REVSTAT – Statistical Journal
Volume 14, Number 4, October 2016, 433–454

OBJECTIVE BAYESIAN ESTIMATORS FOR THE RIGHT CENSORED RAYLEIGH DISTRIBUTION: EVALUATING THE AL-BAYYATI LOSS FUNCTION

Author: J.T. Ferreira
– Department of Statistics, University of Pretoria,
Pretoria, 0002, South Africa
johan.ferreira@up.ac.za

A. Bekker
– Department of Statistics, University of Pretoria,
Pretoria, 0002, South Africa

M. Arashi
– Department of Statistics, University of Pretoria,
Pretoria, 0002, South Africa
and
Department of Statistics, School of Mathematical Sciences,
Shahrood Unviversity of Technology, Shahrood, Iran

Received: December 2013 Revised: March 2015 Accepted: March 2015

Abstract:

- The Rayleigh distribution, serving as a special case of the Weibull distribution, is known to have wide applications in survival analysis, reliability theory and communication engineering. In this paper, Bayesian estimators (including shrinkage estimators) of the unknown parameter of the censored Rayleigh distribution are derived using the Al-Bayyati loss function, whilst simultaneously considering different objective prior distributions. Comparisons are made between the proposed estimators by calculating the risk functions using simulation studies and an illustrative example.

Key-Words:

- *Al-Bayyati loss; Rayleigh distribution; risk efficiency; shrinkage estimation; squared error loss.*

AMS Subject Classification:

- 62501, 62N10.

1. INTRODUCTION

The Rayleigh distribution is a continuous probability distribution serving as a special case of the well-known Weibull distribution. This distribution has long been considered to have significant applications in fields such as survival analysis, reliability theory and especially communication engineering.

When considering the complete Rayleigh model, the probability density function is given by

$$f(x;\theta) = 2\theta x e^{-\theta x^2}\,, \qquad x,\theta > 0\,, \tag{1.1}$$

using the parametrization of the distribution as proposed by Bhattacharya and Tyagi (1990), and is denoted by $X \sim \mathit{Rayleigh}(\theta)$. The parameter θ is a scale parameter, and characterizes the lifetime of the object under consideration in application.

Mostert (1999) did extensive work concerning the censored model, and showed that the censored Rayleigh model is relatively easy to use compared to other more complex models (such as the Weibull- and compound Rayleigh models). In certain types of applications, it is not uncommon that some observations may cease to be observed due to machine failure, budgetary constraints, and the likes. To compensate for such events, right censored analyses utilizes information only obtained from the first d observations. Thus, the right censored sample consists of n observations, where only d lifetimes (d an integer), $x_1 < x_2 < ... < x_d$ are measured fully, while the remainder $n-d$ are censored. These $n-d$ censored observations are ordered separately and are denoted by $x_{d+1} < x_{d+2} < ... < x_n$. In the context of reliability analysis (for example), a lifetime would be the time until a unit / machine fails to operate successfully.

In the paper of Soliman (2000), a family of non-informative priors were introduced:

$$g(\theta) = \frac{1}{\theta^m}\,, \qquad m,\theta > 0\,, \tag{1.2}$$

and was termed a “quasi-density” prior family. This paper explores the application of this prior family with regards to the right censored Rayleigh model. Different known prior densities are contained within (1.2), namely the Jeffreys prior ($m=1$), Hartigan’s prior ($m=3$), and a third prior illustrating the diminishing effect of the prior density family — this is termed a “vanishing” prior (some large value of m, chosen arbitrarily such that $m=10$). The choice of m would be up to the practitioner to determine the extent of the objectivity required. It is worth noting that Hartigan’s prior ($m=3$) is known as an asymptotically invariant prior as well. Liang (2008) provides valuable contributions when considering relevant choices of hyperparameters.

Mostert (1999) showed that the likelihood of the censored Rayleigh model is given by

$$\mathcal{L}(\theta) \propto (2\theta)^d u e^{-\theta T}$$

where $T = \sum_{i=1}^n x_i^2 \sim gamma(\frac{n}{2}, \theta)$. The quantity u is defined as $u = \prod_{i=1}^{d} x_i$, see Mostert (1999) for further details. It can be shown that the posterior distribution results in

$$g(\theta|T) = \frac{T^{d-m+1}}{\Gamma(d-m+1)} \theta^{(d-m+1)-1} e^{-\theta T} \tag{1.3}$$

which characterizes a $gamma(d-m+1, T)$ distribution, where $\Gamma(\cdot)$ denotes the gamma function. Note that, since the posterior distribution is always a proper distribution, it ensures the need of restrictions on the parameter space. In order for (1.3) to be well-defined, it is thus assumed throughout that $m < d+1$.

Together with Soliman (2000), Mostert (1999) compared the Bayesian estimators under the linear exponential (LINEX) loss function and squared-error loss (SEL) function, and Dey and Dey (2011) did similar work for the complete model by applying Jeffreys prior and a loss function as proposed by Al-Bayyati (2002). This paper extends concepts in the literature for the *censored* Rayleigh model by considering this new loss function, namely the Al-Bayyati loss (ABL), and comparing it to other known results.

Gruber (2004) proposed a method where a balanced loss function is used in a Bayesian context. A balanced loss function is where a weighted *loss* value is constructed by substituting each estimate into its corresponding loss function and determining some weighted value thereof. In this paper an extension of this methodology is considered, by obtaining a new *estimator* as a weighted value of the Bayesian estimator under either SEL or ABL, and some other estimate of the unknown parameter (in this case, θ). This is also known as a shrinkage based estimation approach.

The focus of this paper is the evaluation of the ABL estimator in terms of its performance by considering its risk efficiency in comparison to the SEL estimator, and also the effect of the parameter m, the prior density family degree. In this respect the following proposal is adopted:

1. Obtain the Bayes estimator under SEL, and evaluate under ABL;

2. Obtain the Bayes estimator under ABL, and evaluate under SEL; and

3. Obtain shrinkage estimators of both SEL and ABL estimators by combining the Bayesian estimators with some prespecified point estimate of the parameter, and evaluate under SEL.

In Section 2 the respective Bayesian estimators are determined and the risk (expected loss) are studied comparatively. The effect of risk efficiency is also investigated, and a shrinkage approach is also then considered. In section 3 an illustrative example involving a simulation study and a real data analysis presented, and section 4 contains a discussion and some final conclusions.

2. SQUARED-ERROR LOSS (SEL) & AL-BAYYATI LOSS (ABL)

2.1. Parameter estimation under SEL & ABL

This section explores the Bayesian estimators under the loss functions for the model discussed in the introduction. The SEL is defined by

$$L_{SEL}(\hat{\theta},\theta) = (\hat{\theta}-\theta)^2 \tag{2.1}$$

and the loss function proposed by Al-Bayyati (2002):

$$L_{ABL}(\hat{\theta},\theta) = \theta^c(\hat{\theta}-\theta)^2 \,, \qquad c \in \mathbb{R} \,. \tag{2.2}$$

SEL is a widely used loss function due to its attractive feature of symmetry — where the function focuses on the size of the loss rather than the direction (over- or underestimation) of the loss. The ABL introduces the additional parameter c, which assists in determining a flatter loss function (albeit still symmetric) or the alternative, and it specifically generalizes the SEL (2.1). c can also be considered the order of weighting of the quadratic component. Under SEL, the (posterior) risk function has the following form:

$$\begin{aligned} R_{SEL}(\hat{\theta},\theta) &= \int_0^\infty L_{SEL}(\hat{\theta},\theta)g(\theta|T)d\theta \\ &= \hat{\theta}^2_{SEL} - 2\hat{\theta}_{SEL}\frac{\Gamma(d-m+2)}{\Gamma(d-m+1)T} + \frac{\Gamma(d-m+3)}{\Gamma(d-m+1)T^2} \,. \end{aligned}$$

From (1.3) the Bayesian estimator $\hat{\theta}_{SEL}$ is given by the posterior mean of θ:

$$\hat{\theta}_{SEL} = \frac{d-m+1}{T} \,. \tag{2.3}$$

Since (1.1) indicates that the parameter θ must be positive, a restriction implied by (2.3) is that $m < d+1$ (corresponding to the restriction discussed in the Introduction regarding the posterior distribution). Under ABL, the (posterior) risk function has the following form:

$$\begin{aligned} R_{ABL}(\hat{\theta},\theta) &= \int_0^\infty L_{ABL}(\hat{\theta},\theta)g(\theta|T)d\theta \\ &= \hat{\theta}^2_{ABL}\frac{\Gamma(d-m+c+1)}{\Gamma(d-m+1)T^c} - 2\hat{\theta}_{ABL}\frac{\Gamma(d-m+c+2)}{\Gamma(d-m+1)T^{c+1}} \\ &\quad + \frac{\Gamma(d-m+c+3)}{\Gamma(d-m+1)T^{c+2}} . \end{aligned}$$

The Bayesian estimator $\hat{\theta}_{ABL}$ is

$$\hat{\theta}_{ABL} = \frac{d-m+c+1}{T} . \tag{2.4}$$

Similar to the case of the SEL estimator, $m < d+c+1$ for positive c, and $m+c < d+1$ for negative c in order for the gamma function to be well-defined.

2.2. Comparing the risk of SEL and ABL

The three different prior degrees are of interest here, namely the Jeffreys prior ($m = 1$), Hartigan's prior ($m = 3$), and the vanishing prior ($m = 10$). The posterior risk of the two loss functions was compared against each other for certain parameter values — notably for increasing values of θ and for the three different values of m.

The risk was determined empirically by simulating 5000 samples of sizes $n = 30, 40$ and 50 each, using the inverse-transform method and $uniform\,(0,1)$ random variates. From each of these obtained samples, the parameter was estimated under SEL and ABL (with $c = 0.5$), and the average loss of all 5000 samples was determined. The value of d was set at $d = 0.2n$, which implies that 20% of lifetimes have been observed. There are practical examples were a censoring of between 70% and 90% have been observed (see Stablein, Carter, and Novak (1981)), which is why, as an illustration, a censoring of 80% is used.

In Figures 1 to 3 it is seen that the shape of the functions do not change for different values of m, but it is observed that the risk is increasing for larger m values. Also, as the sample size n increases, the magnitude of the risk is decreasing. From the simulation it is evident that for positive c, SEL has least risk and would thus be preferable. An effective way of comparing the risk of different loss functions is by determining the risk efficiency — which is explored in the next section.

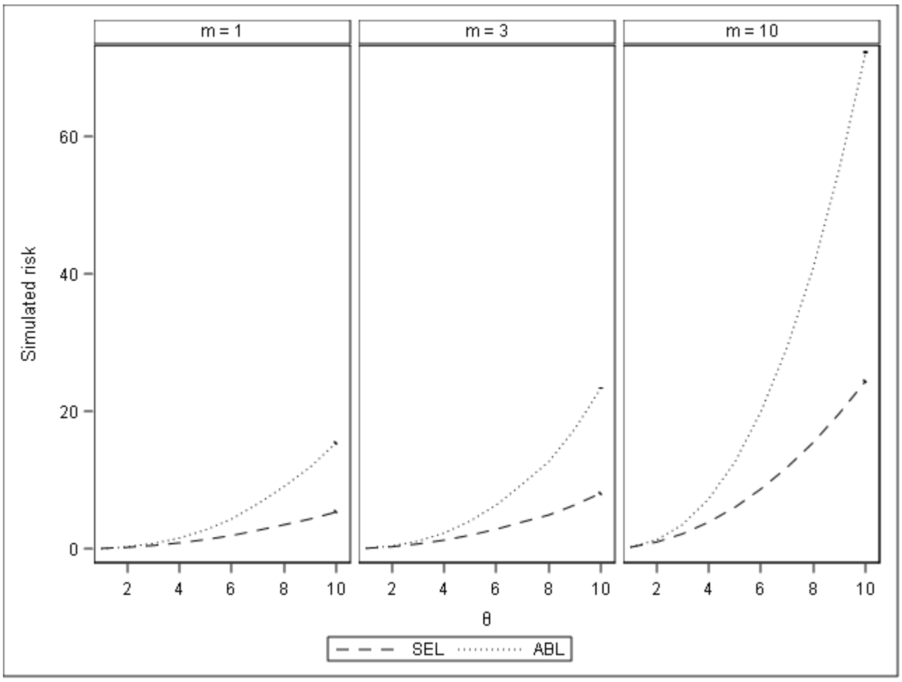

Figure 1: Simulated risk for SEL and ABL ($n = 30$).

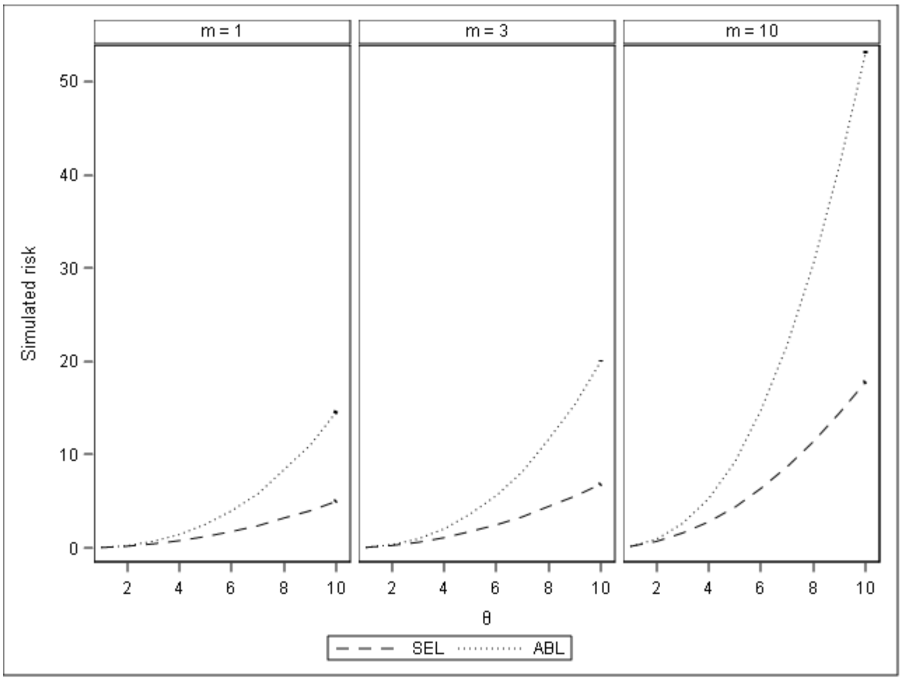

Figure 2: Simulated risk for SEL and ABL ($n = 40$).

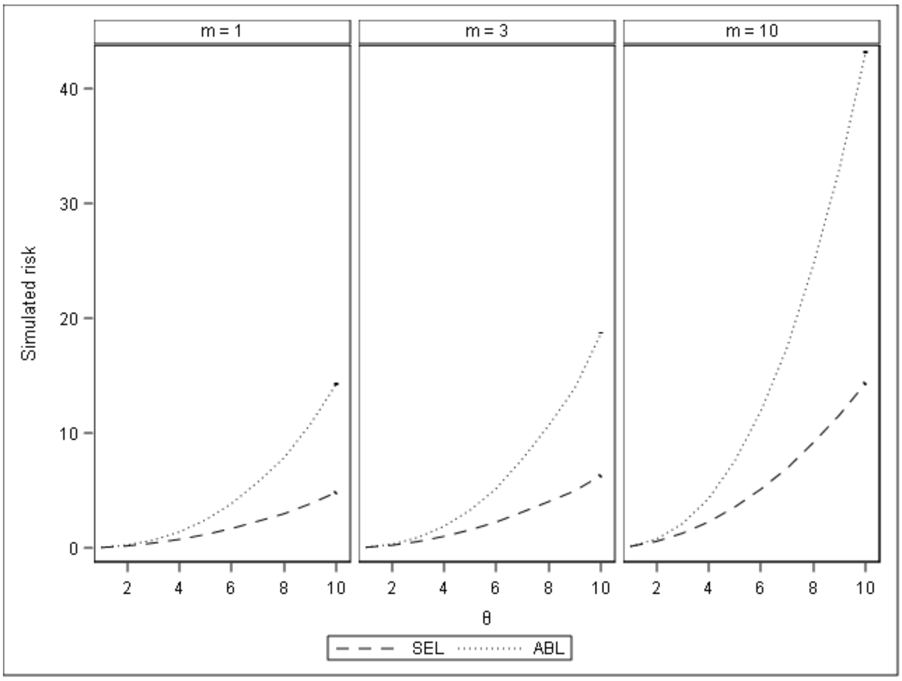

Figure 3: Simulated risk for SEL and ABL ($n = 50$).

2.3. Risk efficiency between SEL and ABL

Risk efficiency is a method that provides an intuitive way of determining which estimator — under a certain loss function — performs better than the other. The form of the risk function considered is

$$R_L^*(\hat{\theta}_{est}, \theta) \;=\; E_T(L(\hat{\theta}_{est}, \theta)) \;=\; \int_0^\infty L(\hat{\theta}_{est}, \theta) f(T) dT$$

using the distribution of T. Here, L denotes the loss function under which the risk efficiency is calculated, and $\hat{\theta}_{est}$ denotes its estimator of θ. The risk efficiency is then given by:

$$RE_L(\hat{\theta}_L, \hat{\theta}_y) \;\equiv\; \frac{R_L^*(\hat{\theta}_y, \theta)}{R_L^*(\hat{\theta}_L, \theta)}$$

translating to, the risk efficiency of $\hat{\theta}_L$ with respect to $\hat{\theta}_y$ under L loss ($\hat{\theta}_y$ denotes an estimator under any other loss function than L). This is similar as the approach by Dey (2011). Now, $\hat{\theta}_L$ denotes the estimator for the parameter that needs to be estimated under loss L, and $\hat{\theta}_y$ denotes the estimator for the parameter under the loss y. The interpretation of this expression is that when $RE_L(\hat{\theta}_L, \hat{\theta}_y) > 1$, the estimator $\hat{\theta}_L$ is preferable under L loss than that of $\hat{\theta}_y$.

2.3.1. SEL vs. ABL under SEL

The risk efficiency for the estimators derived in section 2.1 under SEL are given by:

$$RE_{SEL}(\hat{\theta}_{SEL}, \hat{\theta}_{ABL}) = \frac{R^*_{SEL}(\hat{\theta}_{ABL}, \theta)}{R^*_{SEL}(\hat{\theta}_{SEL}, \theta)} .$$

The expressions required by above equation are obtained as:

$$\begin{aligned} R^*_{SEL}(\hat{\theta}_{ABL}, \theta) &= \int_0^\infty L_{SEL}(\hat{\theta}_{ABL}, \theta) f(T) dT \\ &= \theta^2 \left(\frac{(d-m+1+c)^2}{(\frac{n}{2}-1)(\frac{n}{2}-2)} - 2\frac{d-m+1+c}{(\frac{n}{2}-1)} + 1 \right) \end{aligned}$$

and

$$\begin{aligned} R^*_{SEL}(\hat{\theta}_{SEL}, \theta) &= \int_0^\infty L_{SEL}(\hat{\theta}_{SEL}, \theta) f(T) dT \\ &= \theta^2 \left(\frac{(d-m+1)^2}{(\frac{n}{2}-1)(\frac{n}{2}-2)} - 2\frac{d-m+1}{(\frac{n}{2}-1)} + 1 \right). \end{aligned}$$

The risk efficiency of $\hat{\theta}_{sel}$ with respect to $\hat{\theta}_{abl}$ under SEL is then

$$\begin{aligned} RE_{SEL}(\hat{\theta}_{SEL}, \hat{\theta}_{ABL}) &= \frac{R^*_{SEL}(\hat{\theta}_{ABL}, \theta)}{R^*_{SEL}(\hat{\theta}_{SEL}, \theta)} \\ &= \frac{\left(\frac{(d-m+1+c)^2}{(\frac{n}{2}-1)(\frac{n}{2}-2)} - 2\frac{d-m+1+c}{(\frac{n}{2}-1)} + 1 \right)}{\left(\frac{(d-m+1)^2}{(\frac{n}{2}-1)(\frac{n}{2}-2)} - 2\frac{d-m+1}{(\frac{n}{2}-1)} + 1 \right)} . \end{aligned} \tag{2.5}$$

An interesting characteristic of this equation (2.5) is that it is independent from the sample information i.e. independent of x_i. It is only dependent on n, d, c, and m.

Figure 4 illustrates the risk efficiency (2.5) for arbitrary parameter values. Since the function is not dependent on sample information, no simulation from (1.1) is required. A sample size of $n = 30$ was specified along with $d = 0.2n$ and for different values of c. The risk efficiency values is plotted against values of m, the prior family degree. It is of special interest that for negative values of c, the ABL estimator performs better than that of the SEL counterpart for small values of m. The converse holds when this "threshold" value of m is reached, where the more efficient estimator becomes the SEL estimator.

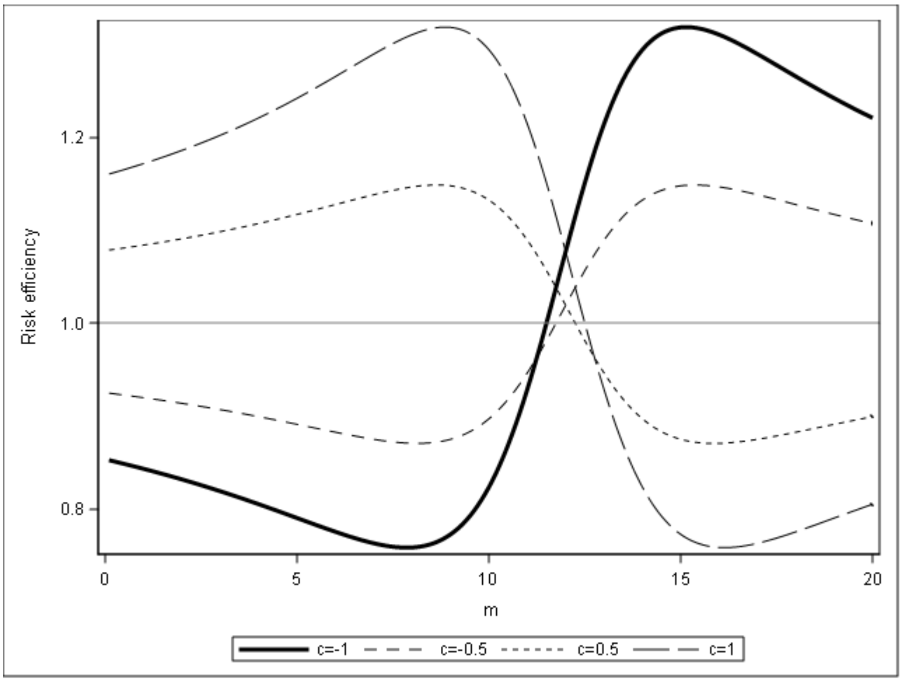

Figure 4: Risk efficiency of SEL- and ABL estimator under SEL.

2.3.2. ABL vs. SEL under ABL

The risk efficiency for SEL and ABL under ABL is given by:

$$RE_{ABL}(\hat{\theta}_{ABL}, \hat{\theta}_{SEL}) = \frac{R^*_{ABL}(\hat{\theta}_{SEL}, \theta)}{R^*_{ABL}(\hat{\theta}_{ABL}, \theta)} .$$

The expressions required by above equation are obtained as:

$$\begin{aligned} R^*_{ABL}(\hat{\theta}_{SEL}, \theta) &= \int_0^\infty L_{ABL}(\hat{\theta}_{SEL}, \theta) f(T) dT \\ &= \theta^{c+2} \left(\frac{(d-m+1)^2}{(\frac{n}{2}-1)(\frac{n}{2}-2)} - 2\frac{d-m+1}{(\frac{n}{2}-1)} + 1 \right) \end{aligned}$$

and

$$\begin{aligned} R^*_{ABL}(\hat{\theta}_{ABL}, \theta) &= \int_0^\infty L_{ABL}(\hat{\theta}_{ABL}, \theta) f(T) dT \\ &= \theta^{c+2} \left(\frac{(d-m+1+c)^2}{(\frac{n}{2}-1)(\frac{n}{2}-2)} - 2\frac{d-m+1+c}{(\frac{n}{2}-1)} + 1 \right) \end{aligned}$$

again using the relations derived in section (2.3.1). The risk efficiency of $\hat{\theta}_{abl}$ versus $\hat{\theta}_{sel}$ under ABL is:

$$
\begin{aligned}
RE_{ABL}(\hat{\theta}_{ABL},\hat{\theta}_{SEL}) &= \frac{R^{*}_{ABL}(\hat{\theta}_{SEL},\theta)}{R^{*}_{ABL}(\hat{\theta}_{ABL},\theta)} \\
&= \frac{\left(\frac{(d-m+1)^2}{(\frac{n}{2}-1)(\frac{n}{2}-2)} - 2\frac{d-m+1}{(\frac{n}{2}-1)} + 1\right)}{\left(\frac{(d-m+1+c)^2}{(\frac{n}{2}-1)(\frac{n}{2}-2)} - 2\frac{d-m+1+c}{(\frac{n}{2}-1)} + 1\right)} .
\end{aligned} \tag{2.6}
$$

It is observed that this last result is the reciprocal of the (2.5). Figure 5 illustrates this result; where the converse of the discussion of (2.5) holds.

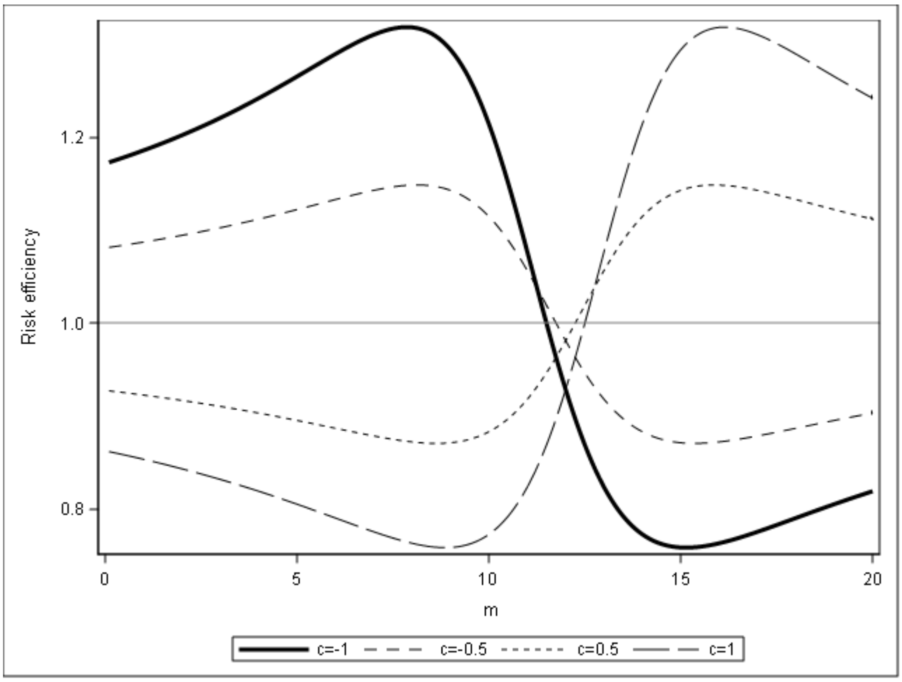

Figure 5: Risk efficiency of SEL- and ABL estimator under ABL.

2.4. Shrinkage estimation approach

Gruber (2004) proposed a method where a balanced loss function is used for Bayesian analysis. A balanced loss function is where a weighted loss value is constructed by substituting each estimate into its corresponding loss function and determining some weighted value thereof. As a slight twist on this approach, consider obtaining a new estimator as a weighted value of the Bayesian estimator under either SEL or ABL, and some other estimate of the unknown parameter (in this case, θ). This is also known as a shrinkage based estimation approach. Define the SEL-based Bayesian shrinkage estimator by

$$\hat{\theta}_{S_1} = \lambda\hat{\theta}_{SEL} + (1-\lambda)\theta_o \,, \qquad 0 \leq \lambda \leq 1 \,, \tag{2.7}$$

and the ABL-based Bayesian shrinkage estimator by

$$\hat{\theta}_{S_2} = \lambda\hat{\theta}_{ABL} + (1-\lambda)\theta_o \,, \qquad 0 \leq \lambda \leq 1 \,. \tag{2.8}$$

where θ_o is a pre-specified point realization of θ. Similar as in the case of the SEL- and ABL estimators, these two newly proposed estimators ((2.7) and (2.8)) is compared in terms of their risk functions. The analysis here is only considered under the SEL. For the SEL-based shrinkage Bayesian estimator we have from (2.1) and (2.7)

$$\begin{aligned} R^*_{SEL}(\hat{\theta}_{S_1},\theta) &= E_T\left(\lambda\hat{\theta}_{SEL} - \lambda\theta + \lambda\theta + (1-\lambda)\theta_0 - \theta\right)^2 \\ &= \lambda^2\left(\theta^2\left(\frac{(d-m+1)^2}{\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)} - \frac{2(d-m+1)}{\left(\frac{n}{2}-1\right)} + 1\right)\right) \\ &\quad + (1-\lambda^2)(\theta_0-\theta)^2 \\ &\quad + 2\lambda(1-\lambda)\left(\theta_0 E_T(\hat{\theta}_{SEL}) - \theta E_T(\hat{\theta}_{SEL}) - \theta\theta_0 + \theta^2\right) \end{aligned}$$

where $E_T(\hat{\theta}_{SEL}) = (d-m+1)\frac{\theta}{\left(\frac{n}{2}-1\right)}$, using the expected value of the gamma distribution of T. The ABL-based shrinkage Bayesian estimator is, from (2.2) and (2.8), given by

$$\begin{aligned} R^*_{SEL}(\hat{\theta}_{S_2},\theta) &= E_T\left(\lambda\hat{\theta}_{ABL} - \lambda\theta + \lambda\theta + (1-\lambda)\theta_0 - \theta\right)^2 \\ &= \lambda^2\left(\theta^2\left(\frac{(d-m+1+c)^2}{\left(\frac{n}{2}-1\right)\left(\frac{n}{2}-2\right)} - \frac{2(d-m+1+c)}{\left(\frac{n}{2}-1\right)} + 1\right)\right) \\ &\quad + (1-\lambda^2)(\theta_0-\theta)^2 \\ &\quad + 2\lambda(1-\lambda)\left(\theta_0 E_T(\hat{\theta}_{ABL}) - \theta E_T(\hat{\theta}_{ABL}) - \theta\theta_0 + \theta^2\right) \end{aligned}$$

where $E_T(\hat{\theta}_{ABL}) = (d-m+1+c)\frac{\theta}{\left(\frac{n}{2}-1\right)}$. When this method is repeated with ABL as the underlying loss functions, similar expressions are obtained but in a scaled form (stemming from the scaling value θ^c from the ABL), and is omitted here.

2.4.1. Risk comparison under SEL and ABL for shrinkage estimators

A similar approach was followed as in Dey (2011) and as discussed in section 2.2, but in this instance the shrinkage estimators were considered with the true risk. Again because of the inferential nature of the ABL, it is only discussed here for the SEL. Two viewpoints were considered: the first of which was for different prior point estimates and for varying λ, and the second was for fixed prior point estimate, different values of m, and for varying λ. This was all considered in the same simulated data setting as in section 2.2, with the addition that the "true" value of θ was 10. An underestimated value, an overestimated value, together with the MLE of θ was considered; i.e. $\theta_0 = 7$, 7.7625, and 15 (here, $\hat{\theta}_{MLE} = \frac{d}{T}$). Figure 6 illustrates the effect of these different prior point estimates and $m = 1$, whilst Figure 7 illustrates for different values of m and the prior point estimate equal to the MLE of the censored Rayleigh distribution. The two figures illustrate these effects.

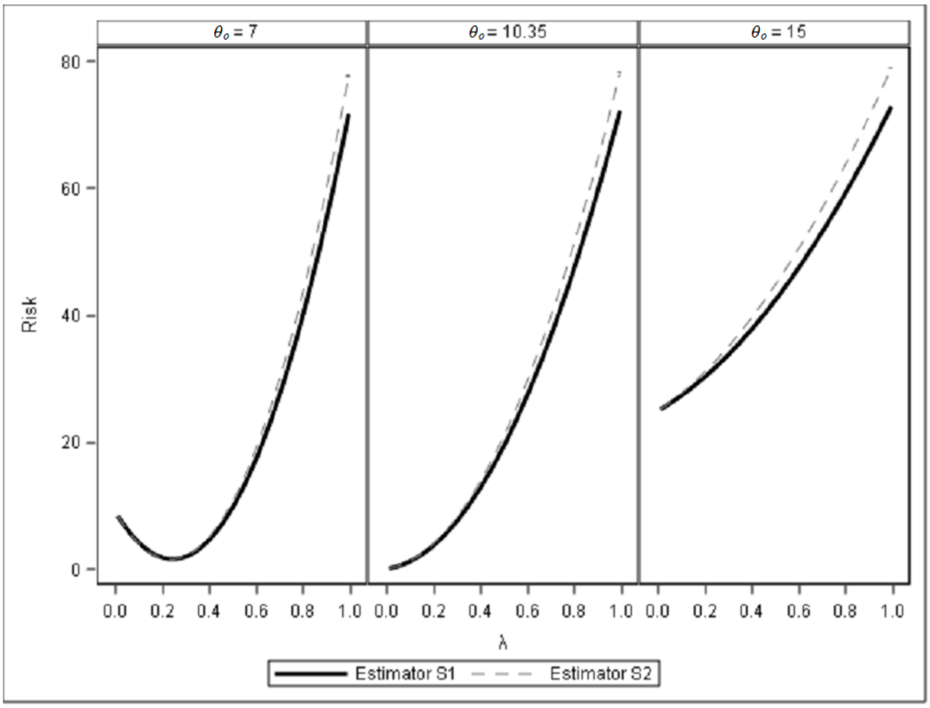

Figure 6: Risk under SEL for shrinkage estimators $\hat{\theta}_{S_1}$ and $\hat{\theta}_{S_2}$, different θ_0, and varying λ ($m = 1$ (fixed)).

As can be seen in both cases, least risk can be obtained for some nonzero and nonunity value of λ, except for the case depicted in Figure 7 when $m = 10$.

This however makes little practical sense if not viewed in comparison with that of the "original" risk for only the Bayesian estimators. In the next section, this comparison is explored with reference to the risk efficiency.

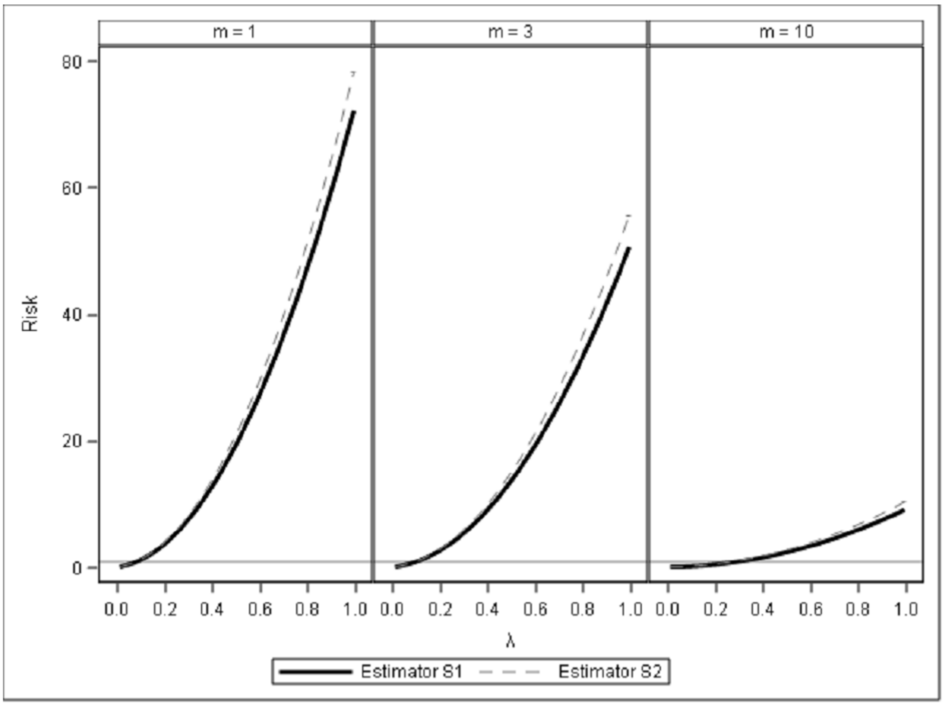

Figure 7: Risk under SEL for shrinkage estimators $\hat{\theta}_{S_1}$ and $\hat{\theta}_{S_2}$, different m, and varying λ ($\theta_0 = MLE$ (fixed)).

2.4.2. Risk efficiency under SEL and ABL for shrinkage estimators

Now, the risk efficiency for the shrinkage estimators was determined under these two loss functions. The following comparisons are considered:

$$RE_{SEL}(\hat{\theta}_{SEL}, \hat{\theta}_{S_1}) = \frac{R^*_{SEL}(\hat{\theta}_{S_1}, \theta)}{R^*_{SEL}(\hat{\theta}_{sel}, \theta)} \tag{2.9}$$

and

$$RE_{ABL}(\hat{\theta}_{ABL}, \hat{\theta}_{S_2}) = \frac{R^*_{ABL}(\hat{\theta}_{S_2}, \theta)}{R^*_{ABL}(\hat{\theta}_{ABL}, \theta)} . \tag{2.10}$$

The same parameter choices as used previously was employed here, and different values of θ_0 were chosen arbitrarily, to assist with the comparison.

The prior density degree was $m = 1$, the Jeffreys prior, and the true value of θ from which the observations were simulated from, is 10. Three values were considered: a value that underestimates the true value of θ, the MLE, and a value that overestimated the true value of θ. Two considerations were examined and is illustrated by the respective figures below. Figure 7 illustrates the risk efficiency under SEL for varying λ, and these different prior point estimates. Figure 8 illustrates the same, but for the case where the underlying loss function is ABL. For these illustrative purposes, the ABL constant c was set to 0.5.

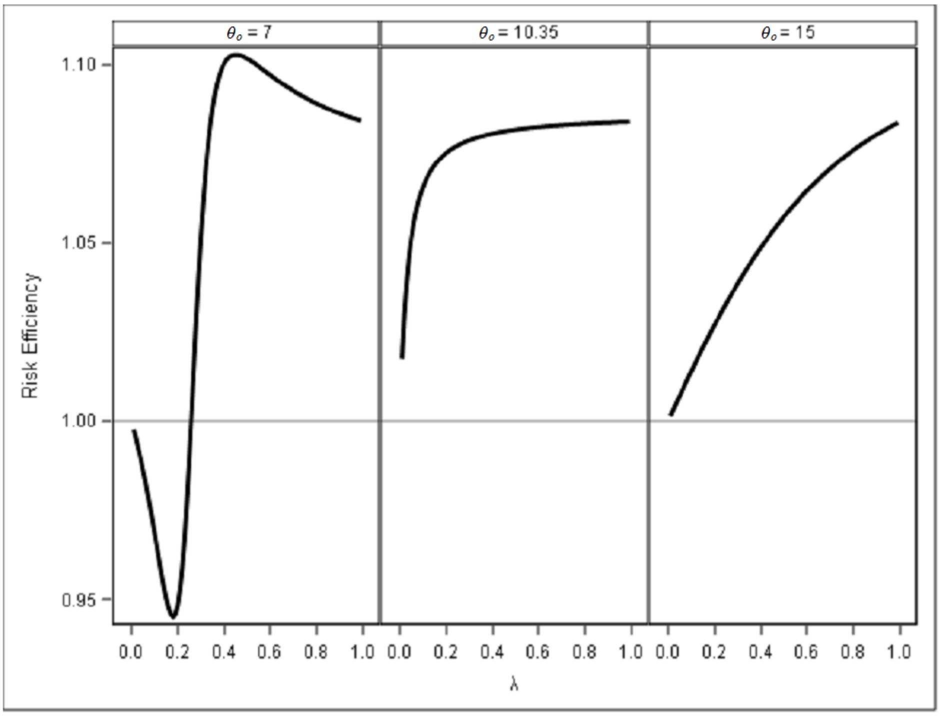

Figure 8: Risk efficiency under SEL for shrinkage estimators $\hat{\theta}_{S_1}$ and $\hat{\theta}_{S_2}$, different θ_0, and varying λ.

Figure 8 clearly shows that there is indeed some shrinkage estimator value (i.e. $0 \leq \lambda \leq 0.25$) that is more appropriate to use than the the true corresponding Bayesian estimator (for a risk efficiency value of < 1). This seems only true for the case of underestimation ($\theta_0 = 7$). For the case of the MLE and overestimation ($\theta_0 = 15$), only the Bayesian estimate seems appropriate. Figure 9 shows the reciprocal results, where the shrinkage estimator seems more appropriate to use in overestimation.

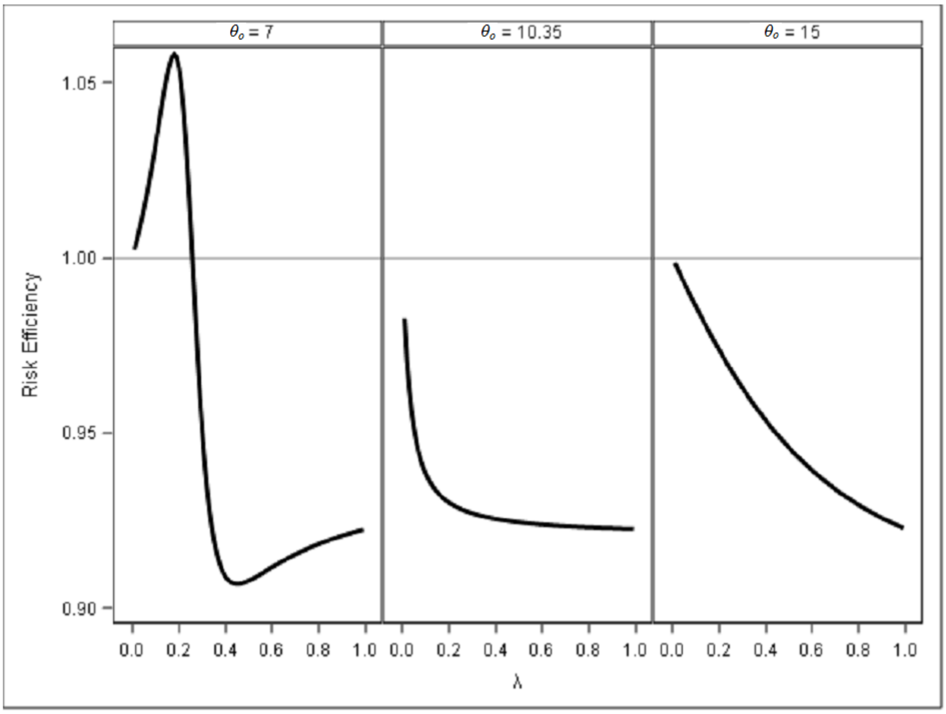

Figure 9: Risk efficiency under ABL for shrinkage estimators $\hat{\theta}_{S_1}$ and $\hat{\theta}_{S_2}$, different θ_0, and varying λ.

3. ILLUSTRATIVE EXAMPLES

3.1. Simulation study

In this section, the $RMSE$ (root mean square error) comparison of the SEL estimator (2.3), the ABL estimator (2.4), and the shrinkage counterparts (2.7) and (2.8) is calculated via simulation. It is known that an estimator with least $RMSE$ is considered preferable. As the parameter θ in (1.1) indicates a lifetime, it is important to use an estimator which estimates the true value of the population parameter as closely as possible, otherwise the chosen estimator may overestimate or underestimate the value too severely, resulting in catastrophic events in real life. For example, when estimating the lifetime of an airplane engine, underestimating the lifetime is much less serious than overestimating the lifetime of the engine. By using the $RMSE$ the estimator which exhibits the smallest error in estimation can be determined.

The $RMSE$ is given by $RMSE = \sqrt{\frac{\sum_{i=1}^{p}\left(\hat{\theta}_{est}-\theta\right)^2}{p}}$, where p denotes the number of observations of θ. $\hat{\theta}_{est}$ denotes the estimated value of θ under a specific loss function. The following steps outline the method followed in this simulation.

1. Simulate $p = 5000$ random samples from (1.1) for a given value of θ. From each simulated sample, determine $\hat{\theta}_{est}$ under SEL, ABL, and both considered shrinkage estimators (for the shrinkage estimators, the value of $\theta_0 = MLE$). Then, calculate the value of the $RMSE$.

2. Repeat Step 1 for a successive range of θ values, in this case, $\theta = 1...40$.

3. Plot the $RMSE$ for all four estimators upon the same set of axis. The estimator with lowest $RMSE$ is considered the preferable estimator.

Figure 10 and 11 shows the results for different choices of λ.

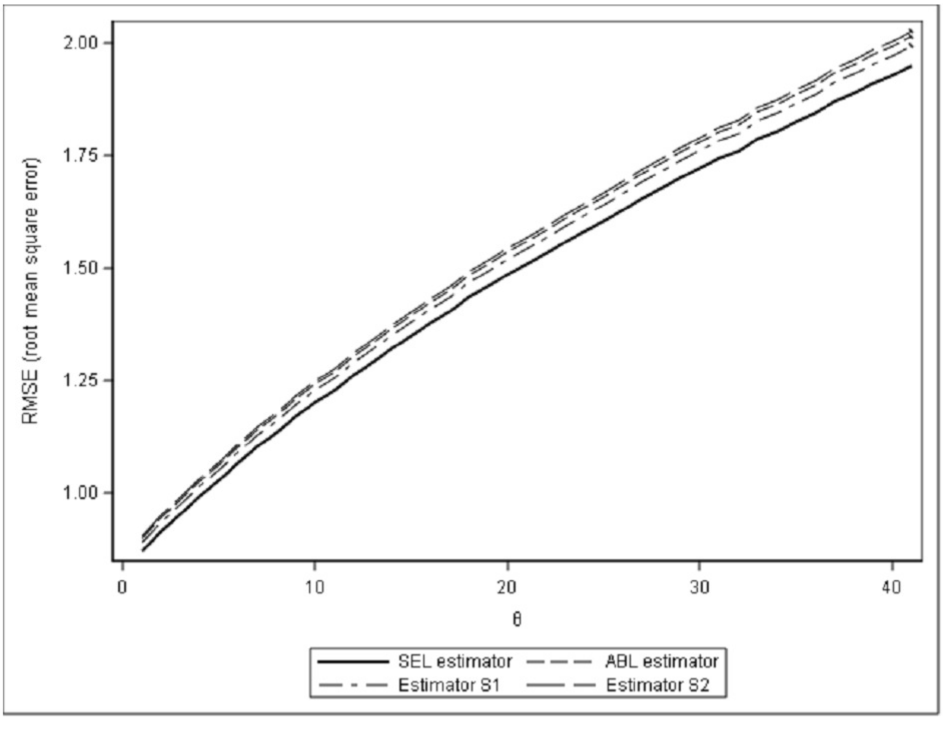

Figure 10: Root mean square error for $\hat{\theta}_{est}$ under SEL, ABL, S_1 and S_2 where $\theta_0 = MLE$, and varying θ ($m = 1$ (fixed), $c = 0.5$, $\lambda = 0.5$).

It is observed that the SEL estimator is preferable for the considered Rayleigh model against that of the ABL estimator, and both considered shrinkage estimators. The SEL estimator is also preferable to its corresponding shrinkage estimator, and the ABL estimator is also preferable to its corresponding shrinkage estimator. These are for the cases when the MLE and the Bayesian estimate carries equal weight in the shrinked estimator.

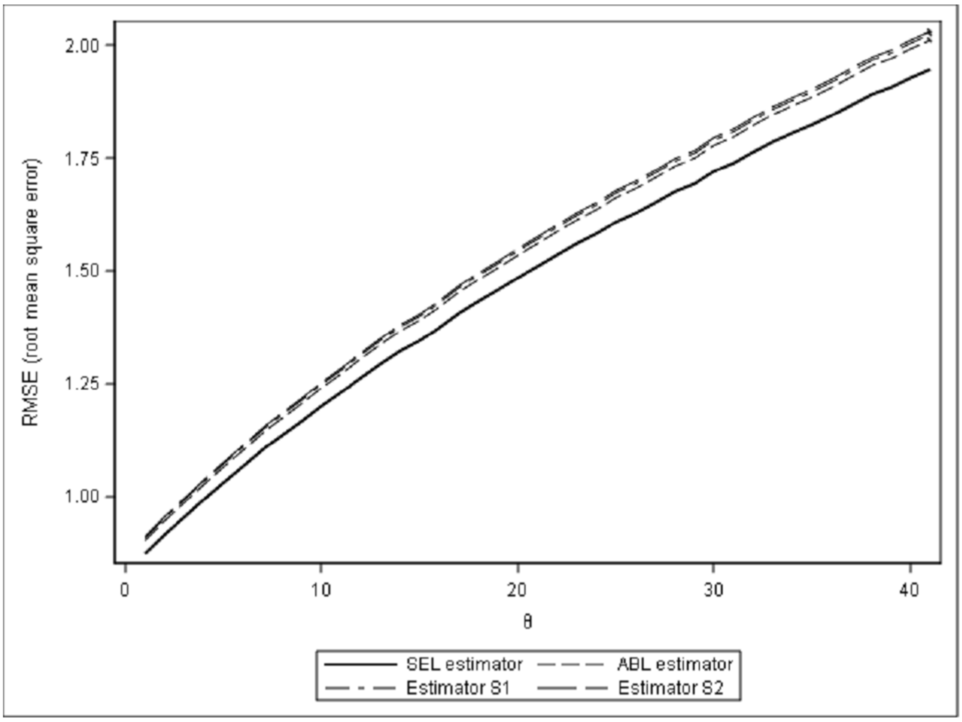

Figure 11: Root mean square error for $\hat{\theta}_{est}$ under SEL, ABL, S_1 and S_2 where $\theta_0 = MLE$, and varying θ ($m = 1$ (fixed), $c = 0.5$, $\lambda = 0.1$).

Figure 11 shows the case when the weight of the shrinkage estimators are skewed toward the MLE. Even in this case, both Bayesian estimates are preferred compared to their respective shrinkage estimators.

3.2. Practical application: gastrointestinal tumor group

The results are illustrated using gastrointestinal tumor study group data, obtained from Stablein, Carter, and Novak (1981) from a clinical trial in the treatment of locally advanced nonresectable gastric carcinoma. Mostert (1999) showed that the Rayleigh model is suitable for this data — it is also of censored nature which applies here. The sample size is $n = 45$, and the number of fully observed lifetimes is $d = 37$, where $T = 133.643$. The MLE of θ was used as the estimate θ_0. Table 1 below gives the parameter estimates under different loss function ((2.3) and (2.4)) for different parameter combinations.

Table 1: Parameter estimates under SEL and ABL for the real data set, for different values of m and c.

Value of m	Estimate value	$c = -1$	$c = -0.5$	$c = 0.5$	$c = 1$
$m = 1$	$\hat{\theta}_{SEL} = 0.27685$				
	$\hat{\theta}_{MLE} = 0.27685$				
	$\hat{\theta}_{ABL}$	0.26937	0.27311	0.28059	0.28433
$m = 3$	$\hat{\theta}_{SEL} = 0.26189$				
	$\hat{\theta}_{MLE} = 0.27685$				
	$\hat{\theta}_{ABL}$	0.25440	0.25815	0.26563	0.26937
$m = 10$	$\hat{\theta}_{SEL} = 0.20951$				
	$\hat{\theta}_{MLE} = 0.27685$				
	$\hat{\theta}_{ABL}$	0.20203	0.20577	0.21325	0.21699

This example aims to emphasize the effect of the shrinkage effect of the respective shrinkage estimators ((2.7) and (2.8)) and was achieved via a bootstrapping approach. By using the bootstrap method, a sampling distribution of the mentioned estimators can be constructed, and determined whether the estimator has a convergent nature — also, to have small standard error. The convergent nature of the bootstrap in parameter estimation is expected to illustrate the shrinkage effect to determine which estimator seems more appropriate for the given data set.

As mentioned, the performance of the estimator was studied via bootstrapping from the sample $k = 1000$ times. Thus, 1000 samples were drawn from the original sample with replacement, and for each of the drawn samples, the estimator under SEL was computed, and the risk value. The risk value was computed via

$$R^*\left(\hat{\theta}_{SEL}, \theta\right) = \frac{1}{k}\sum_{i=1}^{k}\left(\hat{\theta}_{S_1,i} - \theta\right)^2$$

where $\hat{\theta}_{S_1,i}$ is the shrinkage estimator (2.3) for the i^{th} bootstrapped sample, and θ the fixed sample parameter (determined via reparametrization of the mean of the distribution, equal to $\mu = \frac{1}{2}\sqrt{\frac{\pi}{\theta}}$, thus $\theta = \frac{\pi}{(2\mu)^2}$). This risk value was determined for increasing λ and graphed correspondingly, and is presented in Figure 12. It can be concluded that the estimator is indeed accurate and stable; in addition, from visual inspection it is observed that the estimator indeed has a small standard error. However, because of its near-convergent nature as $\lambda \to 1$, in this example, $\hat{\theta}_{SEL}$ is preferred to that of the MLE. This is in accordance with the $RMSE$

study in the preceeding section. This could be attributed to the shrinkage effect present in the shrinkage expression (2.7).

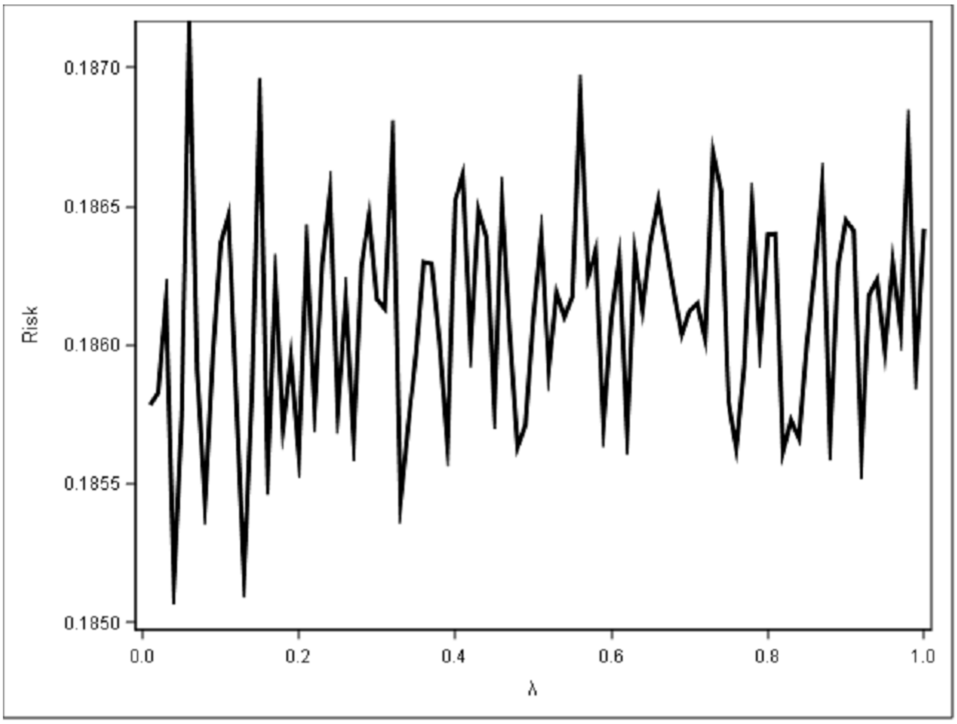

Figure 12: Bootstrap estimated values of $\hat{\theta}_{S_1,i}$, for $m = 1$ and increasing λ.

4. CONCLUSION

This paper explored the behaviour of the loss function proposed by Al-Bayyati (2002) by comparing it to the well-known squared error loss function. Bayes- and shrinkage estimators were derived. Their performance was studied under each of the mentioned loss functions in terms of their respective risk. It was observed that for positive values of c, the Al-Bayyati loss parameter, the risk of SEL was lower than that of ABL. Another focus of this paper was the effect of the prior family degree m. It was observed that the risk of both SEL and ABL became larger as m increased. In a risk efficiency perspective, it was seen that negative values of c results in the ABL estimator being more efficient under SEL since the risk is then smaller. The reciprocal result holds when the underlying loss function is the ABL. When the underlying loss is ABL, then for positive values of c the SEL estimator performs better in terms of risk.

After proposing shrinkage estimators (where the derived Bayesian estimators are combined in linear fashion with some pre-specified point estimate of the parameter) their risk and risk efficiency was also studied. It was observed that for underestimation of the parameter, the shrinkage estimator yielded lower risk than that of only the Bayesian estimator itself. For overestimation, only the Bayesian estimator performed better than the shrinkage estimator. In the risk efficiency setting it was observed that there does exist some values of λ which results in the shrinkage estimator under ABL performing better than the SEL estimator when the underlying loss function is SEL.

As a simulation study the RMSE was determined for each of the proposed estimators and subsequently compared. It was seen that the estimator under SEL remains preferable when considering the RMSE criterion. A numerical example also followed showing the applicational use of the estimators to a real data set.

ACKNOWLEDGMENTS

The authors wish to acknowledge the Office of the Dean of the Faculty of Natural and Agricultural Sciences, University of Pretoria, for their financial assistance toward this study. In addition, the support from STATOMET, Department of Statistics, Faculty of Natural and Agricultural Sciences, University of Pretoria is also humbly acknowledged. Finally the anonymous reviewer is thanked for his/her constructive comments and suggestions for greatly improving the quality of this paper.

REFERENCES

[1] Al-Bayyati, H.N. (2002). Comparing methods of estimating Weibull failure models using simulation, *Unpublished PhD thesis*, College of Administration and Economics, Baghdad University, Iraq.

[2] Bhattacharya, S.K. and Tyagi, R.K. (1990). Bayesian survival analysis based on the Raleigh model, *Trabajos de Estadistica*, **5**(1), 81–92.

[3] Delaportas, P. and Wright, D.E. (1991). Numerical prediction for the two-parameter Weibull distribution, *The Statistician*, **40**, 365–372.

[4] Dey, S. (2011). Comparison of relative risk functions of the Rayleigh distribution under Type II censored samples: Bayesian approach, *Jordan Journal of Mathematics and Statistics*, **4**(1), 61–68.

[5] Dey, S. and Dey, T. (2011). Rayleigh distribution revisited via an extension of Jeffreys prior information and a new loss function, *Revstat*, **9**(3), 213–226.

[6] Gruber, M.H.J. (2004). The efficiency of shrinkage estimators with respect to Zellner's balanced loss function, *Communications in Statistics — Theory and Methods*, **33**(2), 235–249.

[7] Liang, F.; Paulo, R.; German, G.; Clyde, M.A. and Berger, J.O. (2008). Mixtures of g priors for Bayesian variable selection, *Journal of the American Statistical Association*, **103**, 401–414.

[8] Mostert, P.J. (1999). A Bayesian method to analyse cancer lifetimes using Rayleigh models, *Unpublished PhD thesis*, University of South Africa.

[9] Soliman, A.A. (2000). Comparison of LINEX and quadratic Bayes estimators for the Rayleigh distribution, *Communications in Statistics — Theory and Methods*, **29**(1), 95–107.

[10] Stablein, D.M.; Carter, W.H. and Novak, J.W. (1981). Analysis of survival data with nonproportional hazard functions, *Controlled Clinical Trials*, **2**, 149–159.