Journal of Statistics Applications & Probability
An International Journal

Extended Beta Distribution and Mixture Distributions with applications to Bayesian analysis

Jamal A. Al-Saleh[1] and Satish K. Agarwal[2]

[1]*Department of Statistics and Operation Research, Faculty of Science, Kuwait University, P.O. Box 5969, Safat, Kuwait*
Email Address: alsalehj@kuc01.kuniv.edu.kw
[2]*Department of Mathematics, College of Science, University of Bahrain, P.O. Box 32038, Bahrain*
Email Address: satishagarwal123@yahoo.com

Received: 26 Mar. 2012, Revised: 17 Jun. 2012, Accepted: 9 Jul. 2012
Published online: 1 Mar. 2013

Abstract: An extended form of *beta* distribution by Al-Saleh and Agarwal, is further extended which has an additional two shape parameters k and l. Introduction of new shape parameters help to express *extended beta* distribution not only as a mixture of distributions, but also provides extra flexibility to the density function over the interval [0,1]. Certain statistical properties such as the r-th moment are defined explicitly. Some of the shapes of family of the densities are also illustrated for different k and l so that it may help the Bayesians to approximate a wide range of prior beliefs among the members of the suggested extended family. The Bayesian analysis for the posterior of an uncertain parameter for the *Bernoulli* process using *extended beta* prior is also considered with an application of mortality rates in 12 hospitals performing surgery on babies.

Keywords: Bayesian analysis; extended beta distribution; finite mixture distribution; Gibbs sampling; Markov chain Monte Carlo; prior distribution.

1 Introduction

It is well known that by introducing new parameter(s), the generalization of statistical distributions such as *gamma* and *beta* distributions can be defined. In literature, many generalization of *beta* distribution are considered. The main contributors are Pham-Gia and Doung [1], Volodin [2], Armero, and Bayarri [3], McDonald and Xu [4], Wilfling [5], Gordy [6], and Parker [7]. In recent years generalization of *beta* distribution by introducing more parameters are considered by, Barreto-Souza at el, [8], Mahmoudi [9], and Singla at el [10].

The extended beta distribution of Al-Saleh and Agarawal [11] can also be written as a finite mixture of *beta* distributions that provides a flexible skewed density over the interval [0,1].
The finite mixture of *beta* distributions receive attention, in practice, quite often when we come across a situation where it is not possible to approximate a prior belief by a member of the available family of distributions on [0,1]. In such cases, the only suitable choice is the mixture of *beta* distributions [MacLachlan and Basford [12], Titterington et al. [13]].

$$f(p) = \sum_{i=1}^{n} w_i f_i(p), \tag{1}$$

where $\sum_{i=1}^{n} w_i = 1$, and the f_i's are *beta* distributions. For the sake of clarity, one of the uses of mixture *beta* distributions can be illustrated through a problem given by Diaconis and YIvisaker [14], that requires a mixture of *beta* distributions as a prior. The problem is as follows: Diaconis and YIvisaker observed that there is a big difference between spinning a coin on a table and tossing it in the air. While tossing often led to about an even proportion of an 'heads' and 'tails', spinning often led to proportion like 1/3 or 2/3. The reason for the bias, according to them, was attributed to the shape of the edge. They considered many possibilities to explain this phenomenon before arriving at a 50:50 mixture (that is $w_1= w_2=1/2$) of two *beta* distributions, one with Be(10,20) and another with Be(20,10) as a reasonable prior. Therefore, the problems of such nature may arise and worth examining. Another situation where mixture of distribution can be used is, when one wants to obtain a posterior and has a complicated prior which not fully dominated by the data. In such cases if the complicated prior is used, it may not only mislead the posterior but also lead to complicated numerical integration if the prior is non conjugate.

Moreover, equation (1) if considered, requires appropriate knowledge of weights $w_1, \ldots, w_n$. If it is not possible to approximate weights reasonably well, then the number of parameters in the prior distribution will increase. In such a situation the Bayesian analysis needed for evaluation of complex integration may lead to more difficult problems. In order to overcome this problem, there is a need to search for a new class of mixture distributions, on the interval between 0 and 1. The present study is an attempt in this direction.

In this paper, a new *extended beta* distribution is considered which can also be written as a finite mixture of *beta* distributions, and it provides a flexible skewed density over the interval [0,1]. This mixture distribution is useful, particularly in applications where the subjective views of information come from two or more sources, and the views are close to each other. Some examples are: the subject expert's assessment, personal viewpoint and decision making of two or more surgeons on performing a surgical operation, etc. The usefulness of proposed finite mixture of *beta* distributions as a conjugate prior are demonstrated to obtain the posterior distribution, say of uncertain parameter *p* of the *Bernoulli* distribution and *p* follows mixture of *beta* distribution. It not only gives a nice posterior but also provides simplicity of the computations. The specific weights attached to the suggested prior are functions of the parameters and the distribution we define represents a class of mixture of *beta* densities. While equation (1) if used as a prior for *p*, will need the knowledge of weights and hence increase the number of parameters.

In section 2, an *extended beta* distribution and finite mixture of *beta* distributions are derived. The statistical properties such as the rth moments are computed which may be of some interest to Bayesians. The shapes of *extended beta* distribution have also been given in Figure I-II, for certain values of the parameters. These shapes may help in identifying an approximate prior among the family of distributions. In section 3, related distributions such as the *log-extended beta*, *extended Pearson type VI*, and *extended generalized F* distribution are presented. In section 4, its use as a conjugate prior to derive the posterior distribution of uncertain parameter *p* for the *Bernoulli* distribution is obtained and the predictive distribution of the next observation is also given. In section 5, an illustration is given to show the usefulness of extended *beta* and finite mixture of beta distributions as a prior by considering an application of mortality rates in 12 hospitals performing surgery on babies.

2 Extended beta distribution

Agarwal and Al-Saleh [15] defined a *generalized gamma type* distribution with parameters α, β, γ and δ with probability density function:

$$f(x) = \frac{\gamma^{\alpha-\delta} x^{\alpha-1} e^{-\gamma x}}{\Gamma_\delta(\alpha,\beta)\ (x+\beta/\gamma)^\delta};\ x>0, \alpha, \beta, \gamma > 0;\ \delta \in \Re . \tag{2}$$

The equation (2) can be reduced to a two parameter *gamma* distribution when we let $\beta = \gamma = 1$, and is defined as:

$$f(x)=\frac{x^{\alpha-1}e^{-x}}{\Gamma_{\delta}(\alpha,1)\ (x+1)^{\delta}};\ x>0,\alpha>0;\delta\in\Re, \tag{3}$$

where

$$\Gamma_{\delta}(\alpha,1)\ =\ \int_0^{\infty}\frac{x^{\alpha-1}e^{-x}}{(x+1)^{\delta}}dx\ . \tag{4}$$

It can be seen that for $\delta = 0$, equation (3) reduces to ordinary *gamma* distribution. The integral equation (4) can be computed using the transformation concerning the *generalized hypergeometric* function [see Andrews [16], Chapter 9, p 365], and is defined as

$$F([n_1,n_2,..,n_u];[d_1,d_2,..,d_v]:z)=\sum_{k=0}^{\infty}\frac{\left(\prod_{i=1}^{u}\frac{\Gamma(n_i+k)}{\Gamma(n_i)}\right)z^k}{\left(\prod_{i=1}^{v}\frac{\Gamma(d_i+k)}{\Gamma(d_i)}\right)k!}. \tag{5}$$

The equation (4) can be computed numerically as

$$\begin{aligned}\Gamma_{\delta}(\alpha,1)&=\Gamma(\alpha)\Gamma(\delta-\alpha)\Gamma(\delta)^{-1}F([\alpha],[1-\delta+\alpha];1)\\&\quad+\Gamma(\alpha-\delta)F([\delta],[1-\alpha+\delta];1),\end{aligned} \tag{6}$$

for all values of $\alpha>0$, and $\delta\in\Re/\{0\}$. It is to be noted that $F([a],[b];\ 1)=\frac{\Gamma(b)\Gamma(b-a-1)}{\Gamma(b-a)\Gamma(b-1)}$.

To define a new type of finite mixture of *beta* distributions we use (3), by reducing the range of the parameter δ to take only negative integers, $\delta = -k$ (k = 0,1,2, ..n), n is fixed. To do this we first introduce *extended gamma* distribution [$ExGa(\alpha,k)$] with parameters $\alpha > 0$ and k integer, with probability density function (pdf) as

$$f(x)=\frac{x^{\alpha-1}(x+1)^k e^{-x}}{\Gamma_k(\alpha,1)},\quad x>0;\ \alpha>0;\ k=0,1,\ ...,n \tag{7}$$

Here $\Gamma_k(\alpha,1)=\int_0^{\infty}x^{\alpha-1}(x+1)^k e^{-x}dx$. Using the *finite binomial series* $(a+b)^k\ =\ \sum_{i-0}^{k}\binom{k}{i}a^i b^{k-i}$, equation (7) can be redefined as

$$f(x)=\sum_{i=0}^{k}\binom{k}{i}\frac{x^{\alpha+i-1}e^{-x}}{\sum_{j=0}^{k}\binom{k}{j}\int_0^{\infty}x^{\alpha+j-1}e^{-x}dx}$$

$$= \frac{1}{C_{\alpha,k}} \sum_{i=0}^{k} \binom{k}{i} (\alpha)_i StGa(\alpha+i), \tag{8}$$

where $StGa(\alpha+i)$ is the *standard gamma* pdf of X, with parameter $\alpha+i$, and $(\alpha)_j$ is known as "Pochhammer symbol" and

$$C_{\alpha,k} = \sum_{j=0}^{k} \binom{k}{j} (\alpha)_j; \quad (\alpha)_j = \alpha(\alpha+1)\cdots(\alpha+j-1), \text{ and } (\alpha)_0 = 1.$$

Hence the equation (8) is a finite mixture of *gamma* distributions with parameters $\alpha > 0$, and $k = 0,1,\ldots,n$. This equation introduces a class of new mixture of *gamma* distribution to give more flexibility in choice of the mixing k.

Theorem 1. Let X and Y be independent random variables, where X having *gamma type* distribution as defined in equation (7) with parameters α and $k = 0,1,\ldots,n$, and Y having *gamma type* distribution [equation (7)] with parameters β, and $l = 0,1,\ldots,m$. A random variable $U = \frac{X}{X+Y}$ defined on the interval between 0 and 1, is said to have an *extended beta* distribution [$ExBe(\alpha,\beta,k,l)$] with parameters $\alpha, \beta > 0$, $k = 0,1, \ldots ,n$, and $l = 0,1, \ldots ,m$, if its probability density function is:

$$f(u) = \frac{1}{C_{\alpha.k} C_{\beta,l} B(\alpha,\beta)} \sum_{i=0}^{k} \left(\binom{k}{i} \left(\sum_{j=0}^{l} \binom{l}{j} (\alpha+\beta)_{i+j} u^{\alpha+i-1} (1-u)^{\beta+j-1} \right) \right). \tag{9}$$

Proof: Consider the joint pdf of X and Y

$$f(x,y) = \frac{x^{\alpha-1}(x+1)^k e^{-x}}{\Gamma_k(\alpha,1)} \frac{y^{\beta-1}(y+1)^l e^{-y}}{\Gamma_l(\beta,1)}.$$

Defining $X=WU$ and $Y=W(1-U)$ gives the joint pdf of W and U as follows:

$$f(w,u) = \frac{1}{\Gamma_{.k}(\alpha,1)\Gamma_l(\beta,1)} \sum_{i=0}^{k} \left(\binom{k}{i} \left(\sum_{j=0}^{l} \binom{l}{j} u^{\alpha+i-1} (1-u)^{\beta+j-1} w^{\alpha+\beta+i+j-1} e^{-w} \right) \right). \tag{10}$$

Integrating the equation (10) over 'w' gives equation (9).

The rth moment of equation (9) about the origin is:

$$\mu_r' = \frac{1}{C_{\alpha.k} C_{\beta,l} B(\alpha,\beta)} \sum_{i=0}^{k} \left(\binom{k}{i} \left(\sum_{j=0}^{l} \binom{l}{j} (\alpha+\beta)_{i+j} B(\alpha+i+r, \beta+j) \right) \right). \tag{11}$$

Using equation (11) the mean and the variance can be obtained. We should also emphasis that the moment-generating function could be expressed as *confluent hypergeometric* function.

Next, a few special cases of equation (9) for different values of *k* and *l* are considered:

Case 1. For $k=l=0$, the equation. (9) reduces to *conventional beta* distribution.

Case 2. For $k=0$, and $l=1$, then equation (9) reduces to:

$$f(u) = \frac{Be(\alpha,\beta)}{(\alpha+1)}\left[1+(\alpha+\beta)u\right].$$

Case 3. For $k=1$, and $l=0$, then equation (9) reduces to:

$$f(u) = \frac{Be(\alpha,\beta)}{(\beta+1)}\left[1+(\alpha+\beta)(1-u)\right].$$

Case 4. For $k=1$, and $l=1$, then equation (9) reduces to:

$$f(u) = \frac{Be(\alpha,\beta)(\alpha+\beta+1)}{(\beta+1)(\alpha+1)}\left[1+(\alpha+\beta)u(1-u)\right].$$

Case 5. For $k=2$, and $l=0$, then equation (9) reduces to:

$$f(u) = \frac{Be(\alpha,\beta)}{(\beta(\beta+1)+2\beta+1)}\left[1+2(\alpha+\beta)(1-u)+(\alpha+\beta)(\alpha+\beta+1)(1-u)^2\right].$$

Case 6. For $k=0$, and $l=2$, then equation (9) reduces to:

$$f(u) = \frac{Be(\alpha,\beta)}{(\alpha(\alpha+1)+2\alpha+1)}\left[1+2(\alpha+\beta)u+(\alpha+\beta)(\alpha+\beta+1)u^2\right].$$

Where $Be(\alpha,\beta) = \dfrac{u^{\alpha-1}(1-u)^{\beta-1}}{B(\alpha,\beta)}$, is the *conventional beta* probability density function. The Figures 1-3 illustrate the graphs [$f(u)$ eqn (9)] of *conventional beta* [$k=0$, $l=0$] and *extended beta* pdf $f(u;\alpha,\beta,k,l)$ for several choices of α, β, k, and l.

From these graphs it can be seen that for the values $\alpha=0.8$, and $\beta=1.05$ and as k=l increases from 2 to 3 and 4, the graphs tend towards normality and the thickness of tail also reduces. If we increase the value of α and decrease the value of β, [say $\alpha=1.10$, and $\beta= 0.95$], for $k=0$ and by varying the values of l, there is a sharp increase in the shape for u greater than 0.5. These graphs can be helpful not only in getting a general idea of how equation. (9) look like for the different choices of k, and l but also in assuming a conjugate *extended beta* prior for Bayesian analysis.

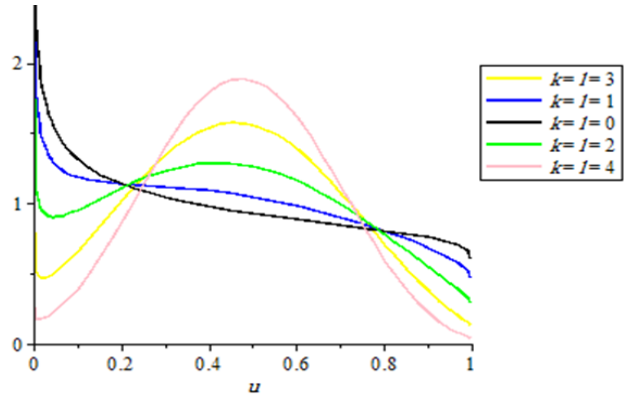

Figure 1: The pdf when $\alpha = 0.8, \beta = 1.05$

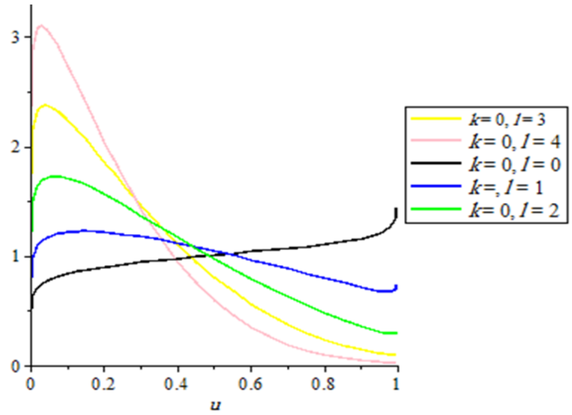

Figure 2: The pdf when $\alpha = 1.10, \beta = 0.95$

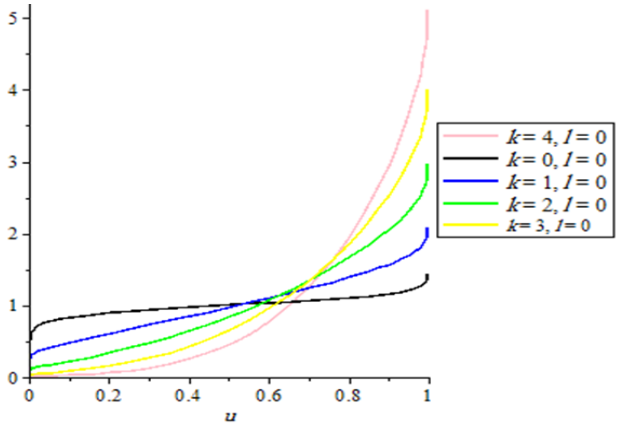

Figure 3: The pdf when $\alpha = 1.10, \beta = 0.95$

3 Related distributions

The moment-generating function of (*-logU*), when *U* has *extended beta* distribution equation (9), is:

$$E[e^{-t\log U}] = \frac{1}{C_{\alpha.k}C_{\beta,l}B(\alpha,\beta)}\sum_{i=0}^{k}\left(\binom{k}{i}\left(\sum_{j=0}^{l}\binom{l}{j}(\alpha+\beta)_{i+j}B(\alpha+i-t,\beta+j)\right)\right). \tag{12}$$

When k=0 and l=0 the distribution of (*-logU*), has been discussed by Barrett, Normand, and Peleg [17]. They suggest the possible use of *log-beta* distributions in place of *log-normal* distributions while fitting the data which come from positively or negatively skewed distribution. The transformation $T = \frac{U}{1-U}$, has the following probability density function

$$f(t) = \frac{1}{C_{\alpha.k}C_{\beta,l}B(\alpha,\beta)}\sum_{i=0}^{k}\left(\binom{k}{i}\left(\sum_{j=0}^{l}\binom{l}{j}(\alpha+\beta)_{i+j}\frac{t^{\alpha+i-1}}{(1+t)^{\alpha+\beta+i+j}}\right)\right),\ t>0. \quad (13)$$

The equation (13) is defined as the *extended Pearson type VI* distribution or a finite mixture of *pearson type VI* distributions. For k=0 and l=0 the distribution is *Pearson type VI* distribution, sometimes called a *beta-prime* distribution [see Keeping [18]]. The *extended Pearson type VI* distribution is related to the *extended generalized F* distribution when $t = g^s$, and the probability density function of this distribution is defined as

$$f(g) = \frac{1}{C_{\alpha.k}C_{\beta,l}B(\alpha,\beta)}\sum_{i=0}^{k}\left(\binom{k}{i}\left(\sum_{j=0}^{l}\binom{l}{j}(\alpha+\beta)_{i+j}\frac{|s|\,g^{s(\alpha+i)-1}}{(1+g^s)^{\alpha+\beta+i+j}}\right)\right),\ g>0. \quad (14)$$

which is a finite mixture of *generalized F* distributions. When k=0 and l=0 it is the *generalized F* distribution and, as a result, is often denoted by *GB2* [see McDonald and Richards [19]]. The cumulative distribution function corresponding to the *extended generalized F* density may be expressed in terms of the *confluent hypergeometric* function. The behavior of the hazard rate of the *generalized F* distribution has been examined by McDonald and Richards [19]. Finite mixtures of *generalized F* distributions are considered by McDonald and Butler [20].

4 Posterior distribution

In this section the parameter of interest is the probability p of success in a number of trials, which can result in success or failure. Suppose there is a fixed number of n trials, with x, number of successes such that x~bin(n,p) a *binomial* distribution of index x and parameter p, thus

$$f(x|p) = \binom{n}{x}p^x(1-p)^{n-x} \quad ,\ x=0,1,\ldots,n. \quad (15)$$

If p is unknown, and the prior for p has the form equation (9), then for known, α, β, k, and l the posterior evidently will have the mixture form,

$$f(p|x) = \frac{\binom{n}{x}}{C_{\alpha.k}C_{\beta,l}B(\alpha,\beta)}\sum_{i=0}^{k}\left(\binom{k}{i}\left(\sum_{j=0}^{l}\binom{l}{j}(\alpha+\beta)_{i+j}p^{\alpha+i+x-1}(1-p)^{\beta+j+n-x-1}\right)\right). \quad (16)$$

If we let n = 1, it is evident that equation (16) belongs to the family of mixture distributions which is closed under sampling with respect to a *Bernoulli* likelihood.

The predictive distribution of the next observation y~bin(m, p) a *binomial* distribution of index y and parameter p, after we have the single observation x on top of our previous background information, is

$$f(y|x) = \int_0^1 f(y \mid p) f(p \mid x) dp$$

$$= \frac{\binom{m}{y}}{C_{\alpha.k} C_{\beta,l} B(\alpha,\beta)} \times \sum_{i=0}^{k} \left(\binom{k}{i} \left(\sum_{j=0}^{l} \binom{l}{j} (\alpha+\beta)_{i+j} B(\alpha+i+x+y, \beta+j+n+m-x-y) \right) \right). \quad (17)$$

This distribution will be defined as the *extended beta binomial* distribution, or the *extended Polya type* distribution [which is related to *Polya* distribution when k=l=0, [see Calvin [21]].

5 Application

Spiegelhalter, et al. [22] analyses the data using WinBUGS (page 15) and was specifically used to demonstrate the use of *beta* prior. The model they considered is based on mortality rates during operation in 12 hospitals [Table 1]. This application may give some insight in how the Bayesian methods for making inference about an uncertain *Bernoulli* parameter p on the basis of prior knowledge that behave as $ExBe(\alpha,\beta,k,l)$ and observed data from *Bernoulli* process would work.

Table 1: The mortality rates in 12 hospitals performing surgery in babies

Hospital	Number of Operations	Number of Deaths	Mortality rates
H_1	47	0	0.0
H_2	148	18	0.12162
H_3	119	8	0.06723
H_4	810	46	0.05679
H_5	211	8	0.03791
H_6	196	13	0.06633
H_7	148	9	0.06081
H_8	215	31	0.14419
H_9	207	14	0.06763
H_{10}	97	8	0.08247
H_{11}	256	29	0.11328
H_{12}	360	24	0.06667

Let x_t and n_t be, the number of deaths, and the number of operations preformed in respective hospital t. This problem can be modeled as a binary response variable with true failure probabilities p_t. Thus x_t can follow $bin(p_t, n_t)$ a *binomial* distribution where t=1, ..,12, that is

$$f(x_t|p_t) = \binom{n_t}{x_t} p_t^{x_t} (1-p_t)^{n_t - x_t} \quad , x_t=0,1,\ldots,n_t,\ t=1,..,12. \quad (18)$$

Now, suppose we have no idea about the prior knowledge of the probabilities p_t for each hospital t. Thus we can suggest the following hierarchical models implemented as follows:

Case (i) k and l known

At the first stage and when k and l are known, we assume a prior belief that follows $ExBe(\alpha_t,\beta_t, k,l)$ for the true failure probabilities p_t for each hospital t. At the second stage, we will assume the following prior specification for the hyperparameters α_{tt} and β_t; α_t~$gamma(\vartheta,\Lambda)$, and β_t~$gamma(\vartheta,\Lambda)$ independent, where ϑ~$exponential(T)$, Λ~$exponential(T)$, T=0.1, and for the values [k=l=0], [k=0,l=1], [k=1, l=1], and [k=2,l=0].

Case (ii) k and l unknown

At the first stage we assume a prior belief that follows $ExBe(\alpha_t,\beta_t,k,l)$ for the true failure probabilities p_t for each hospital t. At the second stage, we will assume the following prior specification for the hyperparameters α_t, β_t; k_t, and l_t, for t=1,…,12, α_t~$gamma(\vartheta,\Lambda)$, β_t~$gamma(\vartheta,\Lambda)$, k_t~$binomial(q_1,100)$, and l_t~$binomial(q_2,100)$ independent, where ϑ~$exponential(T)$, Λ~$exponential(T)$, and q_i~$beta(0.5,1.5)$, i=1.2, and T=0.1.

A Markov Chain Monte Carlo (*MCMC*) *Gibbs* sampling approach implemented in using BUGS[@] computer software can give an analysis of estimates of surgical mortality in each hospital t. A burn in of 1000 updates followed by a further 12000 updates give estimates of $p_{t,}$ for each hospital t=1,…,12, for the case k and l known for different k, and l [Tables 2-5], and for the case when k and l unknown [Table 6].

Examination of the above simulations [tables 2-6] the following observations are noticed:

1. For known differing values of k and l, there is no significant difference in the estimates of p_t, t = 1,2,…..,12.
2. The estimate p_1 for unknown k and l (table 6), there is a dramatic shift to the left in the posterior mean = 0.006345 from the posterior mean of the cases k and l known (table 2-5), which is closer to the true mortality rate of H_1 [table 1]. The posterior standard deviation (SD) remains more or less the same, and a slight decrease in the MC error.
3. For the estimates p_t, t=2,…,12, for k and l unknown (table 6), there is a small shift to the left in the posterior mean from the posterior mean of the cases k and l known (table 2-5), which is closer to the true mortality rate for each hospital. The posterior SD all remained more or less the same, and a slight increase in the MC error as for the cases k known (table 2-5).
4. For the estimates k_t, for t=1,…,12, the posterior means are 8.484, 9.761, 9.541, 9.402, 9.366, 9.592, 9.529, 9.519, 9.548, 9.591, 9.716, 9.5 (prior mean = 25), the posterior SD's about 2.88 (prior SD = 1.37), with the MC errors about 0.04. The posterior mean has a big shift to the left from the prior mean.
5. For the estimate l_t, t=1,…,12, the posterior means are 11.34, 10.47, 10.44, 10.35, 10.5, 10.5, 10.4, 10.2, 10.42, 10.39, 10.41, 10.38 (prior mean = 25), the posterior SD's around 2.9 (prior SD = 1.37), with the MC errors around 0.04. The posterior mean has a big shift to the left from the prior mean.

In brief, the value of the posterior distribution means stay more or less the same across the results for k and l known. For most of the hospitals, the posterior distribution means when k and l are unknown shift slightly to the left closer to the true mortality rates. However, the shift for hospital H_1 is dramatic and is much closer to the true mortality rate compared with the case when k and l are known. Hence, in the above example, when k and l are unknown, the analysis using *extended beta* prior distribution to estimate the mortality rates are more successful compared to [k=0, and l=0, *Beta* prior]. The proposed class of prior distributions offers more flexibility for Bayesian methods to choose among the existing classes of priors.

Table 2: The estimates of surgical mortality for each Hospital when k=0, l=0

Variable	Mean	SD	MC error	2.5%	Median	97.5%
p_1	0.02032	0.01989	4.95E-4	5.62E-4	0.01443	0.07396
p_2	0.1267	0.02674	2.634E-4	0.07905	0.125	0.1849
p_3	0.07441	0.02414	2.263E-4	0.03388	0.07197	0.1285
p_4	0.0579	0.008179	9.032E-5	0.04291	0.05751	0.07502
p_5	0.0421	0.01378	1.434E-4	0.01993	0.04048	0.0736
p_6	0.07082	0.01805	1.862E-4	0.03966	0.06951	0.1095
p_7	0.06681	0.02036	2.058E-4	0.03209	0.06505	0.1116
p_8	0.1475	0.02409	2.136E-4	0.1034	0.1465	0.1973
p_9	0.07181	0.01781	1.89E-4	0.04072	0.0703	0.1099
p_{10}	0.09129	0.02921	2.638E-4	0.04221	0.08845	0.1553
p_{11}	0.1164	0.01996	2.042E-4	0.08016	0.1155	0.1579
p_{12}	0.06925	0.01323	1.434E-4	0.04554	0.06839	0.09723

Table 3: The estimates of surgical mortality for each Hospital when k=0, l=1

Variable	Mean	SD	MC error	2.5%	Median	97.5%
p_1	0.02112	0.02033	4.455E-4	9.553E-5	0.01508	0.0743
p_2	0.1239	0.02688	2.914E-4	0.07645	0.1222	0.1813
p_3	0.07132	0.02345	2.372E-4	0.03273	0.06899	0.124
p_4	0.0575	0.008145	9.156E-5	0.04282	0.0571	0.07448
p_5	0.04021	0.01343	1.727E-4	0.01839	0.03871	0.07064
p_6	0.06832	0.01769	2.079E-4	0.03777	0.06688	0.1066
p_7	0.06403	0.01986	2.352E-4	0.0311	0.06214	0.1086
p_8	0.1458	0.02399	2.843E-4	0.1017	0.1448	0.1951
p_9	0.07005	0.01756	1.849E-4	0.03927	0.06875	0.1073
p_{10}	0.08703	0.02866	3.376E-4	0.03989	0.08455	0.1512
p_{11}	0.1152	0.02001	2.297E-4	0.07919	0.1137	0.1568
p_{12}	0.06778	0.01314	1.498E-4	0.04469	0.06705	0.09576

Table 4: The estimates of surgical mortality for each Hospital when k=1, l=1

Variable	Mean	SD	MC error	2.5%	Median	97.5%
p_1	0.02163	0.02067	4.463E-4	9.553E-5	0.01559	0.07481
p_2	0.1122	0.02387	2.854E-4	0.07027	0.1107	0.1628
p_3	0.06862	0.02024	2.364E-4	0.03413	0.06691	0.1142
p_4	0.05764	0.00801	7.749E-5	0.0426	0.05734	0.07394
p_5	0.04279	0.01325	1.822E-4	0.02029	0.04167	0.0717
p_6	0.06753	0.01654	1.411E-4	0.03874	0.06645	0.1033
p_7	0.06364	0.01792	1.947E-4	0.0324	0.06239	0.1027
p_8	0.1344	0.02221	2.169E-4	0.09396	0.1332	0.1816
p_9	0.06911	0.01627	1.93E-4	0.04074	0.0679	0.1042
p_{10}	0.08061	0.02383	3.065E-4	0.04024	0.07849	0.133
p_{11}	0.1078	0.01839	1.641E-4	0.07551	0.1066	0.1464
p_{12}	0.06765	0.01245	1.278E-4	0.04548	0.06694	0.09375

Table 5 : The estimates of surgical mortality for each Hospital when k=2, l=0

Variable	Mean	SD	MC error	2.5%	Median	97.5%
p_1	0.01082	0.01705	6.013E-4	1.65E-13	0.003435	0.06138
p_2	0.1288	0.02806	5.784E-4	0.07864	0.127	0.1886
p_3	0.07664	0.0257	9.5E-4	0.03489	0.07364	0.1345
p_4	0.06484	0.008224	8.61E-5	0.04959	0.06454	0.08165
p_5	0.06797	0.01478	1.646E-4	0.04249	0.06685	0.1003
p_6	0.09154	0.01846	1.82E-4	0.05903	0.09042	0.1315
p_7	0.09352	0.02106	2.259E-4	0.05736	0.09162	0.1393
p_8	0.1501	0.02479	4.518E-4	0.1046	0.149	0.2007
p_9	0.07183	0.01805	1.736E-4	0.04063	0.07048	0.1109
p_{10}	0.09176	0.02949	6.967E-4	0.04262	0.08899	0.1578
p_{11}	0.1206	0.02075	6.052E-4	0.0823	0.1197	0.163
p_{12}	0.06884	0.0133	1.288E-4	0.04499	0.06796	0.09699

Table 6: The estimates of surgical mortality for each Hospital when k and l unknown

Variable	Mean	SD	MC error	2.5%	Median	97.5%
p_1	0.006345	0.01213	3.871E-4	1.43E-13	0.001162	0.0418
p_2	0.1252	0.02723	3.07E-4	0.077	0.1236	0.1829
p_3	0.07058	0.02335	2.377E-4	0.03191	0.0683	0.1222
p_4	0.05734	0.008094	9.768E-5	0.04232	0.05701	0.07418
p_5	0.03994	0.01348	1.592E-4	0.01777	0.0384	0.06952
p_6	0.06852	0.01823	1.778E-4	0.03729	0.06705	0.1083
p_7	0.0638	0.02016	2.133E-4	0.03012	0.06201	0.1083
p_8	0.1457	0.02382	2.286E-4	0.1025	0.1443	0.1958
p_9	0.06955	0.01775	1.769E-4	0.03889	0.06816	0.1073
p_{10}	0.08662	0.02842	3.124E-4	0.03981	0.08334	0.1492
p_{11}	0.1152	0.02013	2.294E-4	0.07882	0.114	0.1577
p_{12}	0.06768	0.01314	1.214E-4	0.04438	0.067	0.09574

6 Conclusions

The *extended beta* distribution, which can also be expressed as a mixture of *beta* distributions, is used as a conjugate prior. In certain situations it has an advantage over its competitors in the sense that it doesn't require subjective approach of guessing mixing weights. The posterior of uncertain parameter for the *Bernoulli* distribution using the proposed mixture of *beta* prior, is studied by using Markov Chain Monte Carlo (*MCMC*), *Gibbs* sampling approach, on hierarchical models. Another advantage of this prior is that one can use it as a two or three or four parameter mixture of *beta* densities [In literature the mixture distribution requires too much information such as guess weights, besides guess values of the parameters]. Further advantage of the proposed prior is that it gives more flexibility to the users due to the fact that the cases of differing values of k and l [including k=l=0, *conventional beta*] and also the case of unknown k and l are illustrated with the help of an example of real life data set. The illustration gives some idea on how the Bayesian methods for making inference about an uncertain *Bernoulli* parameter, on the basis of prior knowledge that behaves as finite mixture of *beta* distributions, and observed data from *Bernoulli* process, would work. The results reveal that the estimate p_i values remain more or less the same across the results for known k and l. For hospital H_2-H_{12} for unknown k and l, the values of the estimates are closer to the true mortality rates than with the case k and l known, while for hospital H_1 the difference is dramatic. This difference is more evident in the case of k and l unknown than for k and l known. This shows the greater scope of proposed prior.

Acknowledgments

The support by the Research Administration, Kuwait University, Kuwait, is acknowledged by the first author.

References

[1] T. Pham-Gia, Q. P. Duong, The generalized beta and F distributions in statistical modeling, *J. of Mathematical and Computer Modeling*, 13, 1613-1625, (1989).

[2] N. Volodin, Some generalizations of the beta distribution, (1994), *Private Communication.*

[3] C. Armero, M. Bayarri, Prior assessments for prediction in queues, *The Statistician,* 43, 139-153, (1994) .

[4] J.B. McDonald, Y.J. Xu, A generalization of the beta distribution with applications. *Journal of Econometrics*, 69; 2 ,427-428, (1995) .

[5] B. Wilfling, Lorenz ordering of generalized beta income distributions. *Journal of Econometrics*, 71 (1996) 381-388.

[6] M. B. Gordy, A generalization of generalized beta distributions. Finance and Economics Discussion Series 1998-18, *Board of Governors of the Federal Reserve System (U.S.A).*

[7] S. C. Parker, The generalized beta as a model for the distribution of earnings. *Economics Letters*, 62, 197-200, (1999).

[8] W. Barreto-Souza, A.H.S. Santos, G.M. Cordeiro. The beta generalized exponential distribution. *Journal of Statistical Computation and Simulation*. 80, 159-172, (2010).

[9] E. Mahmoudi. The Beta Generalized Pareto distribution with application to lifetime data

Mathematics and computers in Simulation, 81; 11, 2414-2430, (2011) .

[10] N. Singla, K. Jain, S. K. Sharma. The Beta Generalized Weibull distribution: Properties and applications. *Reliability Engineering & System Safety*. 102, 5-15, (2012) .

[11] J. A. Al-Saleh, S. K. Agarwal, Finite mixture of certain distributions. *Communication in Statistics Theory and Method.* 31; 12, 2123 – 2137, (2002) .

[12] G. J. Maclachlan, K. Basford, *Mixture models*, (1988) *Marcel Dekker, New York.*

[13] D. M. Titterington, A. F. M. Smith, U. E. Makov, *Statistical Analysis of Mixture Distributions,* (1985) *Wiley, Chichester, London.*

[14] P. Diaconis, D. Yivisaker Conjugate priors for exponential families, *Annals of Statistics*, 17, 269-281, (1979).

[15] S. K. Agarwal, J. A. Al-Saleh, Generalized gamma type distribution and its hazard rate function. *Communication in Statistics Theory and Method*. 30; 2 , 309-318, (2001).

[16] L. C. Andrews, Special Functions of Mathematics for Engineers (1998). *Oxford University Press, Oxford, UK.*

[17] A. M. Barrett, M. D. Normand, M. Peleg, "Log beta" vs. the log-normal distribution for particle populations with a wide finite size range, *Powder Technology*, 66, 195-199, (1991).

[18] E. S. Keeping, *Introduction to Statistical Inference, New York :* (1962) *Van Nostrand.*

[19] J. B. McDonald, D. O. Richards, Hazard rates and gcneralized beta distribution, *IEEE Transactions on Reliability*, R~36 ,463-466, (1987).

[20] J. B. McDonald, R. J. Buller, Some generalized mixture distributions with an application to unemployment duration, *The Review of Economics and Statistics*, 69, 232-240, (1987).

[21] T. W. Calvin, How and when to perform Bayesian acceptance sampling *(ASQC Basic References in Quality Control: Statistical Techniques, Volume 7 (1984)) Milwaukee, WI: American Society for Quality Control.*

[22] D. J. Spiegelhalter, A. Thomas, N. G. Best, W. R. Gilks, *BUGS Examples Volume 1, Version 0.5, (version ii)* (1996). *Cambridge: Medical Research Council Biostatistics Unit*