Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v6n2p175

A comparative study of loss functions for bayesian control in mixture models

By Hasan, Ali, Khan

Published: October 14, 2013

Electronic Journal of Applied Statistical Analysis
Vol. 06, Issue 02, 2013, 175-185
DOI: 10.1285/i20705948v6n2p175

A comparative study of loss functions for bayesian control in mixture models

Taha Hasan*[a], Sajid Ali[b], and Muhammad Farid Khan[c]

[a,c] *Department of Statistics, Quaid-i-Azam University, Islamabad, 45320, Pakistan*
[b] *Department of Decisions Sciences, Bocconi University, Milan, Italy*

Published: October 14, 2013

Berliner (1987) discussed the issue of controlling the output (response) towards the specified value by choosing the values for independent variables in a regression mixture model, taking it as a Bayesian Decision Problem. The quantification of the potential loss was done with the help of quadratic loss function, which was a symmetric loss function. We have tried to quantify this loss with the help of Precautionary Loss Function and Modified Squared Error Loss Function, in linear Scheffé (1958) mixture model and comparison is established between these loss function. Results are improved as compared to Berliner (1987).

keywords: Bayesian control in mixture, design of mixture experiments, Modified squared error loss function, optimization, Posterior risk, Precautionary loss function.

1 Introduction

A statistical control problem for a regression model is to control the output (response) by selecting the values of the independent variables. Berliner (1987) studied the problem of statistical control in mixture models and posed it as Bayesian decision problem. For optimization, Berliner (1987) used only quadratic loss function. In this article, we have tried to use some alternative loss functions, specially an asymmetric loss functions. A comparative study is established to compare the loss functions for the Bayesian control in mixture experiments. With the study we would be able to control a response value or the target value by choosing the values of mixture components, for which the risk is minimum.

*Corresponding authors: taha.qau@gmail.com

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

For a mixture experiments with q components the proportion of ingredients may be denoted by $z_1, z_2, \cdots, z_q$ where $z_i \geq 0$ for $i = 1, 2, \cdots, q$ and $z_1 + z_2 + \cdots + z_q = 1$. The response depends only on the mixture and not on the total amount of mixture. The factor space is a $(q-1)$-dimensional regular simplex S_{q-1} i.e. $S_{q-1} = \{z : (z_1, z_2, \cdots, z_q) | \sum_{i=1}^{q} z_i = 1, z_i \geq 0\}$. A first order polynomial in q components has *(q+1)* terms. The expected regression model is given as,

$$E(y) = \alpha_0 + \alpha_1 z_1 + \alpha_2 z_2 + \cdots + \alpha_q z_q \tag{1}$$

Since " T is the best Target" situation is considered, so Berliner (1987) calls it a "control" problem. Let y denote the deviations of the response (Y) from the target value T. Due to the constraint $\sum_{i=1}^{q} z_i = 1$, one term must be deleted to obtain a full-rank mixture model, which is a model in which there are no linear dependencies among the repressors and as a result parameters will be unique. Hence following Scheffé (1963), Eq.[1] can be rewritten as,

$$E(y) = \alpha_0 \sum_{i=1}^{q} z_i + \sum_{i=1}^{q} z_i \alpha_i z_i = \sum_{i=1}^{q} (\alpha_0 + \alpha_i) z_i = \sum_{i=1}^{q} \theta_i z_i \tag{2}$$

where $\theta_i = \alpha_0 + \alpha_i$. The preliminary experiments provide n independent observations of y. In matrix form the model becomes,

$$\mathbf{y} = \mathbf{z}\boldsymbol{\theta} + \boldsymbol{\epsilon} \tag{3}$$

where $\mathbf{y} = (y_1, y_2, \cdots, y_n)^{'}$, $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_q)$, $\mathbf{z}$ is $n \times q$ is the design matrix, $\boldsymbol{\epsilon}$ is the $n \times 1$ vector of errors which is assumed to follow multivariate normal distribution with the mean $\mathbf{0} = (0, 0, \cdots, 0)^{'}$ and the variance-covariance matrix $\sigma^2 \boldsymbol{I}_q$. The least square estimator of $\boldsymbol{\theta}$ is $\hat{\boldsymbol{\theta}} = (\mathbf{z}^{'}\mathbf{z})^{-1}\mathbf{z}\mathbf{y}$ and $\hat{\boldsymbol{\theta}} \sim \mathbf{N}(\boldsymbol{\theta}, \sigma^2(\mathbf{z}^{'}\mathbf{z})^{-1})$. Now the optimization problem posed by Berliner is to minimize the square of the future observation $\mathbf{y} = \boldsymbol{\theta}^{'}\mathbf{x} + \epsilon$ where $\mathbf{x} = (x_1, x_2, \cdots, x_q)^{'}$ such that $0 \leq x_i \leq 1$, $\sum_{i=1}^{q} x_i = 1$ and $\epsilon \sim N(0, \eta^2)$. The error term of the future observation is assumed to be independent of $\boldsymbol{\theta}$. For decision rule the expected loss or risk is quadratic,

$$E(\mathbf{y}^{'}\mathbf{y}) = E\{(\boldsymbol{\theta}\mathbf{x})^{'}(\boldsymbol{\theta}\mathbf{x})\} + \eta^2 \tag{4}$$

and the loss $L(\boldsymbol{\theta}, \mathbf{x}) = (\boldsymbol{\theta} - \mathbf{x})^{'}(\boldsymbol{\theta} - \mathbf{x})$. Berger (1985) (pp. 158-163) provided comprehensive study about the Bayesian approach to decision problems.

1.1 Bayesian Control Procedure

The philosophy of Bayesian statistics is based on the updating of belief in the light of prior information about the unknown quantity in which we are interested. We start with a probability distribution reecting our current state of knowledge. When new data become available, we modernize our probability distribution in light of the new data. In a probability framework, there is only one way to do this: via Bayes theorem. The theory of a prior distribution is not much famous in statistics. This controversy is closely linked to the debate regarding the meaning of probability. Some statisticians believe that a prior distribution can be chosen for the parameter in every statistics problem. They believe that this distribution is a subjective probability distribution in the sense that it represents an individual experimenter's information and subjective beliefs about where the true value of parameter is likely to lie. They also believe, however, that a prior distribution is no different from any other probability distribution used in the field of statistics and that all the rules of probability theory apply to a prior distribution. We combine our current information with the prior information and the resulting distribution is called posterior distribution. For more detail discussion about Bayesian philosophy see Berger (1985). Let $\pi(\boldsymbol{\theta}, \sigma^2)$ denote a prior density for the unknown parameters $(\boldsymbol{\theta}, \sigma^2)$. The posterior density for the unknown parameters $(\boldsymbol{\theta}, \sigma^2)$ is given by,

$$\pi(\boldsymbol{\theta}, \sigma^2|\mathbf{x}) = \frac{f(\mathbf{x}|\boldsymbol{\theta}, \sigma^2)\pi(\boldsymbol{\theta}, \sigma^2)}{\int\int f(\mathbf{x}|\boldsymbol{\theta}, \sigma^2)\pi(\boldsymbol{\theta}, \sigma^2)d\boldsymbol{\theta}d\sigma^2} \tag{5}$$

The marginal posterior of $\boldsymbol{\theta}$ is given by $\pi(\boldsymbol{\theta}|\mathbf{x}) = \int \pi(\boldsymbol{\theta}, \sigma^2|\mathbf{x})d\sigma^2$ with the posterior mean vector $\mu(\boldsymbol{\theta}) = \boldsymbol{\theta}_{LS}$ and the posterior Var-Cov matrix $C(\boldsymbol{\theta})$. For known value of σ^2, $C(\boldsymbol{\theta}) = \sigma^2(\mathbf{z}^{'}\mathbf{z})^{-1}$

1.2 Optimization Problem in Bayesian Control for Mixture

Suppose that we want to estimates $\boldsymbol{\theta}$ which is unknown from the realizations $\mathbf{x} \in \Re$. To evaluate the significance of estimator we assign a loss $L(\boldsymbol{\theta}, a) \geq 0$. A loss function $L(\boldsymbol{\theta}, a)$ is said to be symmetric if $L(\boldsymbol{\theta}, a) = L(-\boldsymbol{\theta}, -a)$ otherwise it will be asymmetric. Averaging the loss with respect to the joint probability density function (PDF) $p(\boldsymbol{\theta}, \mathbf{x})$ yields an important characteristic value for an estimator. It is called the Bayes risk/posterior risk (PR), which is $PR = \int\int L(\boldsymbol{\theta}, a(\mathbf{x}))p(\boldsymbol{\theta}, \mathbf{x})d\boldsymbol{\theta}d\mathbf{x}$. Hence the optimal Bayes estimator is an estimator which minimizes the posterior risk i.e. $PR = \underset{a(\mathbf{x})}{\text{argmax}} \int\int L(\boldsymbol{\theta}, a(\mathbf{x}))p(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}d\mathbf{x}$. Suppose that the loss function is differentiable , so taking the derivative with respect to estimator and equating to zero one can get the Bayes estimator and putting this value into expected loss one can get posterior risk. Thus, optimization problem can be stated in terms of posterior expected loss.

$$\text{Minimize} \quad L(.) = \int (\boldsymbol{\theta}\delta)^{'}(\boldsymbol{\theta}\delta)\pi(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta} \tag{6}$$

subject to the constraint $x_i \geq 0 (i = 1, 2, \cdots, q)$ and $\mathbf{x}^{'}\mathbf{1} = \mathbf{1}$ where $\mathbf{1} = (1, 1, \cdots, 1)^{'}$. Since $(\boldsymbol{\theta}\mathbf{x})^2 = \mathbf{x}^{'}(\boldsymbol{\theta}^{'}\boldsymbol{\theta})\mathbf{x}$, we have $L(\hat{\boldsymbol{\theta}}) = \mathbf{x}^{'}\mathbf{B}(\hat{\boldsymbol{\theta}})\mathbf{x}$ where

$$\mathbf{B}(\hat{\boldsymbol{\theta}}) = C(\hat{\boldsymbol{\theta}}) + \mu(\hat{\boldsymbol{\theta}})\mu(\hat{\boldsymbol{\theta}})^{'} \tag{7}$$

Hence, we wish to solve the optimization problem with the objective function,

$$\begin{aligned} &\text{Minimize} \quad \mathbf{x}^{'}\mathbf{B}\mathbf{x} \\ &\text{such that} \quad x_i \geq 0 \quad \text{and} \quad \mathbf{x}^{'}\mathbf{1} = \mathbf{1} \end{aligned} \tag{8}$$

It is a quadratic programming problem which is a branch of mathematical programming involving the minimization or maximization of a continuous real quadratic objective function subject to linear constraints. The quadratic objective function has a unique solution if **B** is positive definite; see Boot (1964)(pp.23-25).

1.3 An example for Bayesian Control in Mixture

To illustrate the control problem Berliner (1987) used an example given by Snee (1981). This example is also given in Castillo (2007) (pp.355-359). Snee (1981) studied the design of mixture experiments concerning gasoline blending. The objective of Gasoline Blending study was to develop a blending model for five-component system with the constraints given in Table [1]. Here the response variable y = research octane at *2.0*

Table 1: Constraints for Gasoline Blending

Components	Range
z_1 =Butane	0.0 − 0.50
z_2 =Isopentane	0.0 − 0.30
z_3 =Reformate	0.0 − 0.35
z_4 =Cat cracked	0.0 − 0.60
z_5 =Alkylate	0.0 − 0.60

grams of lead per gallon. The data in Table [2] is reproduced from Snee (1981) (pp.124). Berliner (1987) focused on a remark given by Snee (1981) i.e. "These components were being used to make a premium grade gasoline, and it was desired that the blends had octanes in the *97-101* range" (pp.123). The standard deviation of the octane are known and taken as *0.25* and *0.4*. The desired octane target value is taken here as $T = 99$. The prior distribution for the unknown parameters $(\boldsymbol{\theta}, \sigma^2)$ is based upon the knowledge of the chemical process involved in gasoline mixing. Berliner (1987) considered only a uniform generalized prior, $\pi(\boldsymbol{\theta}) = 1$ on the regression coefficients. However, he avoided the specification of a prior for σ^2 and simply computed the Bayes rules for various fixed values of σ^2. So, for the given example with the uniform prior on $\boldsymbol{\theta}$, the Bayes rule proved to be insensitive to the prior on σ^2. Using the optimization problem stated in Eq. [8], the future observations obtained for different values of σ^2, with the respective risk values are given in Table [3]. The table is reproduced from Berliner (1987) (pp.459) for Squared Error/ Quadratic Loss Function (SELF/QLF).

Table 2: Gasoline Blending Data

z_1	z_2	z_3	z_4	z_5	Y
0.000	0.000	0.350	0.600	0.060	100.0
0.000	0.300	0.100	0.000	0.600	101.0
0.000	0.300	0.000	0.100	0.600	100.0
0.150	0.150	0.100	0.600	0.000	97.3
0.150	0.000	0.150	0.600	0.100	97.8
0.000	0.300	0.049	0.600	0.051	96.7
0.000	0.300	0.000	0.489	0.211	97.0
0.150	0.127	0.023	0.600	0.100	97.3
0.150	0.000	0.311	0.539	0.000	99.7
0.000	0.300	0.285	0.415	0.000	99.8
0.000	0.080	0.350	0.570	0.000	100.0
0.150	0.150	0.266	0.018	0.600	101.9
0.150	0.150	0.082	0.100	0.600	100.7
0.000	0.000	0.300	0.461	0.239	100.9
0.150	0.034	0.116	0.100	0.600	101.2
0.068	0.121	0.175	0.444	0.192	98.7
0.067	0.098	0.234	0.332	0.270	100.5
0.000	0.300	0.192	0.208	0.300	100.2
0.150	0.150	0.174	0.226	0.300	100.6
0.075	0.225	0.278	0.424	0.000	99.1
0.075	0.225	0.000	0.100	0.600	100.4
0.000	0.126	0.174	0.600	0.100	98.4
0.075	0.000	0.225	0.600	0.100	98.2
0.150	0.150	0.000	0.324	0.376	99.4
0.000	0.300	0.192	0.508	0.000	98.8

2 Bayesian Control for Mixtures under Precautionary and Modified Loss functions

The QLF is a symmetric loss function which allows estimating zero value. This means no risk is expected. But in some cases it is not good to under estimate the potentiality of an event as to overestimate it. Generally in risk analysis we investigate potentiality

Table 3: Bayes Control Rules under SELF/QLF

σ^2	T	Risk Value	x_1	x_2	x_3	x_4	x_5
	97	0.0104	0.0675	0.1872	0.0934	0.6519	0.0000
0.0625	99	0.0026	0.0660	0.1579	0.1554	0.4345	0.1844
	100	0.0058	0.0648	0.1395	0.1970	0.1953	0.4034
	97	0.0625	0.0675	0.1870	0.0939	0.6516	0.0000
0.1600	99	0.0068	0.0660	0.1598	0.1554	0.4343	0.1845
	100	0.0147	0.0648	0.1395	0.1969	0.1957	0.4031

and effects of unwanted events. The probability or a failure rate is used to measure potentiality. Bayes approach is used to estimate this failure rate. The two loss functions are defined below and used in the problem of Bayesian control for mixtures.

2.1 Precautionary Loss Function

In this section, we will consider Bayesian estimation with other loss functions. This dilemma is extremely significant for realistic applications as the loss function should replicate the cost that is associated with a certain estimation error. The following two examples illustrate this more clearly: Consider the problem of constructing a dam. Underestimating the peak water level from older measurements is clearly more serious than overestimating it and this fact should be reflected in the choice of the loss function $L(\theta, a)$. This illustration motivates the use of an asymmetric loss function and it is obvious that the use of asymmetric loss functions as quadratic loss function is not suited for such an estimation problem. Another example that gives rise to other loss functions can be found in the field of image processing. Traditionally, the mean squared error is used to evaluate images and therefore many algorithms are optimized for this loss function. The problem with the MSE is that it does not well represent the human perception. Images which have a small mean squared error may still look very different and therefore it is suggested to use other distance measures. One is the structural similarity (SSIM) index, for more details see Uhlich and Yang (2012) and the references cited therein. However, calculating the optimal Bayesian estimator (OBE) for many non-standard loss functions is not trivial and it can often only be stated in terms of an optimization problem which has to be solved.

Norstram (2012) defined asymmetric loss function called Precautionary loss function (PLF). According to Norstram (2012) these loss functions approach infinitely near the origin to avoid under estimation. It is much useful especially when under estimation may lead to a serious effect (Ali et al. (2012)). The loss function $L(\theta, a)$ is a Precautionary loss function iff:

1. $L(\theta, a)$ is downside damaging.

2. For each fixed θ, $L(\theta, a) \to \infty$ when $a \to \infty$

A loss function is downside damaging if $L(\theta, a - \epsilon) \geq L(\theta, a + \epsilon)$ (for any $\epsilon > 0$) and a is any estimate of θ. The loss function $L(\theta, a) = \frac{(\theta - a)^2}{a}$ satisfies the criteria of being precautionary and its Bayes estimate is easy to calculate. Now, if we consider that $y = (Y - T)/\sqrt{T}$ in Eq. [2] then Eq. [4] becomes:

$$E\left(\frac{Y - T}{\sqrt{T}}\right)^2 = E(\theta x + \epsilon)^{'}(\theta x + \epsilon) \tag{9}$$

The left hand side of Eq. [9] is the PLF and at the right hand side is the SELF. Hence, from Eq. [9] the minimization of the PLF $L(Y, T)$ is equal to the minimization of SELF $L(\theta, x) = E\{(\theta x)^{'}(\theta, x)\}$. So, we use the optimization problem given in Eq. [8]. Using the example given in section 1, we first compute the OLS estimates under PLF as given below:

$$\hat{\theta}_{LS} = \begin{pmatrix} 0.1456 \\ -0.0851 \\ 0.8928 \\ -0.4264 \\ 0.2738 \end{pmatrix}$$

Now $C(\boldsymbol{\theta})$ for $\sigma^2 = 0.0625$ is

$$C(\boldsymbol{\theta}) = \begin{pmatrix} 0.5884 & 0.1113 & 0.0410 & -0.1015 & -0.0923 \\ 0.1113 & 0.2020 & 0.0299 & -0.0641 & -0.0720 \\ 0.0410 & 0.0299 & 0.2113 & -0.0866 & -0.0275 \\ -0.1015 & -0.0641 & -0.0866 & 0.0695 & 0.0284 \\ -0.0923 & -0.0720 & -0.0275 & 0.0284 & 0.0574 \end{pmatrix}$$

Then $B(\hat{\boldsymbol{\theta}}_{LS})$ given in Eq. [8] provides,

$$B(\hat{\boldsymbol{\theta}}_{LS}) = \begin{pmatrix} 0.5886 & 0.1111 & 0.0423 & -0.1022 & -0.0918 \\ 0.1111 & 0.2021 & 0.0292 & -0.0638 & -0.0722 \\ 0.0423 & 0.0292 & 0.2193 & -0.0904 & -0.0251 \\ -0.1022 & -0.0638 & -0.0904 & 0.0713 & 0.0273 \\ -0.0918 & -0.0722 & -0.0251 & 0.0273 & 0.0582 \end{pmatrix}$$

$B(\hat{\boldsymbol{\theta}}_{LS})$ is a positive definite and its unique solution exists. We solve the optimization problem given in Eq. [8] and obtain the future observations which are given below,

$$x^* = \begin{pmatrix} 0.0660 \\ 0.1589 \\ 0.1566 \\ 0.4289 \\ 0.1896 \end{pmatrix}$$

The expected loss or the risk using PLF is *0.0013*. Similarly, the whole process is repeated for $\sigma^2 = 0.16$. We have also tried this control problem for the target values $T = 97, 99, 100$. The Bayes control rules under PLF are given below in Table [4].

Table 4: Bayes Control Rules under PLF

σ^2	T	Risk Value	x_1	x_2	x_3	x_4	x_5
	97	0.0047	0.0664	0.1606	0.1198	0.6302	0.0230
0.0625	99	0.0013	0.0660	0.1589	0.1566	0.4289	0.1896
	100	0.0014	0.0654	0.1502	0.1748	0.3230	0.2866
	97	0.0106	0.0668	0.1737	0.1266	0.6016	0.0313
0.1600	99	0.0064	0.0700	0.1184	0.2326	0.5355	0.0435
	100	0.0035	0.0654	0.1511	0.1730	0.3325	0.2780

2.2 Modified Squared Error Loss Function

The modified squared error loss function (MSELF) may be defined as $L(\theta, a) = \left(\frac{\theta - a}{a}\right)^2$. Next, we use the optimization procedure under MSELF as defined in section 2.1. The optimization problem posed by Berliner is to minimize the expected square of the future observation $y = \boldsymbol{\theta}'\mathbf{x} + \epsilon$ where $y = (Y - T)/T$, $\mathbf{x} = (x_1, x_2, \cdots, x_q)'$. Thus,

$$\min E\left(\frac{Y - T}{T}\right)^2 = E(\theta x + \epsilon)'(\theta x + \epsilon) \tag{10}$$

The left hand side of Eq. [10] is the MSELF and at the right hand side of Eq. [10] is the SELF. Thus, from Eq. [9] the minimization of the MSELF $L(Y, T)$ is equal to the minimization of SELF. Again using the example given in section 1, we first compute the

OLS estimates under the MSELF, with the target value $T = 99$ and $\sigma^2 = 0.0625$.

$$\hat{\theta}_{LS} = \begin{pmatrix} 0.0146 \\ -0.0085 \\ 0.0897 \\ -0.0428 \\ 0.0275 \end{pmatrix}$$

Then $B(\hat{\boldsymbol{\theta}}_{LS})$ given in Eq. [8] provides,

$$B(\hat{\boldsymbol{\theta}}_{LS}) = \begin{pmatrix} 0.5886 & 0.1111 & 0.0423 & -0.1022 & -0.0918 \\ 0.1111 & 0.2021 & 0.0292 & -0.0638 & -0.0722 \\ 0.0423 & 0.0292 & 0.2193 & -0.0904 & -0.0251 \\ -0.1022 & -0.0638 & -0.0904 & 0.0713 & 0.0273 \\ -0.0918 & -0.0722 & -0.0251 & 0.0273 & 0.0582 \end{pmatrix}$$

Again, $B(\hat{\boldsymbol{\theta}}_{LS})$ is a positive definite and its unique solution exists. We solve the optimization problem given in Eq. [8] and obtain the future observations which are given below,

$$x^* = \begin{pmatrix} 0.0658 \\ 0.1556 \\ 0.1636 \\ 0.3882 \\ 0.2268 \end{pmatrix}$$

The expected risk using MSELF is *0.0012*. Again, the whole process is repeated for $\sigma^2 = 0.16$. We have also tried this control problem for the target values $T = 97, 99, 100$. The Bayes control rules under PLF are given below in Table [5].

Table 5: Bayes Control Rules for MSELF

σ^2	T	Risk Value	x_1	x_2	x_3	x_4	x_5
	97	0.0015	0.0659	0.1570	0.1606	0.4057	0.2108
0.0625	99	0.0012	0.0658	0.1556	0.1636	0.3882	0.2268
	100	0.0012	0.0657	0.1550	0.1651	0.3798	0.2344
	97	0.0034	0.0657	0.1562	0.1626	0.3930	0.2225
0.1600	99	0.0031	0.0657	0.1556	0.1638	0.3856	0.2293
	100	0.0031	0.0657	0.1552	0.1644	0.3823	0.2324

3 Discussion and some Final Remarks

The main theme of this article is to use some alternative loss functions in place of QLF, given in Berliner (1987). The SELF, being a symmetric loss function, gives equal weights to over and under estimates. The expected loss obtained through the QLF, for $\sigma^2 = 0.0625$ and with the target value $T = 99$, as given in Castillo (2007), is *0.00263*. When PLF is used, the expected loss reduces to *0.0013* and for MSELF it becomes *0.0012*. So this concludes that if we use asymmetric loss function (PLF) it outperforms SELF. Also, we see that the MSELF is a good replacement for a symmetric SELF, with the reduction in expected loss. Further when we compare the risk values, for two values of σ^2, with the different target values of the Table [3], to the risk values in Table [4] and Table [5], it can be concluded that both, the PLF and MSELF are the better choice for the Bayesian control problem in mixture experiments.

The conclusion about the preference of PLF and MSELF as compared to SELF can

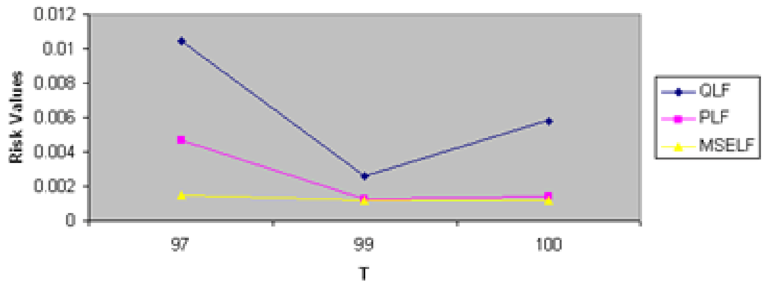

Figure 1: Comparison of Loss Functions for $\sigma^2 = 0.0625$

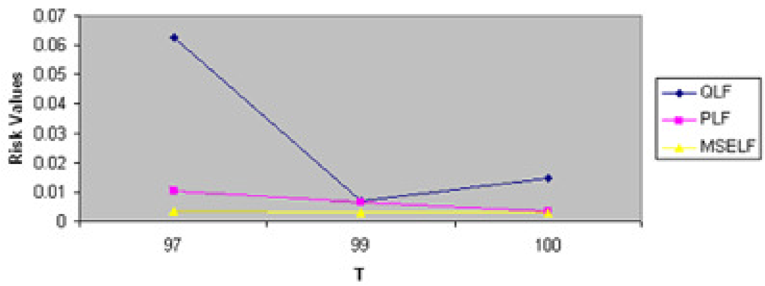

Figure 2: Comparison of Loss Functions for $\sigma^2 = 0.1600$

be verified from Figure [1] and Figure [2]. The results are independent of the choice of data set and also can be verified for other examples. In Figure [1], when the parameter value is small, we get the estimated value of parameter under SELF/QLF has increasing trend but under MSELF (modified quadratic loss function) is smaller than the PLF and

SELF. Similar conclusions can be drawn from Figure [2]. Actually the SELF gives the equal weights to overestimation and underestimation but the other two loss functions prevent this trend. It is assumed the error term in the mixture model follows normal distribution, and hence the response also follows it. We have used the uniform prior for the unknown coefficients of mixture model. But in some mixture experiments the response follows Logistic or Poisson distribution; ultimately the prior distribution will have to be changed for the unknown coefficients of mixture model. As Berliner (1987) avoided the specification of a prior for σ^2 and simply computed the Bayes rules for various fixed values of σ^2 , some future work can be done by using some suitable prior for σ^2. For the quadratic programming solutions we have used MATLAB Ver.7.0.

Aknowledgment

The authors would like to thank the editor and referees who made excellent suggestions that led to considerable improvement in the paper.

References

Ali, S., Aslam, M., Kundu, D., and Kazmi, S. (2012). Mixture of generalized exponential distribution: a versatile lifetime model in industrial engineering processes. *Journal of the Chinese Institute of Industrial Engineers*, 29(4):246–269.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Springer.

Berliner, L. M. (1987). Bayesian control in mixture models. *Technometrics*, 29(4):455–460.

Boot, J. C. G. (1964). *Quadratic Programming*. Amsterdam North Holland.

Castillo, E. D. (2007). *Process optimization: A statistical approach*. Springer-Verlag.

Norstram, J. G. (2012). The use of Precautionary loss functions in Risk Analysis. *IEEE Transaction on Reliability*, 45(3):400–403.

Scheffé, H. (1958). Experiments with mixtures. *Journal of Royal Statistical Society-B*, 20(2):344–360.

Scheffé, H. (1963). Simplex-centroid designs for experiments with mixtures. *Journal of Royal Statistical Society-B*, 25(2):235–263.

Snee, R. D. (1981). Developing blending models for Gasoline and other mixtures. *Technometrics*, 23(2):119–130.

Uhlich, S. and Yang, B. (2012). Bayesian estimation for nonstandard loss functions using a parametric family of estimators. *IEEE Transitions on Signal Processing*, 60(3):1022–1031.