IP, JP, and JG, but for LINEX loss function, all estimates are underestimated and relatively close to their true value. Mixing proportion \( \hat{\pi}_1, \hat{\pi}_2, \) and \( \hat{\pi}_3 \) are overestimated for some values and underestimated for few values. It is observed that the performance of LINEX loss function assuming Jeffreys' gamma prior is better because it has less posterior risk when compared with informative and Jeffreys' prior. From the
graphical representation, it is noted that the maximum value from the data lies on the same point which is obtained from table value.

3.2. Data Application. The data was collected by the IARC [21] and is available at GLOBOCAN which is about cancer incidences and mortality. The cancers responsible for the highest incidence in both the genders (total = 148,041) in Pakistani population includes breast (n = 34038, 23%), lip and oral cavity (n = 12761, 8.6%), lung (n = 6800, 4.6%), non-Hodgkin lymphoma (n = 5964, 4%), and colorectum (n = 5335, 3.6%), respectively, whereas the cancers responsible for the highest deaths (total n = 101,113) in Pakistani population includes breast (n = 16232, 16.1%), lip and oral cavity (n = 7266, 7.2%), lung (n = 6013, 5.9%), oesophagus (n = 4748, 4.7%), and non-Hodgkin lymphoma (n = 4374, 4.3%), respectively. This study is aimed at presenting an analysis of the cancer burden in Pakistan by applying it to a 4-component mixture model, consisting of the estimated number of new cancer cases and deaths in 2012 by age groups. Data is classified into 4 components based on age groups as follows: <45 first group, 45-54 second group, 55-64 third group, and >64 fourth group. Necessary calculations thus obtained are as follows:

Real dataset for the mixture of exponential model incidences of male:

\[ T = 1000.25, n_1 = 57, n_2 = 44, n_3 = 43, n_4 = 65, \sum_{i=1}^{n_1} x_{1i} \]

\[ = 8939, \sum_{i=1}^{n_2} x_{2i} = 8846, \sum_{i=1}^{n_3} x_{3i} = 9539, \sum_{i=1}^{n_4} x_{4i} = 14535. \]  

(31)

Real dataset for the mixture of exponential model incidences of female:

\[ T = 1000.05, n_1 = 59, n_2 = 45, n_3 = 44, n_4 = 69, \sum_{i=1}^{n_1} x_{1i} \]

\[ = 11901, \sum_{i=1}^{n_2} x_{2i} = 9589, \sum_{i=1}^{n_3} x_{3i} = 9758, \sum_{i=1}^{n_4} x_{4i} = 8146. \]  

(32)

Real dataset for the mixture of exponential model deaths of male:

\[ T = 500.25, n_1 = 56, n_2 = 44, n_3 = 40, n_4 = 61, \sum_{i=1}^{n_1} x_{1i} \]

\[ = 4838, \sum_{i=1}^{n_2} x_{2i} = 5831, \sum_{i=1}^{n_3} x_{3i} = 6096, \sum_{i=1}^{n_4} x_{4i} = 11388. \]  

(33)

Real dataset for the mixture of exponential model deaths of female:

\[ T = 500.05, n_1 = 56, n_2 = 44, n_3 = 41, n_4 = 68, \sum_{i=1}^{n_1} x_{1i} \]

\[ = 5348, \sum_{i=1}^{n_2} x_{2i} = 5368, \sum_{i=1}^{n_3} x_{3i} = 5981, \sum_{i=1}^{n_4} x_{4i} = 7747. \]  

(34)

The Bayes estimators and posterior risks using IP, JP, and JG under SELF and LINEX loss functions are presented in Tables 3 and 4. The reciprocal values of Bayes estimators are representing the average no. of incidences and deaths by age in the Pakistani male and female population. \( \pi_1 \) refers to the average no. of incidences and deaths in male and female below the age of 44; similarly, \( \pi_2, \pi_3, \pi_4 \) represent the average no. of incidences and deaths in male and female for the age 45-54, 55-64, and above 65, respectively. And it is also noted that the Bayes estimates under LINEX loss function assuming Jeffreys’ prior are more efficient because their posterior risks are less as compared to IP and JG prior.

4. Conclusion

This study is aimed at developing a 4-component mixture model of exponential distribution using type I censoring under SELF and LINEX loss function and IP, JP, and JG priors. The motivation of this study is to show the application of the exponential mixture model to cancer data under the Bayesian paradigm. It is suggested that mixture models can ideally be applied to analyze cancer data. Bayes estimates are found overestimated for some values and underestimated for few values. In a simulation study under SELF, it is noted that Jeffreys’ gamma prior is best because their posterior risks are less as compared to IP and Jeffreys’ prior. In LINEX loss function, Jeffreys’ gamma prior can be preferred as compared to IP and Jeffreys’ prior at censoring time \( T = 1.05 \). The application of 4 components of exponential mixture distribution is presented using cancer data in which incidences and deaths of the male and female population of Pakistan are studied. The values of Bayes estimates (reciprocals) are representing the average no. of new cases by age in the Pakistani male and female population. \( \pi_1 \) represents the average no. of incidences in male and female below the age of 44; similarly, \( \pi_2, \pi_3, \pi_4 \) represent the average no. of incidences in male and female from the age 45-54, 55-64, and above 65, respectively. And it is also noted that the Bayes estimates under LINEX loss function assuming Jeffreys’ prior is more efficient because their posterior risks are less as compare to IP and JG prior.

For the case of the number of deaths, \( \pi_1 \) represents Bayes estimates and the reciprocal of it represents average no. of death in male and female below the age of 44; similarly, the average no. of deaths in male and female from the age 45-54, 55-64, and above 65 are represented by reciprocal of \( \pi_2, \pi_3, \pi_4 \), respectively. The best loss function is found to be the LINEX loss function assuming Jeffreys’ prior for the male population. In the case of the female
population, the best loss function is SELF assuming Jeffreys’ prior.

Appendix

A. Bayes Estimators and Posterior Risks of IP, JP, and JG under LINEX

The Bayes estimator and posterior risk for LINEX loss function are given by

\[ \widehat{\theta}_{\text{Bayes}} = \frac{1}{c} \ln \left[ \sum_{u, v, w} \left( n - s \right) \left( u \right) \left( v \right) \left( w \right) [f(u)v(w)](1) + \beta \right] \]

where \( h \) = 1 for the IP, \( h \) = 2 for the JP, and \( h \) = 3 for the JG.
Data Availability

All the data are available in the paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] M. A. Nagy, “HIF-1 is the commander of gateways to cancer,” Journal of Cancer Science and Therapy, vol. 3, no. 2, pp. 35–40, 2011.
[2] P. Vanita, V. Subrahmanyam, and K. Jhansi, “A short note on cancer,” Journal of Carcinogenesis and Mutagenesis, vol. 2, p. 128, 2011.
[3] V. Anand and D. Koundal, “Computer-assisted diagnosis of thyroid cancer using medical images: a survey,” in Proceedings of ICRIC 2019, pp. 543–559, Cham, 2020.
[4] D. Koundal, R. Vishraj, S. Gupta, and S. Singh, “An automatic ROI extraction technique for thyroid ultrasound image,” in 2015 2nd international conference on recent advances in Engineering & Computational Sciences (RAECS), pp. 1–5, Chandigarh, India, December 2015.
[5] D. Koundal, B. Sharma, and Y. Guo, “Intuitionistic based segmentation of thyroid nodules in ultrasound images,” Computers in Biology and Medicine, vol. 121, article 103776, 2020.
[6] D. M. Titterington, A. F. Smith, and U. E. Makov, Statistical Analysis of Finite Mixture Distributions, John Wiley and Sons, 1985.
[7] B. Everitt and D. J. Hand, Finite Mixture Distributions, Chapman and Hall, London; New York, 1981.
[8] G. J. McLachlan and D. Peel, Finite Mixture Models, John Wiley & Sons, Inc., USA, 2000.
[9] P. McCullagh, “Exponential mixtures and quadratic exponential families,” Biometrika, vol. 81, no. 4, pp. 721–729, 1994.
[10] A. A. Abu-Taleb, M. M. Smadi, and A. J. Alawneh, “Bayes estimation of the lifetime parameters for the exponential distribution,” Journal of Mathematics and Statistics, vol. 3, no. 3, pp. 106–108, 2007.
[11] F. Noor, A. Sajid, B. Shah, M. Zaman, M. Gheisari, and V. Mariappan, “Bayesian estimation and prediction for Burr-Rayleigh mixture model using censored data,” International Journal of Communication Systems, vol. 32, no. 15, 2019.
[12] H. H. Abu-Zinadah, “A study on mixture of exponentiated-pareto and exponential distributions,” Journal of Applied Sciences Research, vol. 6, no. 4, pp. 358–376, 2010.
[13] N. Feroze and M. Aslam, “Bayesian analysis of Burr type X distribution under complete and censored samples,” International Journal of Pure and Applied Sciences and Technology, vol. 11, no. 2, pp. 16–28, 2012.
[14] F. Noor and M. Aslam, “Bayesian inference of the inverse Weibull mixture distribution using type-I censoring,” Journal of Applied Statistics, vol. 40, no. 5, pp. 1076–1089, 2013.
[15] R. K. Tsutakawa, “Estimation of cancer mortality rates: a Bayesian analysis of small frequencies,” Biometrics, vol. 41, no. 1, pp. 69–79, 1985.
[16] P. C. Lambert, J. R. Thompson, C. L. Weston, and P. W. Dickman, “Estimating and modeling the cure fraction in population-based cancer survival analysis,” Biostatistics, vol. 8, no. 3, pp. 576–594, 2007.