# A BAYESIAN ANALYSIS FOR CENSORED RAYLEIGH MODEL USING A GENERALISED HYPERGEOMETRIC PRIOR

***P. J. Mostert*** [1]
Department of Statistics and Actuarial Science, Stellenbosch University
e-mail: *pjmos@sun.ac.za*

***A. Bekker***
Department of Statistics, University of Pretoria

***J. J. J. Roux***
Department of Statistics, University of South Africa

***Key words:*** Bayesian estimation, Bayesian prediction bounds, Frequentist coverage, Generalised hypergeometric distribution, H-function, Metropolis-Hastings algorithm, Rayleigh model, Squared error loss function, Type II censored data.

***Abstract:*** Based on a type II censored sample, Bayesian estimation for the scale parameter of the Rayleigh model is carried out under the assumption of the squared error loss function. A generalised hypergeometric distribution with its versatile shape of tails is introduced as a prior, and beta special cases are examined. A simulation study is carried out to investigate the sensitivity of four special cases of this beta prior family in terms of bias, frequentist coverage and mean square error and to determine their effect on robustness. Prediction bounds are derived for the lifetime of unused components using this beta prior family. A data set is used to illustrate and support some of the findings.

## 1. Introduction

The Rayleigh distribution has been shown to be the survival time distribution for most cancer patients in some clinical studies as well as for a variety of lifetime models, for which the survival time $T$ is specified by the probability density function (pdf)

$$f(t;\theta) = 2\theta t e^{-\theta t^2}, \tag{1}$$

where $t > 0$ and $\theta > 0$ (see Bhattacharya and Tyagi, 1990). Assume that $n_1$ items from a population specified by (1) are placed on a life test and the experiment is continued until $d$ deaths occur, for

[1] Corresponding author.

*AMS: 62C10, 62N01, 62F15*

some preassigned $d$ (type II censoring). Let $t_1,...,t_d$ be the first $d$ ordered observations in the random sample $(n_1-d)$, being the number of survivors or patients who were still alive at the end of the study. The likelihood function conditional on $\theta$ is given by

$$\ell(t_1,...,t_d|\theta)=\frac{n_1!}{(n_1-d)!}(\Pi_{i=1}^{d}t_i)2^d\theta^d e^{-\theta T}, \tag{2}$$

where $T=\Sigma_{i=1}^{d}t_i^2+(n_1-d)t_d^2$ is a complete and sufficient statistic for estimating $\theta$ and has a gamma distribution with parameters $d$ and $\theta$.

Mostert (1999), Soliman (2000), and Dey and Dey (2011) compared Bayesian estimators under the asymmetric linear exponential loss (LINEX) function and squared error loss (SEL) function for the Rayleigh model. Dey and Dey (2011) focused on an extension of Jeffreys' prior using the loss, introduced by (Al-Bayyati, 2002). Mostert, Bekker and Roux (1998) did a comparative study of the lifetime parameters for the Rayleigh distribution as the underlying model using both the conjugate and Jeffreys' priors under LINEX and squared error loss functions. Dey and Dey (2012) and Dey and Das (2007) studied the conjugate prior for a completely observed sample from the Rayleigh model under different loss functions. Ferreira, Bekker and Arashi (2016) studied objective priors for the Al-Bayyati loss function for this censored model.

As mentioned in Berger (1980), conjugate priors have tails of the same type as the tails of the likelihood function and may cause certain robustness problems. Priors that have flatter tails than that of the likelihood function are generally regarded as superior, at least for inference problems. Conjugate priors as well as classes of non-informative priors have been studied extensively in the literature. Different loss functions point towards different optimal procedures. A fresh approach to the above studies for the Rayleigh model is to assume a generalised hypergeometric distribution (Mathai and Saxena, 1966) as prior for the estimation problem at hand, based on a type II censored sample. As with this prior it is desirable, as mention by Sebastian (2011), to consider priors that exhibit flexibility regarding the shape of their tails.

As Gelman (2006) and Gelman (2009) state, non-informative priors can have strong and undesirable implications for inference. Ignoring important information by assuming priors that cannot elicit that information will lead to inaccurate analysis and inference, hence it is important to assume a parameter-rich model that will enable one to model appropriate prior beliefs about the parameter of interest. Computational limitations within the Bayesian framework also no longer present a problem to analysing any data set with complex prior assumptions; a generalised prior can therefore be handled with ease in Bayesian inference.

In Section 2, some special cases related to the beta family of this subjective generalised prior will be reviewed. Sections 3 and 4 present the framework for inference using this proposed prior. The Bayesian estimator of the parameter of interest is obtained with respect to the squared error loss function, and prediction bounds are derived for the failure times of a future sample of unused components. In Section 5, the sensitivity of the estimators is examined in a simulation study and discussed with respect to the different special cases of this generalised prior, as well as obtaining the prediction bounds in all the cases. A frequentist approach is used to evaluate the performance of the various estimators. A real data set illustrates the value added by using the special cases of the generalised model as prior.

## 2. Generalised hypergeometric prior

The generalised hypergeometric pdf (Mathai and Saxena, 1966) is defined as:

$$g(\theta)=\frac{mb^{-p/m}\Gamma(\alpha)\Gamma(\gamma)\Gamma(\delta-p/m)}{\Gamma(p/m)\Gamma(\delta)\Gamma(\alpha-p/m)\Gamma(\gamma-p/m)}\theta^{p-1}{}_2F_1(\alpha,\gamma;\delta;-\frac{1}{b}\theta^m), \tag{3}$$

where $\theta>0, p>0, \alpha-p/m>0, \gamma-p/m>0, \delta-p/m>0$ and ${}_2F_1(\alpha,\gamma;\delta;-\frac{1}{b}\theta^m)$ is the Gauss hypergeometric function (Erdélyi, 1953, p. 182). The parameters $\alpha,\gamma,\delta,p,m$ and $b$ are restricted to take those values for which $\int_0^\infty g(\theta)d\theta=1$ with $g(\theta)\geq 0$ for $\theta$ positive. It can be observed that if $\theta$ has pdf (3), then

$$E[\theta^h]=\frac{b^{h/m}\Gamma(\delta-p/m)\Gamma((h+p)/m)\Gamma(\alpha-(h+p)/m)\Gamma(\gamma-(h+p)/m)}{\Gamma(p/m)\Gamma(\delta-(h+p)/m)\Gamma(\alpha-p/m)\Gamma(\gamma-p/m)},$$

if $\left|\arg b^{-1}\right|<\pi$, $0<(h+p)/m<1-\max[Re(1-\alpha),Re(1-\gamma)]$.

Four special cases of this generalised pdf (3) are given below and will be discussed in the subsequent sections. These special cases form the basis of the simulation study that is the focus of Section 5.

(i) *Generalised F-distribution*: Let $\gamma=\delta$, and considering reparameterisation $\alpha=p+q$ in (3), then

$$g_1(\theta)=\frac{mb^{-p/m}\Gamma(p+q)}{\Gamma(p/m)\Gamma(p+q-p/m)}\theta^{p-1}(1+\frac{1}{b}\theta^m)^{-(p+q)}, \tag{4}$$

where $\theta>0, p>0, q>0, p+q>p/m, b>0, m>0$. Now

$$E_1[\theta]=\frac{b^{1/m}\Gamma((1+p)/m)\Gamma(p+q-(1+p)/m)}{\Gamma(p/m)\Gamma(p+q-p/m)}$$

and

$$var_1(\theta)=\frac{b^{2/m}\Gamma((2+p)/m)\Gamma(p+q-(2+p)/m)}{\Gamma(p/m)\Gamma(p+q-p/m)}-[E_1[\theta]]^2.$$

(ii) *Type I compound gamma distribution* (Dubey, 1970, according to Bhattacharya and Tyagi, 1990, it is a beta distribution of the second kind): With $\gamma=\delta$, $m=1$, and considering the reparameterisation $\alpha=p+q$ in (3), $\theta$ has pdf

$$g_2(\theta)=\frac{b^{-p}\Gamma(p+q)}{\Gamma(p)\Gamma(q)}\theta^{p-1}(1+\frac{1}{b}\theta)^{-(p+q)}, \tag{5}$$

where $\theta>0, p>0, q>0, b>0$. Thus

$$E_2[\theta]=\frac{pb}{q-1}, \qquad (q>1) \tag{6}$$

and

$$var_2(\theta) = \frac{b^2(p^2+p(q-1))}{(q-1)^2(q-2)}, \qquad (q>1);(q>2),$$

also, a special case of (4) with $m = 1$.

Bhattacharya and Tyagi (1990) obtained the Bayesian estimators of the mean survival time, the hazard function, and the survival function with (5) as prior for $\theta$. In their numerical illustration, the requirement $q > 2$ has not been met.

(iii) *F-distribution*: Let $\gamma = \delta, m = 1, \alpha = p+q$ and $b = \frac{q}{p}$. Then it follows from (3) for $\theta > 0$, that

$$g_3(\theta) = \frac{p^p\Gamma(p+q)}{q^p\Gamma(p)\Gamma(q)}\theta^{p-1}(1+\frac{p}{q}\theta)^{-(p+q)}, \tag{7}$$

where $\theta > 0, p > 2, q > 0$. Hence

$$E_3[\theta] = \frac{q}{q-1}, \qquad (q>1)$$

and

$$var_3(\theta) = \frac{q^2(p^2+p(q-1))}{p^2(q-1)^2(q-2)}, \qquad (q>1);(q>2).$$

(iv) *Beta type II distribution*: With $\gamma = \delta, m = 1, \alpha = p+q$ and $b = 1$ in (3), $\theta$ has pdf

$$g_4(\theta) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}\theta^{p-1}(1+\theta)^{-(p+q)}, \tag{8}$$

where $\theta > 0, p > 0, q > 0$. Hence

$$E_4[\theta] = \frac{p}{q-1}, \qquad (q>1) \tag{9}$$

and

$$var_4(\theta) = \frac{p^2+p(q-1))}{(q-1)^2(q-2)}, \qquad (q>1);(q>2),$$

also here, a special case of (5) with $b = 1$.

From (6) and (9), it follows that $E_2[\theta] = bE_4[\theta]$. Thus it can be said that there are $b$ more (or fewer) units of input information in the prior density (5) (Pham-Gia, 1994, p. 2179).

# 3. Bayesian inference

The likelihood function (2) is combined with the prior density (3) to obtain the posterior density function

$$h(\theta|T) = C^{-1}\theta^{p+d-1}e^{-\theta T}{}_2F_1(\alpha,\gamma;\delta;-\frac{1}{b}\theta^m) \tag{10}$$

and the values of the parameters such that $\int_0^\infty h(\theta|T)d\theta = 1$ with $h(\theta|T) \geq 0$ for $\theta$ positive.

The following integral representation given by Gupta (1965, p. 100) will be used to find the normalisation constant $C$:

$$\begin{aligned}&\int_0^\infty e^{-\lambda x}x^\nu H_{r,s}^{k,l}\left[zx^\sigma \left| \begin{matrix}(a_1,e_1),...,(a_r,e_r)\\(c_1,f_1),...,(c_s,f_s)\end{matrix}\right.\right]dx \\ = \;& \lambda^{-\nu-1}H_{r+1,s}^{k,l+1}\left[z\lambda^{-\sigma}\left| \begin{matrix}(-\nu,\sigma),(a_1,e_1),...,(a_r,e_r)\\(c_1,f_1),...,(c_s,f_s)\end{matrix}\right.\right]\end{aligned} \tag{11}$$

where $\sigma > 0, Re(\lambda) > 0, Re(\nu + 1 + \sigma \min\frac{c_i}{f_i}) > 0$ $(i = 1,...,k)$, the other conditions of validity are given in Gupta (1965), and $H(\cdot)$ is the $H$-function introduced by Fox (1961, p. 408). Therefore

$$\begin{aligned}C &= \int_0^\infty \theta^{p+d-1}e^{-\theta T}{}_2F_1(\alpha,\gamma;\delta;-\frac{1}{b}\theta^m)d\theta \\ &= \frac{\Gamma(\delta)}{\Gamma(\alpha)\Gamma(\gamma)T^{(p+d)}}H_{3,2}^{1,3}\left[\frac{1}{bT^m}\left| \begin{matrix}(1-p-d,m),(1-\alpha,1),(1-\gamma,1)\\(0,1),(1-\delta,1)\end{matrix}\right.\right]\end{aligned} \tag{12}$$

by using (11) and the relation

$${}_2F_1(\alpha,\gamma;\delta;-\frac{1}{b}\theta^m) = \frac{\Gamma(\delta)}{\Gamma(\alpha)\Gamma(\gamma)}T^{-(p+d)}H_{2,2}^{1,2}\left[\frac{1}{b}\theta^m\left| \begin{matrix}(1-\alpha,1),(1-\gamma,1)\\(0,1),(1-\delta,1)\end{matrix}\right.\right]$$

(Mathai and Saxena, 1978, p. 11) where $m > 0$, $Re(T) > 0$, $p+d > 0$.

Hence, under the assumption of the squared error loss function, the Bayesian estimator of the scale parameter $\theta$ is simply the posterior mean

$$\hat{\theta} = T^{-1}\frac{H_{3,2}^{1,3}\left[\frac{1}{bT^m}\left| \begin{matrix}(-p-d,m),(1-\alpha,1),(1-\gamma,1)\\(0,1),(1-\delta,1)\end{matrix}\right.\right]}{H_{3,2}^{1,3}\left[\frac{1}{bT^m}\left| \begin{matrix}(1-p-d,m),(1-\alpha,1),(1-\gamma,1)\\(0,1),(1-\delta,1)\end{matrix}\right.\right]} \tag{13}$$

from (10) and (12).

In a similar manner, the variance of the posterior distribution of $\theta$ is computed as

$$var(\hat{\theta}) = T^{-2}\frac{H_{3,2}^{1,3}\left[\frac{1}{bT^m}\left| \begin{matrix}(-p-d-1,m),(1-\alpha,1),(1-\gamma,1)\\(0,1),(1-\delta,1)\end{matrix}\right.\right]}{H_{3,2}^{1,3}\left[\frac{1}{bT^m}\left| \begin{matrix}(1-p-d,m),(1-\alpha,1),(1-\gamma,1)\\(0,1),(1-\delta,1)\end{matrix}\right.\right]} - \hat{\theta}^2.$$

The posterior density function, the Bayesian estimator and the variance of the posterior distribution of $\theta$ for the special cases will now be discussed.

(i) *Generalised F-distribution* (4) as prior:

The posterior density function is

$$h_1(\theta|T) = C_1^{-1}\theta^{p+d-1}e^{-\theta T}(1+\frac{1}{b}\theta^m)^{-(p+q)} \tag{14}$$

where the normalisation constant is

$$\begin{aligned} C_1 &= \int_0^\infty \theta^{p+d-1}e^{-\theta T}(1+\frac{1}{b}\theta^m)^{-(p+q)}d\theta \qquad (15)\\ &= \frac{T^{-(p+d)}}{\Gamma(p+q)}H_{2,1}^{1,2}\left[\frac{1}{b}T^{-m}\middle| \begin{array}{c}(1-p-d,m),(1-p-q,1)\\(0,1)\end{array}\right] \end{aligned}$$

from (12) with $\gamma=\delta$ and $\alpha=p+q$.

The constant $C_1$ can also be derived from (11) using the relations (see Mathai and Saxena, 1978, p. 10)

$$\begin{aligned} (1+\frac{1}{b}\theta^m)^{-(p+q)} &= {}_1F_0(p+q;-\frac{1}{b}\theta^m)\\ &= \frac{1}{\Gamma(p+q)}H_{1,1}^{1,1}\left[\frac{1}{b}\theta^m\middle| \begin{array}{c}(1-p-q,1)\\(0,1)\end{array}\right]. \end{aligned}$$

The posterior density function (14) can also be rewritten as a mixture of gamma densities by using the series expansion of ${}_1F_0(\cdot)$.

The Bayesian estimator of $\theta$ is obtained as

$$\begin{aligned} \hat{\theta}_1 &= C_1^{-1}\int_0^\infty \theta^{p+d}e^{-\theta T}(1+\frac{1}{b}\theta^m)^{-(p+q)}d\theta\\ &= T^{-1}\frac{H_{2,1}^{1,2}\left[\frac{1}{b}T^{-m}\middle| \begin{array}{c}(-p-d,m),(1-p-q,1))\\(0,1)\end{array}\right]}{H_{2,1}^{1,2}\left[\frac{1}{b}T^{-m}\middle| \begin{array}{c}(1-p-d,m),(1-p-q,1)\\(0,1)\end{array}\right]} \end{aligned}$$

from (13) with $\alpha=p+q$ and $\gamma=\delta$, and $var(\hat{\theta}_1)$ can be derived in a similar way.

(ii) *Type I compound gamma distribution* (5) as prior:

The posterior density function is

$$h_2(\theta|T) = C_2^{-1}\theta^{p+d-1}e^{-\theta T}(1+\frac{1}{b}\theta)^{-(p+q)} \tag{16}$$

with the normalisation constant $C_2 = b^{p+d}\Gamma(p+d)\psi(p+d,d-q+1,bT)$, where $\psi(\cdot)$ is the Type II confluent hypergeometric function introduced by Tricomi (Erdélyi, 1953).

$C_2$ can also easily be derived from (15) with $m = 1$ and known relations (Erdélyi, 1954, p. 375) and (Mathai, 1993, pp. 72, 130, 142). The following are also obtained:

$$\hat{\theta}_2 = b(p+d)\frac{\psi(p+d+1,d-q+2,bT)}{\psi(p+d,d-q+1,bT)} \tag{17}$$

and

$$var(\hat{\theta}_2) = b^2(p+d)(p+d+1)\frac{\psi(p+d+2,d-q+3,bT)}{\psi(p+d,d-q+1,bT)} - \hat{\theta}_2^2.$$

(iii) *F-distribution* (7) as prior:

As in (ii), it can be shown that the posterior density function and the Bayesian estimator of $\theta$ are as follows:

$$h_3(\theta|T) = C_3^{-1}\theta^{p+d-1}e^{-\theta T}(1+\frac{p}{q}\theta)^{-(p+q)}$$

where

$$C_3 = \left(\frac{q}{p}\right)^{p+d}\Gamma(p+d)\psi(p+d,d-q+1,\frac{q}{p}T)$$

and

$$\hat{\theta}_3 = (\frac{q(p+d)}{p})\frac{\psi(p+d+1,d-q+2,\frac{q}{p}T)}{\psi(p+d,d-q+1,\frac{q}{p}T)}.$$

(iv) *Beta type II distribution* (8) as prior:

From (16) and (17), it is clear that the posterior density function of $\theta$ is

$$h_4(\theta|T) = C_4^{-1}\theta^{p+d-1}e^{-\theta T}(1+\theta)^{-(p+q)}$$

where

$$C_4 = \Gamma(p+d)\psi(p+d,d-q+1,T)$$

and

$$\hat{\theta}_4 = (p+d)\frac{\psi(p+d+1,d-q+2,T)}{\psi(p+d,d-q+1,T)}.$$

$Var(\hat{\theta}_3)$ and $Var(\hat{\theta}_4)$ can be derived with relative ease.

**Remark 1** The normalisation constants in all the above expressions of the posterior distributions can be obtained in a number of mathematical software programs using, for example, the built-in functions for the type II confluent hypergeometric function $\psi(\cdot)$ and the Gauss hypergeometric function ${}_2F_1(\cdot)$. An alternative method that is preferred in this paper is to use MCMC methods and the Metropolis-Hastings algorithm to obtain variates of the posterior distributions, and hence Bayesian estimators of the parameters.

# 4. Bayesian prediction bounds

Consider predicting $Y_1$, the time to first failure, in a future sample of size $n_2$. For given $\theta$, the density of $Y_1$ is

$$\begin{aligned} f(y_1|\theta) &= \binom{n_2}{1} f(y_1)[1-F(y_1)]^{n_2-1} \\ &= 2n_2\theta y_1 e^{-\theta n_2 y_1^2}, \end{aligned}$$

which is again recognised as a Rayleigh distribution with parameter $n_2\theta$.

Forming the product of $f(y_1|\theta)$ with the posterior density function of $\theta$ in (10), the predictive density function of $Y_1$ is

$$f(y_1|t_1,...,t_d) = C^{-1}\int_0^\infty 2n_2 y_1 \theta^{p+d} e^{-\theta(T+n_2 y_1^2)} {}_2F_1(\alpha,\gamma;\delta;-\frac{1}{b}\theta^m)d\theta.$$

The predictive survival function of $Y_1$ is

$$\begin{aligned} P(Y_1 &> y) = 1 - P(Y_1 \leq y) \\ &= 1 - \int_0^y C^{-1}\int_0^\infty 2n_2 y_1 \theta^{p+d} e^{-\theta(T+n_2 y_1^2)} {}_2F_1(\alpha,\gamma;\delta;-\frac{1}{b}\theta^m)d\theta dy_1 \\ &= 1 - C^{-1}\int_0^\infty \theta^{p+d-1} e^{-\theta T} {}_2F_1(\alpha,\gamma;\delta;-\frac{1}{b}\theta^m)\int_0^y 2n_2 y_1 \theta e^{-\theta n_2 y_1^2)} dy_1 d\theta \\ &= 1 - C^{-1}\int_0^\infty \theta^{p+d-1} e^{-\theta T} {}_2F_1(\alpha,\gamma;\delta;-\frac{1}{b}\theta^m)(1-e^{-\theta n_2 y^2)})d\theta \\ &= C^{-1}\int_0^\infty \theta^{p+d-1} e^{-\theta(T+n_2 y^2)} {}_2F_1(\alpha,\gamma;\delta;-\frac{1}{b}\theta^m)d\theta \end{aligned}$$

and then the lower $100\Lambda\%$ prediction bound for $Y_1$ is the value $y$ satisfying the equation

$$C^{-1}\int_0^\infty \theta^{p+d-1} e^{-\theta(T+n_2 y^2)} {}_2F_1(\alpha,\gamma;\delta;-\frac{1}{b}\theta^m)d\theta = \Lambda,$$

which can be solved for $y$ using a numerical integration technique or MCMC methods.

**Remark 2** Where MCMC methods are used to obtain the posterior distribution of $\theta$ (i.e. where 10 000 variates of the posterior have been simulated using the Metropolis-Hastings algorithm), the prediction density can conveniently be obtained using the fact that $Y_1$ is Rayleigh distributed with parameter $n_2\theta$. The prediction of a future lifetime is obtained by letting $n_2 = 1$ in this procedure.

# 5. Illustrations

## 5.1. Simulation study

A Monte Carlo simulation is carried out in this section to study the performance of the proposed Bayesian estimators and the characteristics of the posterior distributions using the priors discussed in Section 2 for different sample sizes ($n_1 = n = 18$ and $n_1 = n = 35$), as well as different values for a number of actual events $d$ in a sample. Firstly, the performance is evaluated based on the bias and the mean squared error (MSE) criteria, where the MSE of $\hat{\theta}$ (the Bayesian estimator of $\theta$) is defined by

$$MSE(\hat{\theta}) = E[\hat{\theta} - \theta]^2 = var(\hat{\theta}) + [bias(\hat{\theta})]^2. \tag{18}$$

Secondly, coverage probabilities are determined for different values of $\theta$ under the hypothesis

$$H_0 : \theta = \theta_0.$$

Random samples from a Rayleigh distribution are generated for different numbers of events, $d$, to study the performance of the estimators using bias, MSE and coverage. The hyperparameters of the different prior distributions are chosen in such a way that the expected value of the prior (see (i) – (iv)) equates to the true value of the parameter under consideration, with less variation around the expected value (see, amongst others, Duran and Booker, 1988). This paper studies the sensitivity of the values of the parameters to errors in specification.

In order to evaluate the performance, 1 000 random samples are generated, each of sizes $n_1 = n = 18$, with $d = 8, 10, 12, 14, 16, 18$; and $n_1 = n = 35$, with $d = 10, 15, 20, 25, 30, 35$, from a Rayleigh distribution with parameter $\theta = 2.5$. For each generated sample of size $n$, the sufficient statistic $T$ is determined for given $d$ values. The 1 000 posterior distributions for $\theta$ are simulated using the Metropolis-Hastings algorithm by first simulating 1 000 burn-in variates in order to obtain the 10 000 variates of the relevant posterior distribution. Table 1 summarises the assigned values of the hyperparameters in the prior with an expected value around 2.5 to be used in the simulation studies.

**Table 1:** Hyperparameter values of the prior.

| | $p$ | $q$ | $b$ | $m$ |
|---|---|---|---|---|
| $g_1(\theta)$ : Generalised F-distribution, see (4) | 5 | 4.0 | 0.526 | 0.8 |
| $g_2(\theta)$ : Type I compound gamma distribution, see (5) | 5 | 4.0 | 1.5 | 1.0 |
| $g_3(\theta)$ : F-distribution, see (7) | 1 | 1.667 | 1.667 | 1.0 |
| $g_4(\theta)$ : Beta type II distribution, see (8) | 8 | 4.2 | 1.0 | 1.0 |

The posterior means were determined, and MSE and bias were calculated using (18) for the $\theta_0$ values, chosen as increments between 1.7 and 3.0, which are indicated above the markers on all the graphs reflecting MSE and bias. Figure 1 summarises the results for prior (4) when using the posterior mean as Bayesian estimator with sample sizes of 18, while Figure 2 summarises the results for sample sizes of 35.

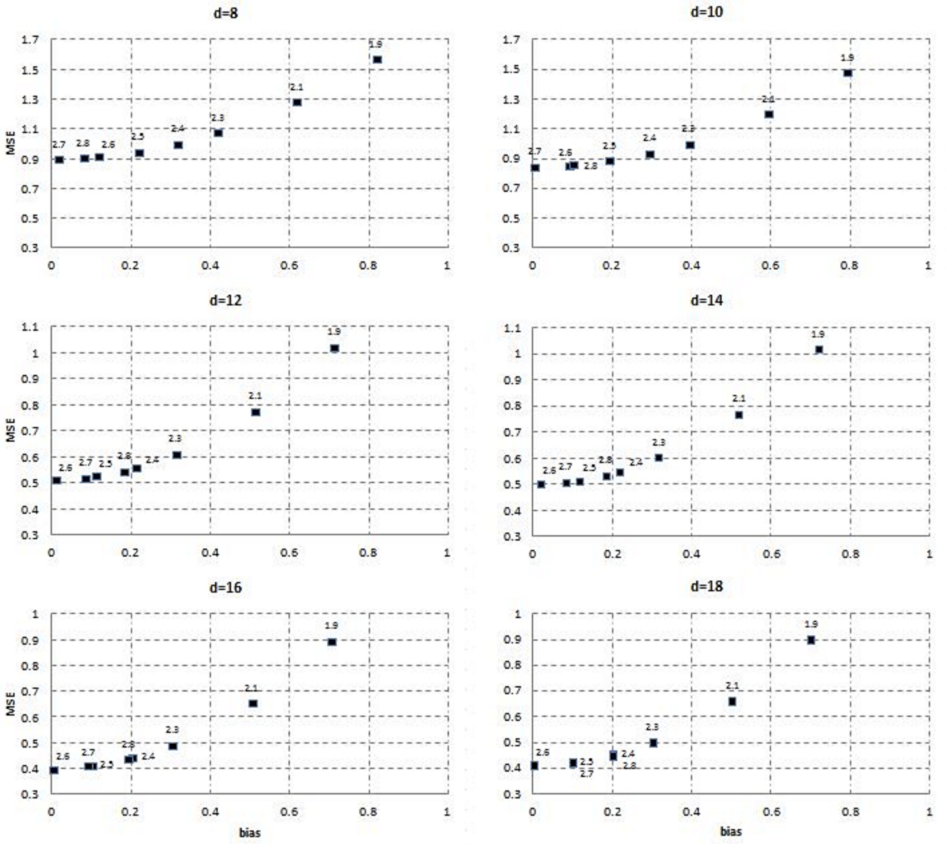


**Figure 1**: Bias versus MSE for the posterior mean using prior (4), adjusted for censoring ($n = 18$).

It is evident that bias and MSE vary less and are closer to zero for (a) larger sample sizes and (b) if $\theta_0$ under $H_0$ is close to the true parameter (2.5) in (18), when considering different censoring scenarios (numbers of events $d$). It was also evident in the simulation results that for larger number of events $d$, the MSE remained fairly constant for a particular $d$ value, but bias decreased overall $\theta_0$ in (18). The larger $d$ values in these graphs are the ones indicated with the smaller MSEs.

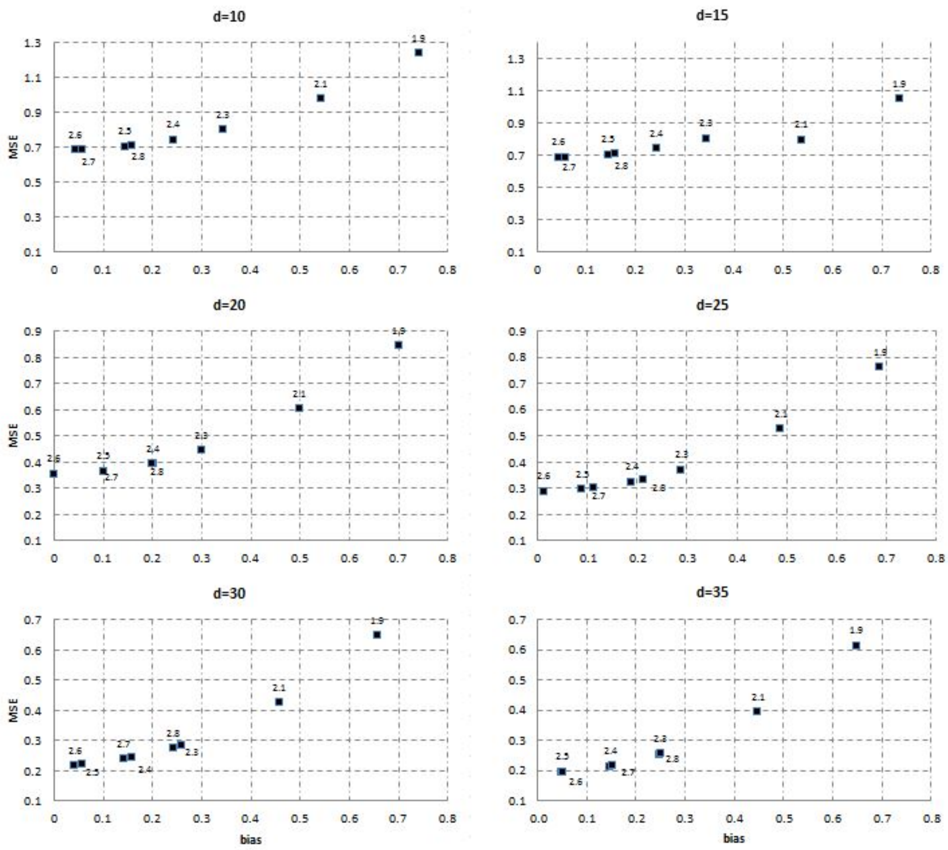


**Figure 2**: Bias versus MSE for the posterior mean using prior (4), adjusted for censoring ($n = 35$).

For each of the initial generated 1 000 random samples of size $n$, a $(1-\alpha)100\%$ credible interval from the posterior was obtained to determine whether a particular $\theta_0$ under $H_0$ is contained in that interval.

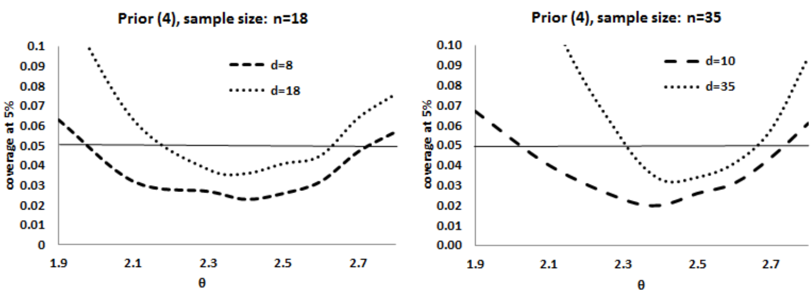


**Figure 3**: Coverage at $\alpha = 5\%$ using prior (4).

The relative frequency was determined at $\alpha = 0.05$, and Figure 3 shows the corresponding coverage at 5% for different sample size scenarios and only the extreme censoring schemes in this study, meaning that for other values of $d$, the coverage was observed between these extremes.

Figure 3 shows that for smaller numbers of events ($d$), the posterior seems less sensitive towards the true parameter, $\theta = 2.5$.

Comparisons are also done for the other priors, and Figures 4 and 5 summarise the results for prior (5) when using the posterior mean as Bayesian estimator for sample sizes of 18 and 35 respectively.

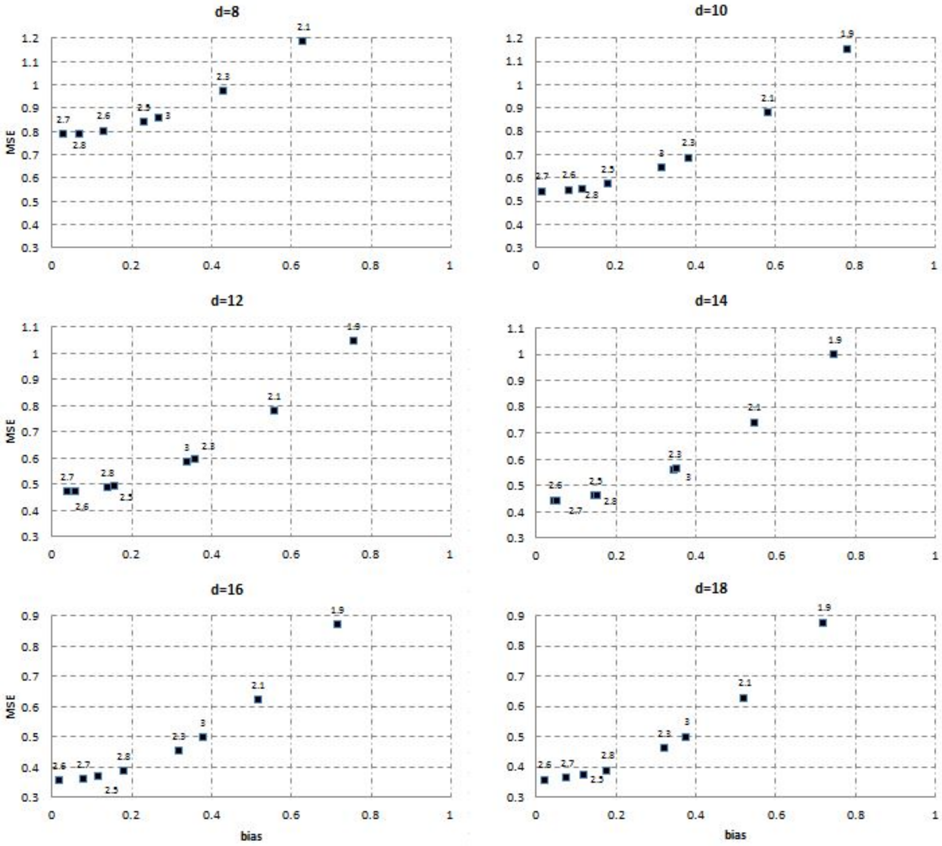


**Figure 4**: Bias versus MSE for the posterior mean using prior (5), adjusting for censoring ($n = 18$).

It is evident, as was the case with prior (4), that bias and MSE vary less and are closer to zero for (a) larger sample sizes and (b) if $\theta_0$ is close to the true parameter in (18) when considering the different censoring scenarios (number of events $d$). It was also evident in the simulation results that for larger numbers of events $d$, the MSE and bias decreased overall $\theta_0$ in (18). It is also evident that bias and MSE yield lower values for than for prior (5) than for prior (4). Figure 6 shows the corresponding coverage for different sample size scenarios and the extreme censoring schemes in this study.

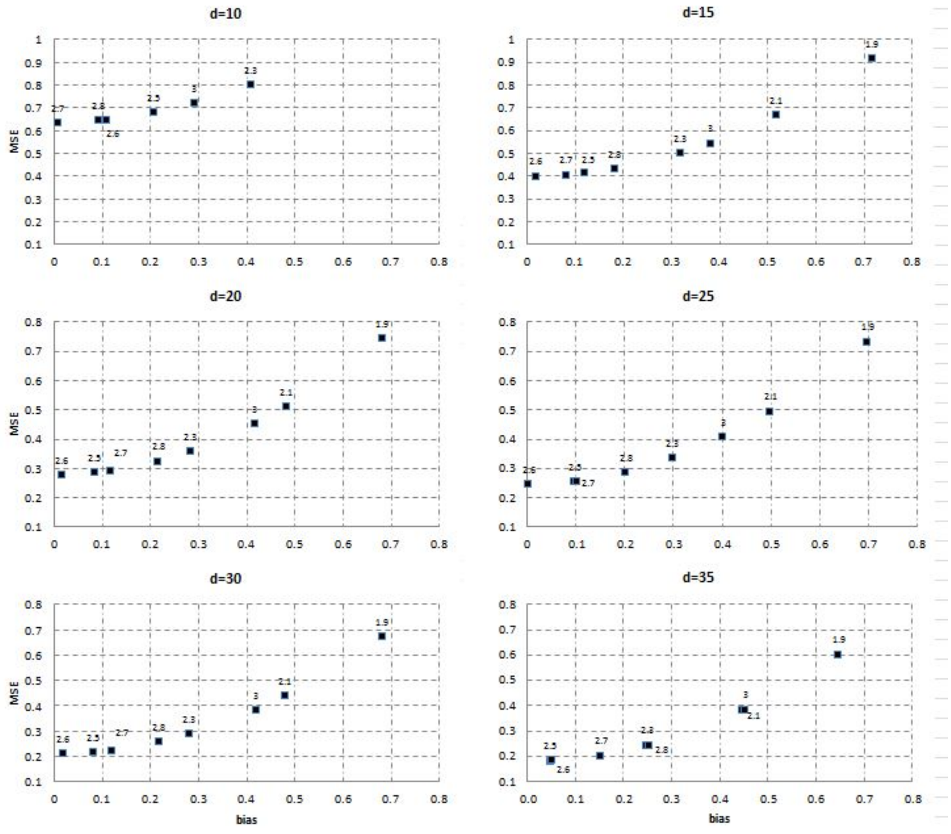

**Figure 5**: Bias versus MSE for the posterior mean using prior (5), adjusting for censoring ($n = 35$).

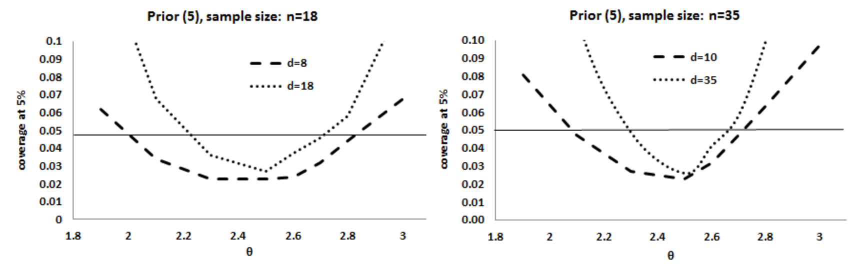

**Figure 6**: Coverage at $\alpha = 5\%$ using prior (5).

Figure 6 illustrates the sensitivity of Bayesian inference for $\theta$ outside the interval $[2.25; 2.70]$ with the effect of a larger sample size.

The posterior mean was then determined using prior (7), and MSE and bias were calculated using (18) for different and predetermined $\theta_0$ values in $H_0$. Figures 7 and 8 summarise the results for prior (7) when using the posterior mean as Bayesian estimator for sample sizes 18 and 35 respectively, while Figure 9 shows the corresponding coverage for different sample size scenarios and the extreme censoring schemes in this study.

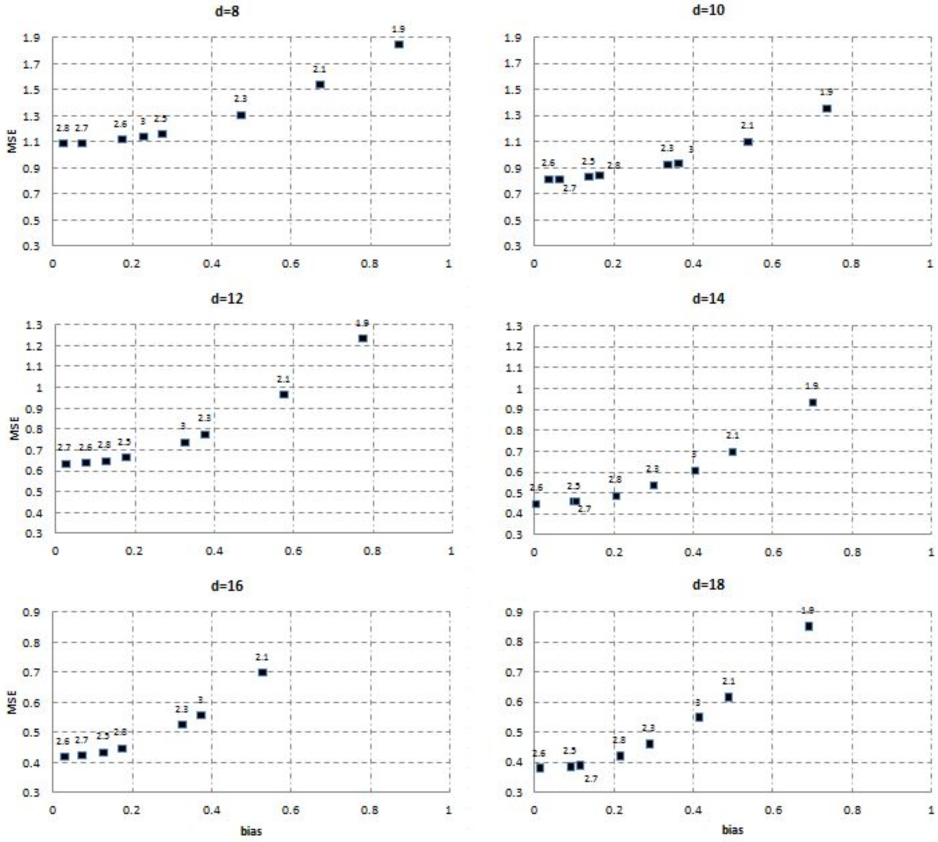


**Figure 7**: Bias versus MSE for the posterior mean using prior (7), adjusting for censoring ($n = 18$).

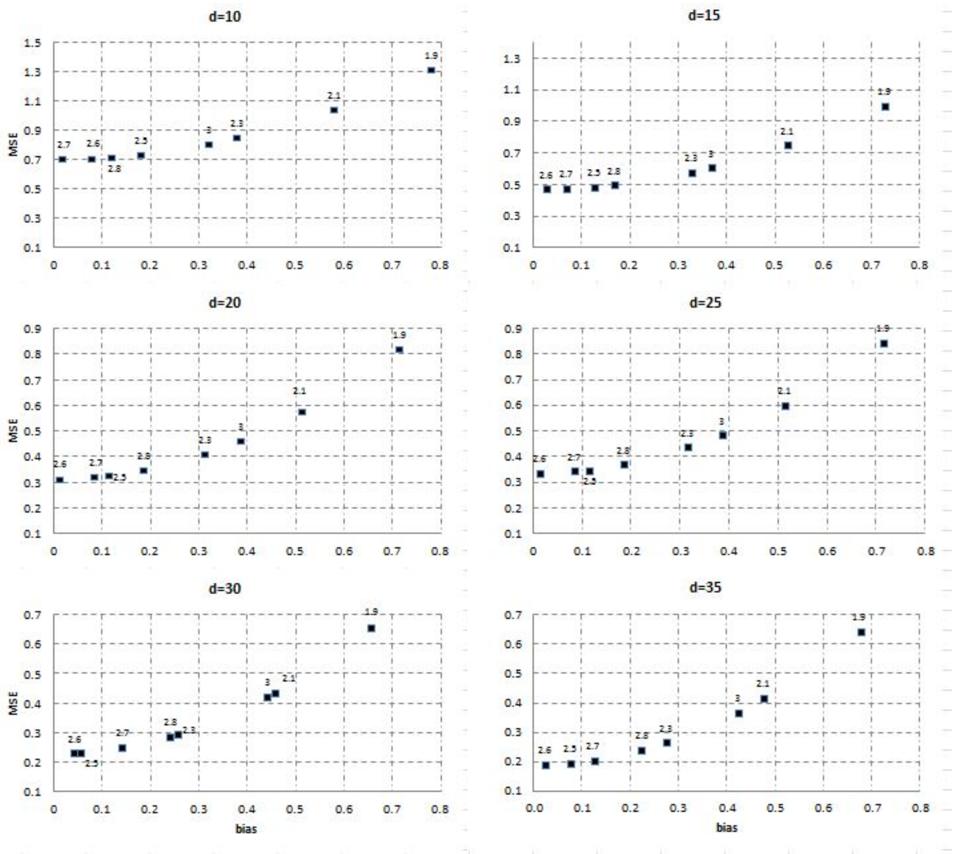


**Figure 8**: Bias versus MSE for the posterior mean using prior (7), adjusting for censoring ($n = 35$).

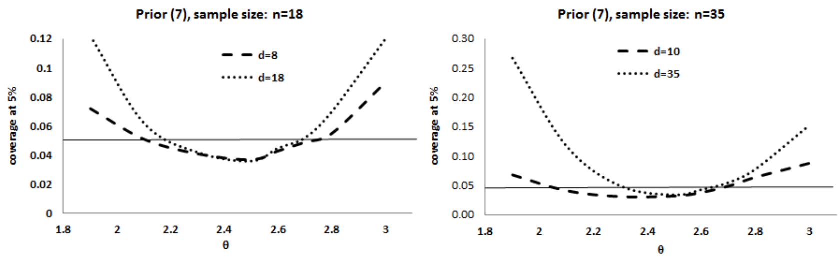


**Figure 9**: Coverage at $\alpha = 5\%$ using prior (7).

Figure 9 illustrates the sensitivity of Bayesian inference for $\theta$ outside the interval $[2.10; 2.60]$ with the effect of a larger sample size. It also shows a smaller coverage range compared with using other priors.

The posterior mean using prior (8) was finally determined, and MSE and bias were calculated using (18) for predetermined $\theta_0$ values in $H_0$. Figures 10 and 11 summarise the results for prior (8) using the posterior mean as Bayesian estimator for sample sizes of 18 and 35 respectively, while Figure 12 shows the corresponding coverage for different sample size scenarios and the extreme censoring schemes in this study, where similar conclusions can be drawn, as mentioned previously.

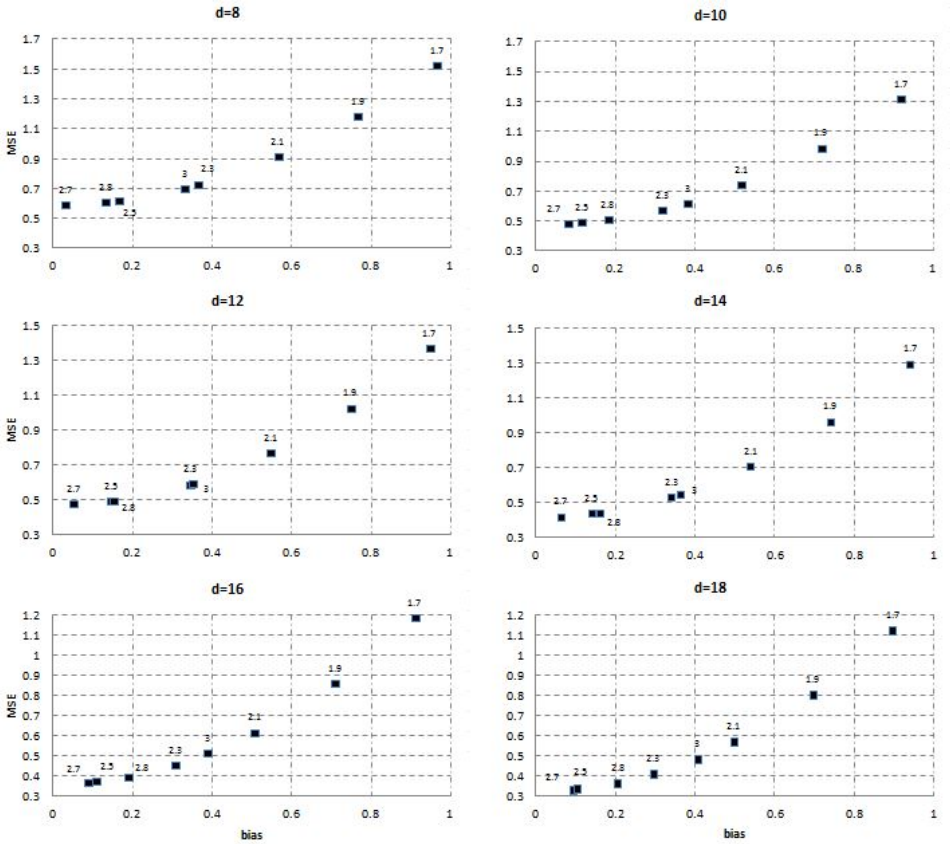


**Figure 10**: Bias versus MSE for the posterior mean using prior (8), adjusting for censoring ($n = 18$).

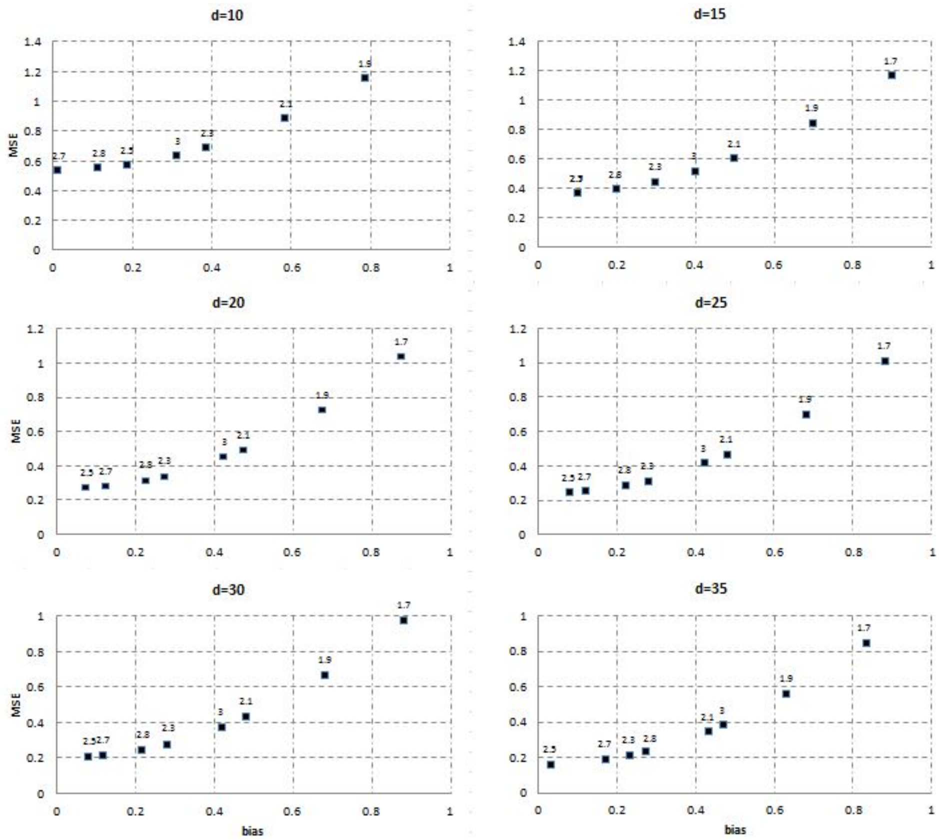


**Figure 11**: Bias versus MSE for the posterior mean using prior (8), adjusting for censoring ($n = 35$).

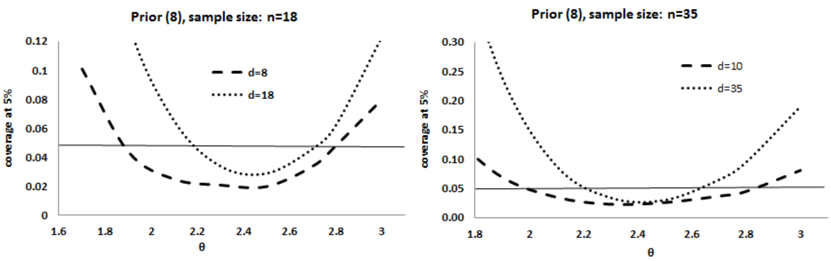


**Figure 12**: Coverage at $\alpha = 5\%$ using prior (8).

The graphs reflecting coverage illustrate the sensitivity of Bayesian inference for $\theta$ to sample size. For values of $\theta$ outside the intervals $[2.20; 2.70]$, the coverage decreases faster for larger sample sizes.

The simulation is extended to incorporate other characteristics of the different posteriors when censoring is present. Again, sample sizes of $n_1 = n = 18$ and $n_1 = n = 35$ were generated with the same set-up of events ($d$) mentioned earlier from the Rayleigh distribution with parameter $\theta = 2.5$. One thousand such samples were then generated, and the mean sufficient statistic $T$ was calculated for each data scenario that was used in the simulations. Tables 2 and 3 summarise the results for different sample sizes, including the posterior mean (Average), the posterior variance (Var), the Bayesian 95% credible interval (L-2.5% and U-97.5%) and the length of the credible interval (Length). These results are based on 10 000 variates simulated from the relevant posterior distributions using the Metropolis-Hastings algorithm.

**Table 2:** Posterior analyses based on random samples of size $n_1 = 18$.

| | d | T | Average | Var | L-2.5% | U-97.5% | Length |
|---|---|---|---|---|---|---|---|
| Prior (4): p=5; q=4; b=0.526; m=0.8 | 8 | 3.1669 | 2.4443 | 0.7274 | 1.0883 | 4.3592 | 3.2710 |
| | 10 | 3.9618 | 2.4381 | 0.6052 | 1.1869 | 4.1575 | 2.9705 |
| | 12 | 4.8167 | 2.4093 | 0.5085 | 1.2443 | 3.9772 | 2.7330 |
| | 14 | 5.6252 | 2.4436 | 0.4593 | 1.3081 | 3.9003 | 2.5922 |
| | 16 | 6.3849 | 2.4624 | 0.4038 | 1.3766 | 3.8091 | 2.4325 |
| | 18 | 7.2074 | 2.4452 | 0.3622 | 1.4327 | 3.7255 | 2.2927 |
| Prior (5): p=5; q=4; b=1.5; m=1.0 | 8 | 3.1669 | 2.4794 | **0.6566** | 1.1583 | 4.2797 | **3.1214** |
| | 10 | 3.9618 | 2.5088 | 0.5654 | 1.2711 | 4.1032 | **2.8320** |
| | 12 | 4.8167 | 2.4835 | 0.4925 | 1.3312 | 4.0547 | 2.7235 |
| | 14 | 5.6252 | 2.4613 | 0.4296 | 1.3758 | 3.8607 | 2.4849 |
| | 16 | 6.3849 | 2.4857 | 0.3911 | 1.4278 | 3.8366 | 2.4088 |
| | 18 | 7.2074 | 2.4857 | 0.3668 | 1.4670 | 3.7608 | 2.2938 |
| Prior (7): p=1; q=1.667; b=1.667; m=1.0 | 8 | 3.1669 | 2.4850 | 0.7808 | 1.1050 | 4.3884 | 3.2834 |
| | 10 | 3.9618 | 2.4829 | 0.6261 | 1.1890 | 4.1792 | 2.9902 |
| | 12 | 4.8167 | 2.4498 | 0.5276 | 1.2645 | 4.0374 | 2.7729 |
| | 14 | 5.6252 | 2.4571 | 0.4560 | 1.3246 | 3.9060 | 2.5813 |
| | 16 | 6.3849 | 2.4650 | 0.4231 | 1.3689 | 3.8443 | 2.4754 |
| | 18 | 7.2074 | 2.4717 | 0.3932 | 1.4021 | 3.7830 | 2.3809 |
| Prior (8): p=8; q=4.2; b=1.0; m=1.0 | 8 | 3.1669 | 2.5127 | 0.6727 | 1.2057 | 4.3488 | 3.1431 |
| | 10 | 3.9618 | 2.4874 | **0.5480** | 1.2783 | 4.1230 | 2.8447 |
| | 12 | 4.8167 | 2.4724 | **0.4869** | 1.3271 | 4.0045 | **2.6774** |
| | 14 | 5.6252 | 2.4750 | **0.4110** | 1.3951 | 3.8792 | **2.4841** |
| | 16 | 6.3849 | 2.4754 | **0.3730** | 1.4458 | 3.7911 | **2.3453** |
| | 18 | 7.2074 | 2.4755 | **0.3536** | 1.4712 | 3.7350 | **2.2638** |

**Table 3:** Posterior analyses based on random samples of size $n_1 = 35$.

| | d | T | Average | Var | L-2.5% | U-97.5% | Length |
|---|---|---|---|---|---|---|---|
| Prior (4): p=5; q=4; b=0.526; m=0.8 | 10 | 4.0201 | 2.4050 | 0.5849 | 1.1445 | 4.0808 | 2.9363 |
| | 15 | 5.9663 | 2.4792 | 0.4605 | 1.3524 | 3.9400 | 2.5875 |
| | 20 | 7.9561 | 2.4754 | **0.3319** | 1.5038 | 3.6822 | **2.1784** |
| | 25 | 9.9507 | 2.4891 | 0.2793 | 1.5884 | 3.5939 | 2.0054 |
| | 30 | 12.0038 | 2.4829 | 0.2455 | 1.6237 | 3.5036 | 1.8799 |
| | 35 | 14.0322 | 2.4740 | 0.2149 | 1.6696 | 3.4434 | 1.7738 |
| Prior (5): p=5; q=4; b=1.5; m=1.0 | 10 | 4.0201 | 2.4767 | 0.5633 | 1.2779 | 4.1233 | **2.8454** |
| | 15 | 5.9663 | 2.4853 | 0.4163 | 1.4127 | 3.8610 | **2.4484** |
| | 20 | 7.9561 | 2.4900 | 0.3335 | 1.4856 | 3.7193 | 2.2338 |
| | 25 | 9.9507 | 2.4937 | 0.2826 | 1.5827 | 3.6296 | 2.0469 |
| | 30 | 12.0038 | 2.4885 | 0.2373 | 1.6563 | 3.5056 | 1.8494 |
| | 35 | 14.0322 | 2.4802 | 0.2044 | 1.6916 | 3.4186 | 1.7270 |
| Prior (7): p=1; q=1.667; b=1.667; m=1.0 | 10 | 4.0201 | 2.4299 | 0.6115 | 1.1661 | 4.1242 | 2.9581 |
| | 15 | 5.9663 | 2.4975 | 0.4653 | 1.3571 | 3.9715 | 2.6144 |
| | 20 | 7.9561 | 2.4881 | 0.3454 | 1.4889 | 3.7109 | 2.2220 |
| | 25 | 9.9507 | 2.4970 | 0.2899 | 1.5629 | 3.6231 | 2.0602 |
| | 30 | 12.0038 | 2.4961 | 0.2519 | 1.6315 | 3.5246 | 1.8931 |
| | 35 | 14.0322 | 2.4899 | 0.2141 | 1.7026 | 3.4654 | 1.7628 |
| Prior (8): p=8; q=4.2; b=1.0; m=1.0 | 10 | 4.0201 | 2.4699 | **0.5551** | 1.2732 | 4.1458 | 2.8726 |
| | 15 | 5.9663 | 2.4812 | **0.4068** | 1.4128 | 3.8844 | 2.4717 |
| | 20 | 7.9561 | 2.4936 | 0.3378 | 1.5088 | 3.7273 | 2.2186 |
| | 25 | 9.9507 | 2.4942 | **0.2735** | 1.6070 | 3.5945 | **1.9875** |
| | 30 | 12.0038 | 2.4784 | **0.2321** | 1.6515 | 3.4909 | **1.8393** |
| | 35 | 14.0322 | 2.4887 | **0.2027** | 1.7051 | 3.4220 | **1.7169** |

Values highlighted in bold in Tables 2 and 3 indicate minima. Censoring and sample size play an important role in the interpretation of results. Light censoring (i.e. more observed actual events) indicates that less complicated priors may be used, while with heavy censoring some of the priors with more hyperparameters are more appropriate. Heavier censoring results in greater variance of assumed parameters under $H_0$ with relative good coverage, provided that the prior is correctly specified. Correctly specified priors around the true parameter should have only a small variability for parameters to yield an accurate inference.

The results are in line with the definition of the information ratio between two distributions. (Let $f_1(\theta)$ and $f_2(\theta)$ be two different pdfs for $\theta$. The information ratio between the second and the first distribution, denoted $R_{dist}^{2,1}$ , or more conveniently $R_{dist}$, is:

$$R_{dist} = V_1/V_2,$$

where $V_i$ is the variance of $f_i(\theta), i = 1,2$ (Pham-Gia, 1994, p.2180). ) Regardless of the complexity of a specific prior in this family (3), the results are more or less the same for a given mean and variance. With the prior information ratio as the criterion, it follows that for prior densities (5) and (8), $R_{prior} = \frac{var_2(\theta)}{var_4(\theta)} = b^2$. If $b < 1(> 1)$, then $g_2(\theta)$ provides an increase (decrease) in prior information over $g_4(\theta)$. Tables 2 and 3 also confirm as expected with $var_2(\hat{\theta}) > var_4(\hat{\theta})$, e.g. 0.5633 > 0.5551. Similar remarks can be made between different combinations of the different prior pdfs. Heavy censoring yields shorter intervals for priors (4) and (5); in contrast, light censoring yields yields shorter intervals (i.e. precise inference) for prior (7).

The simulation study about a future lifetime in a second/future sample of size $n_2$ involves a comparison in which samples are generated (of size $n = n_1 = 35$ and $n = n_1 = 18$) from a Rayleigh distribution with parameter $\theta = 2.5$. One thousand random samples of size $n_1$ (and $d$ known) were generated; for each of these 1 000 random samples, the predictive density of $Y_1$ is determined by applying the technique mentioned in remark (Section 4) to each of the 1 000 simulated posteriors. The average quantities over the 1 000 repetitions are summarised in Table 4 for a future sample of size $n_2 = 10$, in Table 5 for a future sample of size $n_2 = 20$ and in Table 6 for a future sample of size $n_2 = 30$. The characteristics displayed in these tables are the sample information ($n_1$, $d$ and $T$), the lower and upper boundaries of a 95% Bayesian credible interval for future lifetime $Y_1$ (L-2.5% and U-97.5%), the median of a future lifetime in a sample of size $n_2$ ($Med[Y_1]$), the mean of the predictive density ($E[Y_1]$), and the probability that a future lifetime is larger than an arbitrary value of $Y_1$, say $y = 10$.

**Table 4:** Predictive density characteristics of $Y_1$; sample size $n_2 = 10$.

| Prior | $n_1$ | d | T | L-2.5% | U-97.5% | Med[$Y_1$] | E[$Y_1$] | P($Y_1$>10) |
|---|---|---|---|---|---|---|---|---|
| Generalised F-distribution (4)<br>p=5; q=4; b=0.526; m=0.8 | 18 | 12 | 4.858 | 0.7694 | 10.0467 | 4.0915 | 4.4331 | 0.0315 |
| | 18 | 18 | 7.189 | 0.7804 | 9.9743 | 4.1331 | 4.4547 | 0.0290 |
| | 35 | 20 | 8.003 | 0.7806 | 9.9208 | 4.1288 | 4.4449 | 0.0273 |
| | 35 | 35 | 13.883 | 0.7891 | 9.8361 | 4.1602 | 4.4589 | 0.0246 |
| Type I compound gamma distribution (5)<br>p=5; q=4; b=1.5; m=1.0 | 18 | 12 | 4.783 | 0.7808 | 10.1514 | 4.1539 | 4.4936 | 0.0329 |
| | 18 | 18 | 7.098 | 0.7900 | 10.0566 | 4.1784 | 4.5009 | 0.0302 |
| | 35 | 20 | 7.973 | 0.7846 | 9.9528 | 4.1492 | 4.4646 | 0.0275 |
| | 35 | 35 | 14.017 | 0.7881 | 9.8047 | 4.1511 | 4.4482 | 0.0240 |
| F-distribution (7)<br>p=1; q=1.667; b=1.667; m=1.0 | 18 | 12 | 4.779 | 0.7778 | 10.1853 | 4.1428 | 4.4889 | 0.0340 |
| | 18 | 18 | 7.168 | 0.7850 | 10.0314 | 4.1554 | 4.4792 | 0.0300 |
| | 35 | 20 | 8.005 | 0.7860 | 9.9789 | 4.1502 | 4.4692 | 0.0290 |
| | 35 | 35 | 13.971 | 0.7883 | 9.8262 | 4.1527 | 4.4515 | 0.0243 |
| Beta-prime distribution (8)<br>p=8; q=4.2; b=1.0; m=1.0 | 18 | 12 | 4.803 | 0.7806 | 10.1122 | 4.1446 | 4.4808 | 0.0321 |
| | 18 | 18 | 7.204 | 0.7836 | 9.9649 | 4.1435 | 4.4629 | 0.0283 |
| | 35 | 20 | 8.091 | 0.7812 | 9.8949 | 4.1277 | 4.4407 | 0.0270 |
| | 35 | 35 | 14.030 | 0.7886 | 9.7972 | 4.1464 | 4.4435 | 0.0239 |

**Table 5:** Predictive density characteristics of $Y_1$; sample size $n_2 = 20$.

| Prior | $n_1$ | d | T | L-2.5% | U-97.5% | Med[$Y_1$] | E[$Y_1$] | P($Y_1$>10) |
|---|---|---|---|---|---|---|---|---|
| Generalised F-distribution (4) | 18 | 12 | 4.858 | 1.0880 | 14.2082 | 5.7862 | 6.2693 | 0.1450 |
| p=5; q=4; b=0.526; m=0.8 | 18 | 18 | 7.190 | 1.1037 | 14.1056 | 5.8450 | 6.2998 | 0.1449 |
| | 35 | 20 | 8.003 | 1.1040 | 14.0301 | 5.8390 | 6.2860 | 0.1432 |
| | 35 | 35 | 13.883 | 1.1160 | 13.9104 | 5.8834 | 6.3059 | 0.1425 |
| Type I compound gamma distribution (5) | 18 | 12 | 4.783 | 1.1043 | 14.3562 | 5.8745 | 6.3549 | 0.1510 |
| p=5; q=4; b=1.5; m=1.0 | 18 | 18 | 7.098 | 1.1172 | 14.2222 | 5.9092 | 6.3653 | 0.1500 |
| | 35 | 20 | 7.973 | 1.1096 | 14.0754 | 5.8678 | 6.3139 | 0.1451 |
| | 35 | 35 | 14.017 | 1.1145 | 13.8660 | 5.8706 | 6.2907 | 0.1412 |
| F-distribution (7) | 18 | 12 | 4.779 | 1.1000 | 14.4042 | 5.8588 | 6.3483 | 0.1515 |
| p=1; q=1.667; b=1.667; m=1.0 | 18 | 18 | 7.168 | 1.1102 | 14.1866 | 5.8766 | 6.3346 | 0.1480 |
| | 35 | 20 | 8.005 | 1.1116 | 14.1122 | 5.8692 | 6.3204 | 0.1464 |
| | 35 | 35 | 13.971 | 1.1148 | 13.8963 | 5.8727 | 6.2953 | 0.1417 |
| Beta-prime distribution (8) | 18 | 12 | 4.803 | 1.1039 | 14.3008 | 5.8613 | 6.3367 | 0.1493 |
| p=8; q=4.2; b=1.0; m=1.0 | 18 | 18 | 7.204 | 1.1081 | 14.0925 | 5.8598 | 6.3115 | 0.1456 |
| | 35 | 20 | 8.091 | 1.1048 | 13.9936 | 5.8374 | 6.2801 | 0.1425 |
| | 35 | 35 | 14.030 | 1.1153 | 13.8553 | 5.8639 | 6.2840 | 0.1405 |

**Table 6:** Predictive density characteristics of $Y_1$; sample size $n_2 = 30$.

| Prior | $n_1$ | d | T | L-2.5% | U-97.5% | Med[$Y_1$] | E[$Y_1$] | P($Y_1$>10) |
|---|---|---|---|---|---|---|---|---|
| Generalised F-distribution (4) | 18 | 12 | 4.858 | 1.3326 | 17.4014 | 7.0867 | 7.6783 | 0.2601 |
| p=5; q=4; b=0.526; m=0.8 | 18 | 18 | 7.189 | 1.3517 | 17.2759 | 7.1587 | 7.7157 | 0.2640 |
| | 35 | 20 | 8.003 | 1.3521 | 17.1833 | 7.1513 | 7.6988 | 0.2634 |
| | 35 | 35 | 13.883 | 1.3668 | 17.0366 | 7.2057 | 7.7231 | 0.2662 |
| Type I compound gamma distribution (5) | 18 | 12 | 4.783 | 1.3524 | 17.5827 | 7.1947 | 7.7831 | 0.2690 |
| p=5; q=4; b=1.5; m=1.0 | 18 | 18 | 7.098 | 1.3683 | 17.4186 | 7.2373 | 7.7958 | 0.2711 |
| | 35 | 20 | 7.973 | 1.3589 | 17.2387 | 7.1866 | 7.7329 | 0.2663 |
| | 35 | 35 | 14.017 | 1.3650 | 16.9823 | 7.1899 | 7.7045 | 0.2647 |
| F-distribution (7) | 18 | 12 | 4.779 | 1.3473 | 17.6415 | 7.1755 | 7.7751 | 0.2681 |
| p=1; q=1.667; b=1.667; m=1.0 | 18 | 18 | 7.168 | 1.3597 | 17.3750 | 7.1973 | 7.7582 | 0.2680 |
| | 35 | 20 | 8.005 | 1.3614 | 17.2839 | 7.1883 | 7.7409 | 0.2665 |
| | 35 | 35 | 13.971 | 1.3653 | 17.0195 | 7.1926 | 7.7102 | 0.2652 |
| Beta-prime distribution (8) | 18 | 12 | 4.803 | 1.3522 | 17.5173 | 7.1799 | 7.7621 | 0.2673 |
| p=8; q=4.2; b=1.0; m=1.0 | 18 | 18 | 7.204 | 1.3572 | 17.2597 | 7.1768 | 7.7300 | 0.2658 |
| | 35 | 20 | 8.091 | 1.3531 | 17.1385 | 7.1494 | 7.6915 | 0.2625 |
| | 35 | 35 | 14.030 | 1.3659 | 16.9692 | 7.1817 | 7.6963 | 0.2638 |

It is evident from Tables 4 to 6 that the lower and upper bounds are relatively insensitive to the assumed values of $p, q, b$ and $m$ regardless of the size of a future sample. The shorter 95% predictive intervals for these examples are found to be based on data observed from larger samples and more actual events $d$. The posterior probability in the approximate right quartile area is summarised under $P(Y_1 > 10)$, which indicates relatively low sensitivity to the chosen prior.

## 5.2. Insulating fluid example

The time to breakdown of an insulating fluid for a given voltage is used to illustrate the findings of this paper. The original data comprise time to breakdown at five levels of voltage (Nelson, 1982, p. 252). The data used for this paper represent the insulating fluid at 30kV and are given in Table 7 per 4 hour units in time. The original data set was measured in seconds; by rescaling the observed time units in hours, the simulation study results can coincide with the measurement of time in the data example. Lifetime values indicated by (*) are type II right censored observations (i.e. $d = 10$ with $n = 12$). The sufficient statistic is calculated using the likelihood of the Rayleigh model, (2), as $T = 3.787221$.

**Table 7:** Time to insulating fluid breakdown at 30kV (per 4 hour unit).

| | | | | | |
|---|---|---|---|---|---|
| 0.003472 | 0.009306 | 0.012986 | 0.061250 | 0.100694 | 0.102083 |
| 0.159028 | 0.203472 | 0.290278 | 1.097222 | 1.097222* | 1.097222* |

Table 8 summarises the posterior characteristics of the posterior distributions using the four different priors with the same hyperparameter values as in the simulation studies (see Table 1), while Figure 13 displays the Box-and-Whisker plots of the Rayleigh parameter and confirms the deduction about higher posterior probabilities when using certain priors.

**Table 8:** Posterior characteristics of the Rayleigh parameter.

| Prior | Median | Average | Var | L-2.5% | U-97.5% | Length |
|---|---|---|---|---|---|---|
| (4) | 2.44400 | 2.53769 | 0.65862 | 1.22700 | 4.32610 | 3.09910 |
| (5) | 2.51600 | 2.58507 | 0.59857 | 1.31095 | 4.26005 | 2.94910 |
| (7) | 2.48100 | 2.57617 | 0.68407 | 1.24095 | 4.37903 | 3.13808 |
| (8) | 2.49900 | 2.57632 | 0.59030 | 1.34093 | 4.22605 | 2.88513 |

Table 8 also confirms as before that $var_2(\hat{\theta}) > var_4(\hat{\theta})$ (i.e. 0.59857 > 0.59030). The same trend of inference is observed as in the simulation study.

Table 9 presents the summary statistics of the first failure in a future sample of size $n_2$. The corresponding posterior distributions of the first failure time were obtained from the posterior distributions of the Rayleigh parameter as described in remark 2 (Section 4).

Again, 95% predictive intervals are obtained when future sample sizes increase. It is evident from Table 9 that the lower and upper bounds are relatively insensitive to the assumed priors regardless of the size of a future sample.

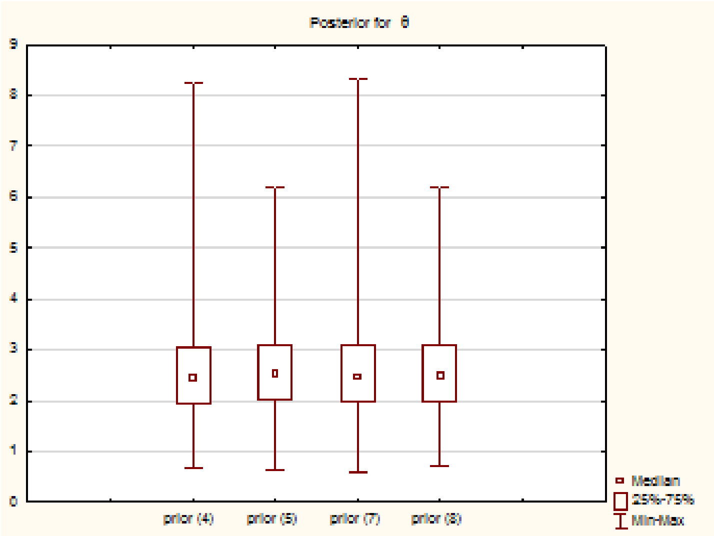


**Figure 13**: Box-and -Whisker plots representing the posteriors.

**Table 9:** Posterior characteristics of first future lifetime $Y_1$.

| Prior | $n_2$ | $\mathrm{Med}[Y_1]$ | $E[Y_1]$ | $\mathrm{Var}[Y_1]$ | L-2.5% | U-97.5% | Length |
|---|---|---|---|---|---|---|---|
| (4) | 10 | 0.16830 | 0.18346 | 0.01025 | 0.03260 | 0.41673 | 0.38413 |
| | 20 | 0.11901 | 0.12973 | 0.00512 | 0.02305 | 0.29467 | 0.27162 |
| | 30 | 0.09717 | 0.10592 | 0.00342 | 0.01882 | 0.24060 | 0.22178 |
| (5) | 10 | 0.16735 | 0.18159 | 0.01006 | 0.03186 | 0.41312 | 0.38126 |
| | 20 | 0.11833 | 0.12840 | 0.00503 | 0.02253 | 0.29212 | 0.26959 |
| | 30 | 0.09662 | 0.10484 | 0.00335 | 0.01840 | 0.23852 | 0.22012 |
| (7) | 10 | 0.16711 | 0.18163 | 0.01017 | 0.03267 | 0.41838 | 0.38571 |
| | 20 | 0.11817 | 0.12843 | 0.00508 | 0.02310 | 0.29584 | 0.27274 |
| | 30 | 0.09648 | 0.10486 | 0.00339 | 0.01886 | 0.24155 | 0.22269 |
| (8) | 10 | 0.16512 | 0.17994 | 0.01003 | 0.03202 | 0.41423 | 0.38221 |
| | 20 | 0.11676 | 0.12723 | 0.00502 | 0.02264 | 0.29290 | 0.27026 |
| | 30 | 0.09533 | 0.10389 | 0.00334 | 0.01849 | 0.23916 | 0.22067 |

## 6. Conclusion

In this paper, a generalised hypergeometric prior, with beta special cases, was proposed for the censored Rayleigh model. A simulation study illustrated that for light censoring (i.e. when more actual events are observed), less complicated priors may be used, while a more parameter-rich prior is appropriate for the heavy censoring cases. In conclusion, the sensitivity of Bayesian inference for the unknown parameter is illustrated when coverage decreases faster for larger sample sizes outside a range of incorrectly assumed parameter values, $\theta_0$ under $H_0$. Even with heavy censoring, the coverage remains high around the true parameter. More censoring present in a data set means that it is crucial that the correct prior is assumed, which has the ability to correct the estimates.

### Acknowledgements

We would like to thank the reviewer and the associate editor for the valuable comments and recommendations. This work is based upon research supported by the National Research Foundation South Africa (CPRR 13090132066 grant no. 91497).

## References

AL-BAYYATI, H. N. (2002). *Comparing Methods of Estimating Weibull Failure Models using Simulation*. Unpublished PhD thesis, Baghdad University: Iraq.

BERGER, J. (1980). A robust generalized Bayes estimator and confidence region for a multivariate normal mean. *The Annals of Statistics*, **8** (4), 716–761.

BHATTACHARYA, S. K. AND TYAGI, R. K. (1990). Bayesian survival analysis based on the Rayleigh model. *Trabajos de Estadistica*, **5** (1), 81–92.

DEY, S. AND DAS, M. K. (2007). A note on prediction interval for a Rayleigh distribution: Bayesian approach. *American Journal of Mathematical and Management Science*, **1&2**, 43–48.

DEY, S. AND DEY, T. (2011). Rayleigh distribution revisited via extension of Jeffreys prior information and a new loss function. *REVSTAT*, **9** (3), 213–226.

DEY, S. AND DEY, T. (2012). On Bayesian estimation and prediction intervals for a Rayleigh distribution under conjugate prior. *Journal of Statistics Computation and Simulation*, **82** (11), 1651–1660.

DUBEY, S. D. (1970). Compound gamma, beta and F distributions. *Metrika*, **16** (1), 27–31.

DURAN, B. S. AND BOOKER, J. M. (1988). A Bayesian sensitivity analysis when using the beta distribution as prior. *IEEE Transactions on Reliability*, **37** (2), 239–247.

ERDÉLYI, A. (1953). *Higher Transcendental Functions*, volume 1. McGraw-Hill: New York.

ERDÉLYI, A. (1954). *Tables of Integral Transforms*, volume 1. McGraw-Hill: New York.

FERREIRA, J., BEKKER, A., AND ARASHI, M. (2016). Objective Bayesian estimators for the right censored Rayleigh distribution: Evaluating the Al-Bayyati loss function. *To appear in REVSTAT*.

FOX, C. (1961). The G and H-functions as symmetrical Fourier kernels. *Transactions of the American Mathematical Society*, **98**, 395–429.

GELMAN, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, **1**, 515–533.

GELMAN, A. (2009). Bayes, Jeffreys, prior distributions and the philosophy of Statistics. *Statistical Science*, **24** (2), 176–178.

GUPTA, K. C. (1965). On the H-function. *Annales de la Société scientifique de Bruxelles*, **T79** (II), 97–106.

MATHAI, A. M. (1993). *A Handbook of Generalized Special Functions for Statistical and Physical Sciences*. Clarendon Press: Oxford.

MATHAI, A. M. AND SAXENA, R. K. (1966). On a generalized hypergeometric distribution. *Metrika*, **11** (2), 127–132.

MATHAI, A. M. AND SAXENA, R. K. (1978). *The H-function with Applications in Statistics and Other Disciplines*. John Wiley & Sons: New York.

MOSTERT, P. J. (1999). *A Bayesian Method to Analyse Cancer Lifetimes using Rayleigh Models*. Unpublished PhD thesis, University of South Africa: Pretoria, South Africa.

MOSTERT, P. J., BEKKER, A., AND ROUX, J. J. J. (1998). Bayesian analysis of survival data using the Rayleigh model and LINEX loss. *South African Statistical Journal*, **32** (1), 19–42.

NELSON, W. (1982). *Applied Data Analysis*. John Wiley & Sons: Toronto, Canada.

PHAM-GIA, T. (1994). Value of the beta prior information. *Communications in Statistics – Theory and Methods*, **23** (8), 2175–2195.

SEBASTIAN, N. (2011). A generalized gamma model associated with a Bessel function. *Integral Transforms and Special Functions*, **22** (9), 631–645.

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.

Manuscript received, 2015-09-15, revised, 2016-03-16, accepted, 2016-03-21.