STATISTICA, anno LXXVIII, n. 1, 2018

BAYESIAN INFERENCE AND PREDICTION FOR NORMAL DISTRIBUTION BASED ON RECORDS

Akbar Asgharzadeh [1]
Department of Statistics, Faculty of Mathematical Sciences, University of Mazandaran, Babolsar, Iran
Reza Valiollahi
Department of Statistics, Statistics and Computer Science, Semnan University, Semnan, Iran
Adeleh Fallah
Department of Statistics, Payame Noor University, P. O. Box 19395-3697, Tehran, Iran
Saralees Nadarajah
School of Mathematics, University of Manchester, Manchester M13 9PL, UK

1. Introduction

Let $X_1, X_2, \ldots$ be independent and identical random variables with the cumulative distribution function (cdf) $F(x)$ and the probability density function (pdf) $f(x)$. Define

$$Y_n = \max\{X_1, \ldots, X_n\}$$

for $n \geq 1$. Then, X_j is an upper record value of this sequence if $X_j > Y_{j-1}$, $j > 1$. Generally, if we define the sequence $\{U(n), n \geq 1\}$ as

$$U(1) = 1, \quad U(n) = \min\left\{j : j > U(n-1), X_j > X_{U(n-1)}\right\}$$

for $n \geq 2$, then $\left\{X_{U(n)}, n \geq 1\right\}$ is a sequence of upper record values. The sequence $\{U(n), n \geq 1\}$ represents the record times.

Chandler (1952) defined the theory of records as a model for successive extremes in a sequence of independent and identical random variables. Record data arise in many real life problems, such as in destructive stress testing, weather, hydrology, economics and sporting and athletic events. For more details and applications, see Ahsanullah (1995), Arnold *et al.* (1998) and Nevzorov (2000).

[1] Corresponding Author. E-mail: a.asgharzadeh@umz.ac.ir

In the frequentist set up, the estimation and prediction problems for normal distribution based on record data have been discussed by several authors. Balakrishnan and Chan (1998) obtained the best linear unbiased estimators (BLUEs) of the normal location and scale parameters, μ and σ, based on the first few upper record values. Using these BLUEs, they then developed a prediction interval for a future record value. Chacko and Mary (2013) discussed classical estimation and prediction for the normal distribution based on k-records. Sajeevkumar and Irshad (2014) estimated the location parameter of distributions with known coefficient of variation by record values.

The main aim of this paper is to consider estimation and prediction for normal distribution based on record data in the Bayesian set up. To the best of our knowledge, this problem has not been studied before in the literature. We compute Bayes estimators of μ and σ under squared error and Linex loss functions. It is observed that Bayes estimators can not be obtained in closed forms. We use the importance sampling procedure to generate samples from the posterior distributions and then compute the Bayes estimators. We then compare Bayes estimators with the maximum likelihood estimators (MLEs) and BLUEs by Monte Carlo simulations. We observe that Bayes estimators work quite well. Bayesian prediction of future records based on the first few upper records is also discussed. We use the importance sampling method to estimate the predictive distribution and then compute the Bayesian predictors.

The contents of this paper are organized as follows. In Section 2, we provide a brief review of frequentist estimators and predictors. In Section 3, the Bayes estimators of μ and σ are obtained using squared error and Linex loss functions. In Section 4, we discuss Bayesian prediction for future records based on the first few upper records. In Section 5, a real data set is analyzed for illustrative purposes. Monte Carlo simulations are performed to compare the proposed estimators and predictors in Section 6. Finally, we conclude the paper in Section 7.

2. Frequentist estimation and prediction: A review

Suppose that we observe the first n upper record values $X_{U(1)}=x_1, X_{U(2)}=x_2, \ldots, X_{U(n)}=x_n$ from the normal $N\left(\mu,\sigma^2\right)$ distribution. For notational simplicity, we will write X_i for $X_{U(i)}$. The likelihood function is given (see Arnold *et al.*, 1998) by

$$L(\mu,\sigma|\mathbf{x})=f(x_n;\mu,\sigma)\prod_{i=1}^{n-1}\frac{f(x_i;\mu,\sigma)}{1-F(x_i;\mu,\sigma)},$$

where $f(x_n;\mu,\sigma)$ and $F(x_n;\mu,\sigma)$ denote, respectively, the pdf and cdf of the $N\left(\mu,\sigma^2\right)$ distribution.

The likelihood function can be rewritten as

$$\begin{aligned} L(\mu,\sigma|\mathbf{x}) &= \frac{1}{\sigma}\phi\left(\frac{x_n-\mu}{\sigma}\right)\prod_{i=1}^{n-1}\left(\frac{\frac{1}{\sigma}\phi\left(\frac{x_i-\mu}{\sigma}\right)}{1-\Phi\left(\frac{x_i-\mu}{\sigma}\right)}\right) \\ &= \left(\frac{1}{\sigma}\right)^n\prod_{i=1}^{n}\phi\left(\frac{x_i-\mu}{\sigma}\right)\left(\prod_{i=1}^{n-1}\left[1-\Phi\left(\frac{x_i-\mu}{\sigma}\right)\right]\right)^{-1}, \end{aligned} \tag{1}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote, respectively, the pdf and cdf of a standard normal distribution.

The log-likelihood function is

$$L=\ln L(\mu,\sigma|\mathbf{x})=-n\ln\sigma+\sum_{i=1}^{n}\ln\phi\left(\frac{x_i-\mu}{\sigma}\right)-\sum_{i=1}^{n-1}\ln\left[1-\Phi\left(\frac{x_i-\mu}{\sigma}\right)\right]. \tag{2}$$

From (2), we obtain the likelihood equations for μ and σ as

$$\frac{\partial L}{\partial\mu}=\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)-\frac{1}{\sigma}\sum_{i=1}^{n-1}\frac{\phi\left(\frac{x_i-\mu}{\sigma}\right)}{1-\Phi\left(\frac{x_i-\mu}{\sigma}\right)}=0 \tag{3}$$

and

$$\frac{\partial L}{\partial\sigma}=-\frac{n}{\sigma}+\frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-\mu)^2-\frac{1}{\sigma^2}\sum_{i=1}^{n-1}\frac{(x_i-\mu)\phi\left(\frac{x_i-\mu}{\sigma}\right)}{1-\Phi\left(\frac{x_i-\mu}{\sigma}\right)}=0. \tag{4}$$

The equations (3) and (4) can be solved analytically to obtain $\widehat{\mu}_{ML}$ and $\widehat{\sigma}_{ML}$, the MLEs of μ and σ.

Following the generalized least-squares approach, the BLUEs of μ and σ can be derived as (see Balakrishnan and Cohen, 1991)

$$\widehat{\mu}_{BLU}=\sum_{i=1}^{n}a_iX_i, \qquad \widehat{\sigma}_{BLU}=\sum_{i=1}^{n}b_iX_i, \tag{5}$$

where

$$\boldsymbol{a}=\frac{\boldsymbol{\alpha}'\boldsymbol{\beta}^{-1}\boldsymbol{\alpha}\mathbf{1}'\boldsymbol{\beta}^{-1}-\boldsymbol{\alpha}'\boldsymbol{\beta}^{-1}\mathbf{1}\boldsymbol{\alpha}'\boldsymbol{\beta}^{-1}}{(\boldsymbol{\alpha}'\boldsymbol{\beta}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\boldsymbol{\beta}^{-1}\mathbf{1})-(\boldsymbol{\alpha}'\boldsymbol{\beta}^{-1}\mathbf{1})^2}$$

and

$$\boldsymbol{b}=\frac{\mathbf{1}'\boldsymbol{\beta}^{-1}\mathbf{1}\boldsymbol{\alpha}'\boldsymbol{\beta}^{-1}-\mathbf{1}'\boldsymbol{\beta}^{-1}\boldsymbol{\alpha}\mathbf{1}'\boldsymbol{\beta}^{-1}}{(\boldsymbol{\alpha}'\boldsymbol{\beta}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\boldsymbol{\beta}^{-1}\mathbf{1})-(\boldsymbol{\alpha}'\boldsymbol{\beta}^{-1}\mathbf{1})^2},$$

where $\boldsymbol{\alpha}' = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ is the moment vector with $\alpha_i = E(X_i)$ and $\beta = \left(\beta_{i,j}\right)$, $1 \leq i \leq j \leq n$ is the covariance matrix with $\beta_{i,j} = \text{Cov}\left(X_i, X_j\right)$, and $\mathbf{1}' = (1, 1, \ldots, 1)_{1\times n}$. The variances of these BLUEs are given by

$$\text{Var}(\widehat{\mu}_{BLU}) = \left[\frac{\boldsymbol{\alpha}'\beta^{-1}\boldsymbol{\alpha}}{(\boldsymbol{\alpha}'\beta^{-1}\boldsymbol{\alpha})(\mathbf{1}'\beta^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\beta^{-1}\mathbf{1})^2}\right]\sigma^2 = V_1\sigma^2$$

and

$$\text{Var}(\widehat{\sigma}_{BLU}) = \left[\frac{\mathbf{1}'\beta^{-1}\mathbf{1}}{(\boldsymbol{\alpha}'\beta^{-1}\boldsymbol{\alpha})(\mathbf{1}'\beta^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\beta^{-1}\mathbf{1})^2}\right]\sigma^2 = V_2\sigma^2.$$

The coefficients $\boldsymbol{a}$, $\boldsymbol{b}$ and the values of V_1 and V_2 can be found in Balakrishnan and Chan (1998, Tables 3 to 5). See also Arnold *et al.* (1998, Table 5.3.1, pages 139 and 140). From those tables, we can see that $\sum_{i=1}^{n} a_i = 1$ and $\sum_{i=1}^{n} b_i = 0$.

By using the BLUEs, one can construct confidence intervals (CIs) for the location and scale parameters, μ and σ, through pivotal quantities given by

$$R_1 = \frac{\widehat{\mu}_{BLU} - \mu}{\widehat{\sigma}_{BLU}\sqrt{V_1}}, \qquad R_2 = \frac{\widehat{\sigma}_{BLU} - \sigma}{\widehat{\sigma}_{BLU}\sqrt{V_2}}. \tag{6}$$

For constructing such CIs, we require the percentage points of R_1 and R_2, which can be computed by using the BLUEs $\widehat{\mu}_{BLU}$ and $\widehat{\sigma}_{BLU}$ via Monte Carlo simulations. Table 1 gives the percentage points of R_1 and R_2 based on ten thousand replications and different choices of n. The following algorithm was used to determine the percentage points:

1. set a value for n;
2. simulate $X_{U(1)}, X_{U(2)}, \ldots, X_{U(n)}$ from a standard normal distribution;
3. compute $\widehat{\mu}_{BLU}$ and $\widehat{\sigma}_{BLU}$ from (5);
4. then compute R_1 and R_2 from (6) by taking $\mu = 0$ and $\sigma = 1$;
5. repeat steps 2 to 4 ten thousand times, obtaining ten thousand estimates for R_1 and ten thousand estimates for R_2;
6. compute the percentage points of R_1 as the quantiles of the empirical distribution of the ten thousand estimates of R_1;
7. similarly, compute the percentage points of R_2 as the quantiles of the empirical distribution of the ten thousand estimates of R_2.

Let $R_{1(\alpha)}$ and $R_{2(\alpha)}$ denote the percentage points at α determined through simulation for the pivotal quantities R_1 and R_2, respectively. Then,

$$\left(\widehat{\mu}_{BLU} - \widehat{\sigma}_{BLU} R_{1(1-\alpha/2)}\sqrt{V_1},\ \widehat{\mu}_{BLU} - \widehat{\sigma}_{BLU} R_{1(\alpha/2)}\sqrt{V_1}\right)$$

and

$$\left(\widehat{\sigma}_{BLU} - \widehat{\sigma}_{BLU} R_{2(1-\alpha/2)}\sqrt{V_2},\ \widehat{\sigma}_{BLU} - \widehat{\sigma}_{BLU} R_{2(\alpha/2)}\sqrt{V_2}\right)$$

form $100(1-\alpha)$ percent CIs for μ and σ based on the pivotal quantities R_1 and R_2, respectively.

Note that if we define $Y = X_{U(n+1)}$ as the next upper record value, then we can predict this value by using the best linear unbiased prediction (BLUP) method. The BLUP of the next upper record value can be derived as (see Balakrishnan and Chan, 1998)

$$\widehat{Y}_{BLUP} = \widehat{\mu}_{BLU} + \widehat{\sigma}_{BLU}\alpha_{n+1}.$$

TABLE 1
Simulated percentage points of R_1 and R_2.

				Percentage points of R_1.				
n	1%	2.5%	5%	10%	90%	95%	97.5%	99%
2	-4.318	-2.765	-1.766	-1.165	2.975	4.926	7.187	9.945
3	-2.240	-1.633	-1.292	-1.003	2.587	3.970	5.662	8.125
4	-1.706	-1.430	-1.223	-0.997	2.253	3.506	4.906	6.795
5	-1.704	-1.443	-1.240	-0.999	2.038	3.107	4.193	5.974
6	-1.592	-1.375	-1.185	-0.944	2.140	3.136	4.218	5.572
7	-1.598	-1.390	-1.224	-1.030	1.825	2.776	3.778	4.930
8	-1.653	-1.441	-1.254	-1.033	1.876	2.717	3.542	4.709
9	-1.664	-1.442	-1.281	-1.048	1.767	2.548	3.384	4.368
10	-1.631	-1.453	-1.264	-1.045	1.740	2.439	3.219	4.222
				Percentage points of R_2.				
n	1%	2.5%	5%	10%	90%	95%	97.5%	99%
2	-8.822	-7.391	-5.666	-3.847	0.649	0.747	0.808	0.856
3	-8.137	-6.242	-4.624	-3.039	0.763	0.896	0.984	1.062
4	-7.120	-5.396	-3.934	-2.613	0.828	0.982	1.104	1.0198
5	-6.560	-4.822	-3.503	-2.332	0.877	1.033	1.160	1.275
6	-6.035	-4.411	-3.282	-2.214	0.898	1.064	1.186	1.311
7	-5.353	-3.938	-2.972	-2.024	0.928	1.112	1.247	1.387
8	-4.833	-3.623	-2.845	-1.949	0.927	1.123	1.276	1.723
9	-4.720	-3.591	-2.740	-1.960	0.948	1.145	1.292	1.440
10	-4.642	-3.437	-2.651	-1.864	0.935	1.138	1.295	1.471

3. BAYES ESTIMATION

In Bayesian inference, a loss function $L\left(\theta,\widehat{\theta}\right)$ describes the loss incurred by making an estimate $\widehat{\theta}$ when the true value of the parameter is θ. Here, we consider two different loss functions: first is the squared error loss function which is symmetric and second is the linear-exponential (Linex) loss function which is an asymmetric function. In the literature, the most commonly used loss function is the squared error. The symmetric nature of this function gives equal weight to overestimation as well as underestimation. However, for many situations, overestimation may be more serious than underestimation or vice-versa (see, for example, Feynman, 1987). Therefore, in order to make statistical inferences more practical and applicable, we often need to choose an asymmetric loss function. Many authors have considered asymmetric loss functions in reliability and life testing. See, for example, Basu and Ebrahimi (1991), Ahmadi *et al.* (2005), Ren *et al.* (2006), Raqab *et al.* (2007), Soliman and Al-Aboud (2008), Asgharzadeh and Fallah (2011), Kundu and Raqab (2012), Asgharzadeh *et al.* (2015).

One of the most popular asymmetric loss function is the Linex loss function

$$L\left(\theta,\widehat{\theta}\right)=e^{c\left(\widehat{\theta}-\theta\right)}-c\left(\widehat{\theta}-\theta\right)-1, \tag{7}$$

where $c\neq 0$. This loss function was introduced by Varrian (1975) and was extensively discussed by Zellner (1986). The sign and magnitude of the shape parameter c represents the direction and degree of asymmetry, respectively. If $c>0$, overestimation is more serious than underestimation, and vice-versa. When c is close to zero,

$$L\left(\theta,\widehat{\theta}\right)=\sum_{k=0}^{\infty}\frac{c^k\left(\widehat{\theta}-\theta\right)^k}{k!}-c\left(\widehat{\theta}-\theta\right)-1=\sum_{k=2}^{\infty}\frac{c^k\left(\widehat{\theta}-\theta\right)^k}{k!}\approx\frac{c^2}{2}\left(\widehat{\theta}-\theta\right)^2.$$

Hence, when c is close to zero, the Linex loss is approximately equal to the squared error loss and therefore almost symmetric.

The posterior expectation of the Linex loss function (7) is

$$E_\theta\left[L\left(\theta,\widehat{\theta}\right)\right]=e^{c\widehat{\theta}}E_\theta\left[e^{-c\theta}\right]-c\left(\widehat{\theta}-E_\theta(\theta)\right)-1, \tag{8}$$

where $E_\theta(\cdot)$ denotes the posterior expectation with respect to the posterior pdf of θ. The Bayes estimator of θ under the Linex loss function, denoted by $\widehat{\theta}_{BL}$, is the value of $\widehat{\theta}$ which minimizes (8). It is

$$\widehat{\theta}_{BL}=-\frac{1}{c}\ln\left\{E_\theta\left(e^{-c\theta}\right)\right\},$$

provided that the expectation $E_\theta\left[\exp(-c\theta)\right]$ exists and is finite.

Under the assumption that both the parameters μ and σ are unknown, specifying a general joint prior for μ and σ leads to computational complexities for the Bayes estimators. To solve this problem and simplify the Bayesian analysis, we can consider the joint prior pdf as a product of a conditional pdf of μ for given σ (which is taken to be the $N(\mu_0, \sigma^2)$ pdf, where $\mu_0 \in \Re$ is known) and a square-root inverted-gamma pdf for σ which has the form

$$\pi(\sigma) = \frac{a^b}{\Gamma(b)2^{b-1}} \sigma^{-2b-1} \exp\left(-\frac{a}{2\sigma^2}\right), \tag{9}$$

where $\sigma > 0$, $a > 0$ and $b > 0$. Note that the choice of a square-root inverted-gamma prior for σ is equivalent to selecting a gamma prior for $\lambda = 1/\sigma^2$. An improper prior for σ is $\pi(\sigma) \propto \sigma^{-1}$ (An improper prior is one that does not integrate to 1. There are no unique improper priors for a given prior. For the prior in (9), there are uncountably infinite ways of defining improper priors. The improper priors need not be particular cases of (9).) For more details on the square-root inverted-gamma distribution, see Raqab and Madi (2002), Wu *et al.* (2006) and Soliman and Al-Aboud (2008).

So, the joint prior pdf of μ and σ can be written as

$$\pi(\mu, \sigma) = \pi(\mu|\sigma)\pi(\sigma) \propto \sigma^{-2b-2} \exp\left\{\frac{-1}{2\sigma^2}\left[a + (\mu - \mu_0)^2\right]\right\}.$$

Now, by multiplying the likelihood function in (1) by the joint prior pdf, the posterior pdf of μ and σ can be derived as

$$\begin{aligned} \pi(\mu, \sigma|\mathbf{x}) \quad &\propto \quad \sigma^{-n-2b-2} \exp\left(\frac{-1}{2\sigma^2}\left[(n-1)s^2 + n(\overline{x} - \mu)^2\right]\right) \\ &\quad \cdot \exp\left(\frac{-1}{2\sigma^2}\left[a + (\mu - \mu_0)^2\right]\right) h(\mu, \sigma) \\ &= \quad \sigma^{-1} \exp\left(-\frac{n+1}{2\sigma^2}\left[\mu - \frac{n\overline{x} + \mu_0}{n+1}\right]^2\right) \\ &\quad \cdot \sigma^{-2\left(\frac{n}{2}+b\right)-1} \exp\left(\frac{-1}{2\sigma^2}\left[a + (n-1)s^2 + \frac{n(\overline{x} - \mu_0)^2}{n+1}\right]\right) h(\mu, \sigma), \end{aligned} \tag{10}$$

where

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

and

$$h(\mu, \sigma) = \left(\prod_{i=1}^{n-1}\left[1 - \Phi\left(\frac{x_i - \mu}{\sigma}\right)\right]\right)^{-1}. \tag{11}$$

The Bayes estimator of any function of μ and σ, say $k(\mu, \sigma)$, under the squared error loss is

$$\widehat{k}_{BS}(\mu, \sigma) = E_\pi\left[k(\mu, \sigma)|\mathbf{x}\right] = \int_0^\infty \int_{-\infty}^\infty k(\mu, \sigma)\pi(\mu, \sigma|\mathbf{x})\, d\mu d\sigma,$$

where $E_\pi(\cdot)$ denotes the expectation with respect to the joint posterior pdf $\pi(\mu,\sigma|\mathbf{x})$ in (10). This can not be reduced to a closed form.

3.1. Importance sampling method

Here, we consider the importance sampling method to generate samples from the posterior distribution and then compute the Bayes estimators of μ and σ under the squared error and Linex loss functions.

Based on the joint posterior pdf of μ and σ, (10) can be written as

$$\pi(\mu,\sigma|\mathbf{x}) \propto g_1(\mu|\sigma,\mathbf{x})\, g_2(\sigma|\mathbf{x})\, h(\mu,\sigma), \tag{12}$$

where $g_1(\mu|\sigma,\mathbf{x})$ denotes a normal pdf with mean $\frac{n\bar{x}+\mu_0}{n+1}$ and variance $\frac{\sigma^2}{n+1}$. Also $g_2(\sigma|\mathbf{x})$ denotes a square-root inverted-gamma pdf with shape parameter $\frac{n}{2}+b$ and scale parameter $\left[a+(n-1)s^2+\frac{n(\bar{x}-\mu_0)^2}{n+1}\right]$, and $h(\mu,\sigma)$ is given by (11).

Since $g_2(\sigma|\mathbf{x})$ is a square-root inverted-gamma pdf, it is simple to generate σ. Then, by using the generated σ, μ can be simulated from $g_1(\mu|\sigma,\mathbf{x})$ for a given $\mathbf{x}$. Now, similarly to Kundu and Pradhan (2009) and Kundu and Howlader (2010), we can use the importance sampling procedure to compute Bayes estimators. Using (12), the Bayes estimator of $k(\mu,\sigma)$ can be written as

$$\begin{aligned}\widehat{k}_{BS}(\mu,\sigma)=E_\pi[k(\mu,\sigma)|\mathbf{x}] &= \frac{\displaystyle\int_0^\infty\int_{-\infty}^\infty k(\mu,\sigma)h(\mu,\sigma)\, g_1(\mu|\sigma,\mathbf{x})\, g_2(\sigma|\mathbf{x})\,d\mu d\sigma}{\displaystyle\int_0^\infty\int_{-\infty}^\infty h(\mu,\sigma)g_1(\mu|\sigma,\mathbf{x})\, g_2(\sigma|\mathbf{x})\,d\mu d\sigma}\\ &= \frac{E_{\pi'}[k(\mu,\sigma)h(\mu,\sigma)]}{E_{\pi'}[h(\mu,\sigma)]},\end{aligned}$$

where $E_{\pi'}(\cdot)$ denotes the expectation with respect to the joint pdf

$$\pi'(\mu,\sigma|\mathbf{x}) = g_1(\mu|\sigma,\mathbf{x})\, g_2(\sigma|\mathbf{x}).$$

Now, the Bayes estimator of $k(\mu,\sigma)$ can be approximated by

$$\widehat{k}_{BS}(\mu,\sigma)=\frac{\displaystyle\frac{1}{N}\sum_{i=1}^N k(\mu_i,\sigma_i)\,h(\mu_i,\sigma_i)}{\displaystyle\frac{1}{N}\sum_{i=1}^N h(\mu_i,\sigma_i)}$$

for a random sample $(\mu_1,\sigma_1),\ldots,(\mu_N,\sigma_N)$ generated from $\pi'(\mu,\sigma|\mathbf{x})$. Here, by making an adjustment to π, we compensate for sampling from π' instead of π. The distribution

π' is referred to as the importance distribution. The distribution π is referred to as the nominal distribution.

Therefore, Bayes estimators can be computed using the following algorithm:

step 1. Generate σ_1 from $g_2(\sigma|\mathbf{x})$ and μ_1 from $g_1(\mu|\sigma_1,\mathbf{x})$.

step 2. Repeat step 1, N times to obtain $(\mu_1,\sigma_1),\ldots,(\mu_N,\sigma_N)$.

step 3. Bayes estimators of μ and σ under the squared error loss function are

$$\widehat{\mu}_{BS}^{MC}=\sum_{i=1}^{N}\mu_i w_i, \qquad \widehat{\sigma}_{BS}^{MC}=\sum_{i=1}^{N}\sigma_i w_i.$$

The estimators based on the Linex loss function are

$$\widehat{\mu}_{BL}^{MC}=-\frac{1}{c}\ln\left(\sum_{i=1}^{N}e^{-c\mu_i}w_i\right), \qquad \widehat{\sigma}_{BL}^{MC}=-\frac{1}{c}\ln\left(\sum_{i=1}^{N}e^{-c\sigma_i}w_i\right),$$

where

$$w_j=\frac{h\left(\mu_j,\sigma_j\right)}{\sum_{j=1}^{N}h\left(\mu_j,\sigma_j\right)} \tag{13}$$

for $j=1,2,\ldots,N$.

We implemented the above algorithm in the R software (R Development Core Team, 2016), using our own codes. The samples needed for construction of highest posterior density (HPD) credible intervals were also generated by using our own codes in the R software. All computational procedures for this paper were implemented using our own codes in the R software. None of the contributed packages in the R software were used.

3.2. HPD credible intervals

To construct HPD intervals, we use the Monte Carlo procedure proposed by Chen and Shao (1999). Given the Monte Carlo samples $\left(\mu_j,\sigma_j\right)$, $j=1,2,\ldots,N$, we compute the HPD interval for μ as follows:

Step 1. Sort $\left\{\mu_j, j=1,\ldots,N\right\}$ as

$$\mu_{(1)}\le\mu_{(2)}\le\cdots\le\mu_{(N)}.$$

Step 2. Compute the $100(1-\alpha)$ percent credible intervals

$$L_i(N)=\left(\widetilde{\mu}^{\left(\frac{i}{N}\right)},\ \widetilde{\mu}^{\left(\frac{i+[(1-\alpha)N]}{N}\right)}\right)$$

for $i = 1,2,\ldots,N-[(1-\alpha)N]$, where

$$\tilde{\mu}^{(\gamma)} = \begin{cases} \mu_{(1)}, & \text{if } \gamma = 0, \\ \mu_{(i)}, & \text{if } \sum_{j=1}^{i-1} w_j < \gamma \le \sum_{j=1}^{i} w_j. \end{cases}$$

Here, $[(1-\alpha)N]$ denotes the integer part of $(1-\alpha)N$.

Step 3. The $100(1-\alpha)$ percent HPD interval is the one with the smallest interval width among all $L_i(N)$'s.

The same procedure can be applied to calculate the HPD interval for σ.

4. BAYESIAN PREDICTION

Suppose that we observe only the n upper record observations $X_{U(1)} = x_1, X_{U(2)} = x_2, \ldots, X_{U(n)} = x_n$, and the aim is to predict the sth upper record value, $s > n$. Let $Y \equiv X_{U(s)}$ be the sth upper record value.

The conditional distribution of Y given $X = (x_1, x_2, \ldots, x_n)$ is just the distribution of Y given $X_{U(n)} = x_n$, due to the well-known Markovian property of record statistics. It follows (see Arnold *et al.* (1998)) that

$$f(y|x_n;\mu,\sigma) = \frac{[H(y)-H(x_n)]^{s-n-1}}{\Gamma(s-n)} \frac{f(y;\mu,\sigma)}{1-F(x_n;\mu,\sigma)},$$

where $y > x_n$ and $H(y) = -\ln[1-F(y)]$. For the normal distribution, $f(y|x_n;\mu,\sigma)$ is given by

$$f(y|x_n;\mu,\sigma) = \left[\ln\left(\frac{1-\Phi\left(\frac{x_n-\mu}{\sigma}\right)}{1-\Phi\left(\frac{y-\mu}{\sigma}\right)}\right)\right]^{s-n-1} \frac{\phi\left(\frac{y-\mu}{\sigma}\right)}{\sigma\Gamma(s-n)\left[1-\Phi\left(\frac{x_n-\mu}{\sigma}\right)\right]}.$$

The Bayes predictive pdf of Y, $f_s^*(y|x_n)$, can be calculated as

$$f_s^*(y|x_n) = \int_0^\infty \int_{-\infty}^\infty f(y|x_n;\mu,\sigma)\,\pi(\mu,\sigma|\mathbf{x})\,d\mu\,d\sigma. \tag{14}$$

Substituting (12) into (14), the Bayes predictive pdf $f_s^*(y|x_n)$ can be obtained as

$$f_s^*(y|x_n) \propto \int_0^\infty \int_{-\infty}^\infty f(y|x_n;\mu,\sigma)\,g_1(\mu|\sigma,\mathbf{x})\,g_2(\sigma|\mathbf{x})\,h(\mu,\sigma)\,d\mu\,d\sigma.$$

4.1. Bayesian point prediction

The Bayesian point predictors of Y under the squared error loss function, $\widehat{Y}_{SEP}$, and under the Linex loss function, $\widehat{Y}_{LEP}$, are

$$\widehat{Y}_{SEP} = \int_{x_n}^{\infty} y f_s^*(y|x_n)\,dy \tag{15}$$

and

$$\widehat{Y}_{LEP} = -\frac{1}{c}\ln\left[\int_{x_n}^{\infty} e^{-cy} f_s^*(y|x_n)\right]dy. \tag{16}$$

Since $f_s^*(y|x_n)$ can not be expressed in closed form, (15) and (16) can not computed explicitly. As before, based on the Monte Carlo sample $\{(\mu_i, \sigma_i), i = 1, 2, \ldots, M\}$, a simulation consistent estimator of $f_s^*(y|x_n)$ can be obtained as

$$\widehat{f_s^*}(y|x_n) = \sum_{i=1}^{M} f(y|x_n; \mu_i, \sigma_i)\, w_i, \tag{17}$$

where w_i, $i = 1, 2, \ldots, N$ are as defined in (13). By using (17), $\widehat{Y}_{SEP}$ and $\widehat{Y}_{LEP}$ can be computed as

$$\widehat{Y}_{SEP} = \int_{x_n}^{\infty} y \sum_{i=1}^{M} f(y|x_n; \mu_i, \sigma_i)\, w_i\, dy = \sum_{i=1}^{M} w_i I(x_n, \mu_i, \sigma_i)$$

and

$$\widehat{Y}_{LEP} = -\frac{1}{c}\ln\left(\int_{x_n}^{\infty} e^{-cy} \sum_{i=1}^{M} f(y|x_n; \mu_i, \sigma_i)\, w_i\, dy\right) = -\frac{1}{c}\ln\left(\sum_{i=1}^{M} w_i J(x_n, \mu_i, \sigma_i)\right),$$

where

$$I(x_n, \mu, \sigma) = \int_{x_n}^{\infty} y \left[\ln\left(\frac{1-\Phi\left(\frac{x_n-\mu}{\sigma}\right)}{1-\Phi\left(\frac{y-\mu}{\sigma}\right)}\right)\right]^{s-n-1} \frac{\phi\left(\frac{y-\mu}{\sigma}\right)}{\sigma\Gamma(s-n)\left[1-\Phi\left(\frac{x_n-\mu}{\sigma}\right)\right]}dy$$

and

$$J(x_n, \mu, \sigma) = \int_{x_n}^{\infty} e^{-cy} \left[\ln\left(\frac{1-\Phi\left(\frac{x_n-\mu}{\sigma}\right)}{1-\Phi\left(\frac{y-\mu}{\sigma}\right)}\right)\right]^{s-n-1} \frac{\phi\left(\frac{y-\mu}{\sigma}\right)}{\sigma\Gamma(s-n)\left[1-\Phi\left(\frac{x_n-\mu}{\sigma}\right)\right]}dy.$$

4.2. Bayesian interval prediction

Bayesian prediction intervals can be obtained from the Bayes predictive pdf $f^*(y|\mathbf{x})$. Bayesian prediction bounds can be obtained by evaluating

$$P(Y>\lambda|x_n)=\int_\lambda^\infty f_s^*(y|x_n)\,dy$$

for some positive λ. Now, the $100(1-\alpha)$ percent Bayesian prediction interval for Y is given by $(L(\mathbf{x}),U(\mathbf{x}))$, where $L(\mathbf{x})$ and $U(\mathbf{x})$ can be obtained by solving the following non-linear equations simultaneously

$$P(Y>L(\mathbf{x})|x_n)=\int_{L(\mathbf{x})}^\infty f_s^*(y|x_n)\,dy=1-\frac{\alpha}{2}$$

and

$$P(Y>U(\mathbf{x})|x_n)=\int_{U(\mathbf{x})}^\infty f_s^*(y|x_n)\,dy=\frac{\alpha}{2}.$$

By substituting $\widehat{f_s^*}(y|x_n)$ in (17) for $f_s^*(y|x_n)$, we can obtain the Bayesian prediction bounds $L(\mathbf{x})$ and $U(\mathbf{x})$ from the following equations:

$$1-\frac{\alpha}{2}=\sum_{i=1}^M w_i K(L(\mathbf{x}),\mu_i,\sigma_i), \qquad \frac{\alpha}{2}=\sum_{i=1}^M w_i K(U(\mathbf{x}),\mu_i,\sigma_i),$$

where

$$K(L,\mu,\sigma)=\int_L^\infty \left[\ln\left(\frac{1-\Phi\left(\frac{x_n-\mu}{\sigma}\right)}{1-\Phi\left(\frac{y-\mu}{\sigma}\right)}\right)\right]^{s-n-1} \frac{\phi\left(\frac{y-\mu}{\sigma}\right)}{\sigma\Gamma(s-n)\left[1-\Phi\left(\frac{x_n-\mu}{\sigma}\right)\right]}dy.$$

5. Real data analysis

In this section, we consider a real data set to illustrate all the estimation and prediction methods described in the preceding sections. The data are the total annual rainfall in inches during March recorded at Los Angeles Civic Center from 1997 to 2006 (see the website of Los Angeles Almanac, www.laalman-ac.com/weather/we08aa.htm):

```
0.00,  4.06, 1.24, 2.82, 1.17, 0.32, 4.31, 1.17, 2.14, 2.87.
```

We observed that the normal distribution with $\mu=2.01$ and $\sigma=1.483$ fits to above data set. We checked the validity of the normal distribution based on the Kolmogorov-Smirnov (K-S) test. The K-S distance was 0.198 and the corresponding p-value was 0.827.

From this data, we observe the three upper record values as 0.00, 4.06, 4.31. For this record data, we estimated μ and σ using MLE, BLUE of Balakrishnan and Chan (1998) and Bayes estimators. For computing Bayes estimators and predictors, since we do not have any prior information, we assume that the prior on σ is improper. Since μ_0 is the location hyper parameter, without loss of generality, we assume that $\mu_0 = 0$. The integrals I, J in Section 4.1 and K in Section 4.2 were computed numerically using the function integrate in the R software (R Development Core Team, 2016). integrate is based on QUADPACK routines dqags and dqagi by R. Piessens and E. deDoncker-Kapenga, available from Netlib. The number of subdivisions for integration was taken to be 100.

Figure 1 plots the p-value of the K-S test versus c when Bayes estimator is used with the Linex loss function. Also plotted are the p-values corresponding to the MLE, BLUE and Bayes estimator with the squared error loss function. We see that BLUE has the smallest p-value and MLE has the second smallest p-value. The p-values for Bayes estimators are larger. The asymmetric Linex loss always produces larger p-values than the squared error loss.

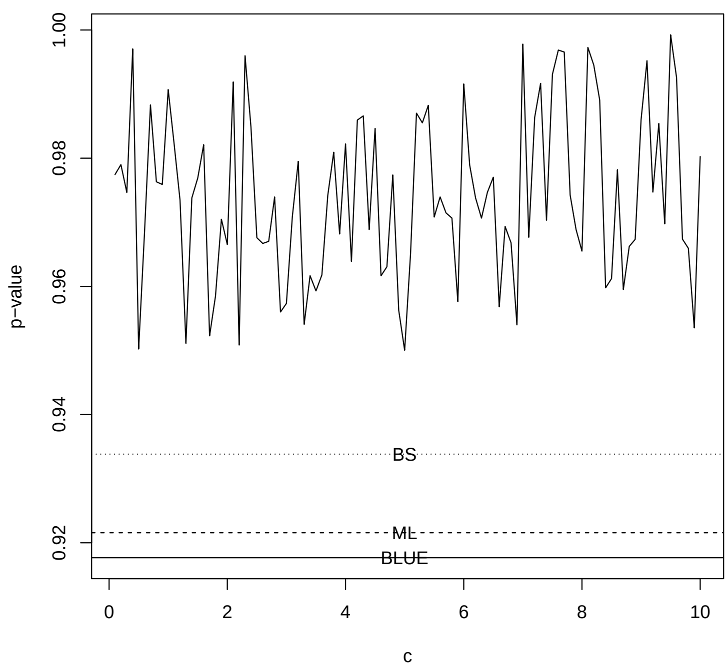

Figure 1 – p-value of the K-S test versus c of the Linex loss function.

We also computed CIs for μ and σ. The 95 percent HPD intervals using the Monte Carlo method were (1.95, 2.11) and (1.481, 1.500) for μ and σ, respectively. The 95 percent CIs based on the BLUEs through the pivotal quantities R_1 and R_2 were (1.91, 2.22) and (1.476, 1.511) for μ and σ, respectively. We see that the former intervals are shorter. All of the intervals contain the MLEs of μ and σ based on the complete sample ($\mu = 2.01$ and $\sigma = 1.483$).

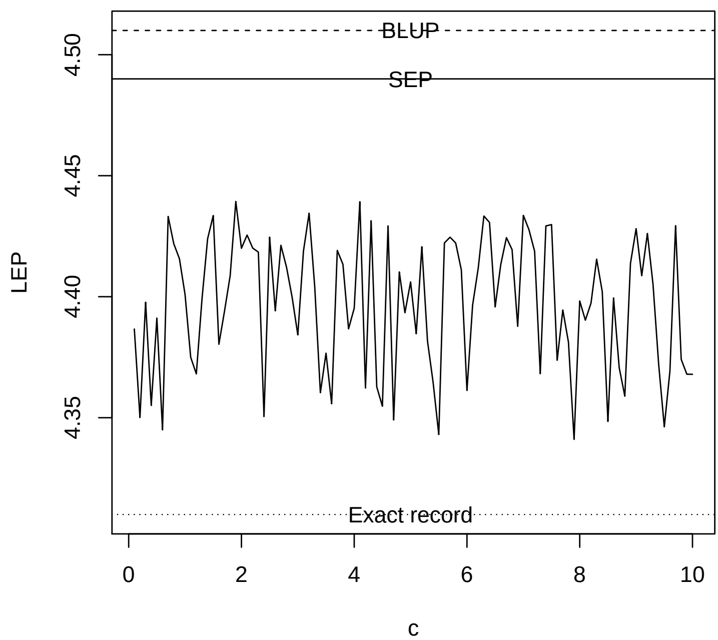

Figure 2 – Estimates of last record versus c of the Linex loss function.

Let us now compare the Bayesian prediction and the BLUP given by Balakrishnan and Chan (1998). In order to make the comparison, the first two upper records are used to predict the last one. Based on the first two upper records, the BLUP of $= X_{U(3)}$ was 4.51. Now, under the assumption that the prior distribution is as before, we computed the Bayesian point predictors with respect to the two loss functions. Figure 2 plots the predicted value versus c when Bayes estimator is used with the Linex loss function. Also plotted are the predicted values corresponding to the BLUP and Bayes estimator with the squared error loss function. We see that BLUP is the furthest from the exact record. The one based on the squared error loss function is the second furthest from the exact record. The predictions given by the asymmetric Linex loss are the closest to the exact record.

We also computed the 95 percent Bayesian prediction interval of the last record. It was $(4.22, 4.56)$, containing the exact value.

6. SIMULATION AND DISCUSSION

In this section, different estimation and prediction methods are compared using a Monte Carlo simulation. We compare the performances of the MLEs, the BLUEs and Bayes point estimators (with respect to the squared error and Linex loss functions) in terms of biases, and mean squared errors (MSEs). We also compare two CIs, namely, the CIs based on BLUEs, asymptotic MLEs, bootstrapping and the HPD intervals based on the Monte Carlo method in terms of average confidence lengths, and coverage probabilities. For computing Bayes estimators and predictors, we assume two priors:

$$\text{Prior 1:}\quad \mu_0 = 0,\quad a = 0,\quad b = 0,$$
$$\text{Prior 2:}\quad \mu_0 = 0,\quad a = 2,\quad b = 3.$$

Obviously, prior 2 is more informative than prior 1.

The simulations were performed as follows:

1. simulate n upper record values from the standard normal distribution;
2. compute the MLEs, BLUEs and Bayes estimators based on the squared error and Linex loss functions. For the Linex loss function, we took $c = -1, 0.1, 1$;
3. compute the 95 percent CIs based on BLUEs, asymptotic MLEs, bootstrapping and the 95 percent HPD intervals based on Monte Carlo simulations;
4. iterate over steps 1 to 3 ten thousand times.

This scheme gives for a given n the biases and MSEs of the MLEs, BLUEs and Bayes estimators. It also gives for a given n the average confidence lengths and coverage probabilities of the intervals based on BLUEs, asymptotic MLEs, bootstrapping and the HPD intervals based on Monte Carlo simulations.

Figure 3 plots the biases and MSEs of the estimators of μ and σ versus $n = 5, 6, \ldots, 10$. Figure 4 plots the confidence lengths and coverage probabilities of the estimators of μ and σ versus $n = 5, 6, \ldots, 10$.

We can observe the following from Figure 3: the biases for the BLUE are the closest to zero as expected; the biases for the other estimators generally decrease to zero as n increases; of these, the MLEs have the largest biases, the Bayes estimators based on the squared error loss have the second largest biases and the Bayes estimators based on the Linex loss have the smallest biases; the MSEs of all the estimators generally decrease to zero as n increases; the BLUEs have the largest MSEs, the MLEs have the second largest MSEs, the Bayes estimators based on the squared error loss have the third largest MSEs and the Bayes estimators based on the Linex loss have the smallest MSEs; the use of prior 2 leads to smaller biases and smaller MSEs.

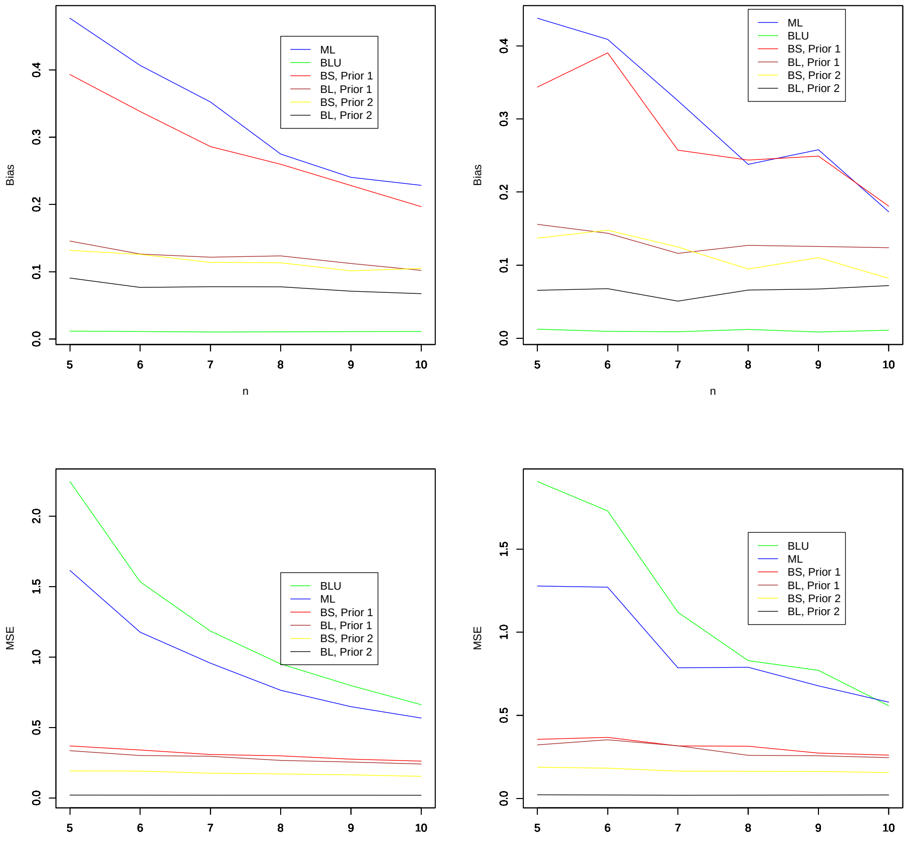

Figure 3 – Biases and MSEs versus n: bias for μ (top left), bias for σ (top right), MSE for μ (bottom left) and MSE for σ (bottom right).

We can observe the following from Figure 4: confidence length generally decreases with increasing n; they appear largest for the BLUE CIs, second largest for the asymptotic MLE intervals, third largest bootstrap based intervals, fourth largest for the HPD intervals based on prior 1 and smallest for the HPD intervals based on prior 2; coverage probabilities appear furthest from the nominal level for the BLUE CIs and closest to the nominal level for the HPD intervals based on prior 1 and prior 2.

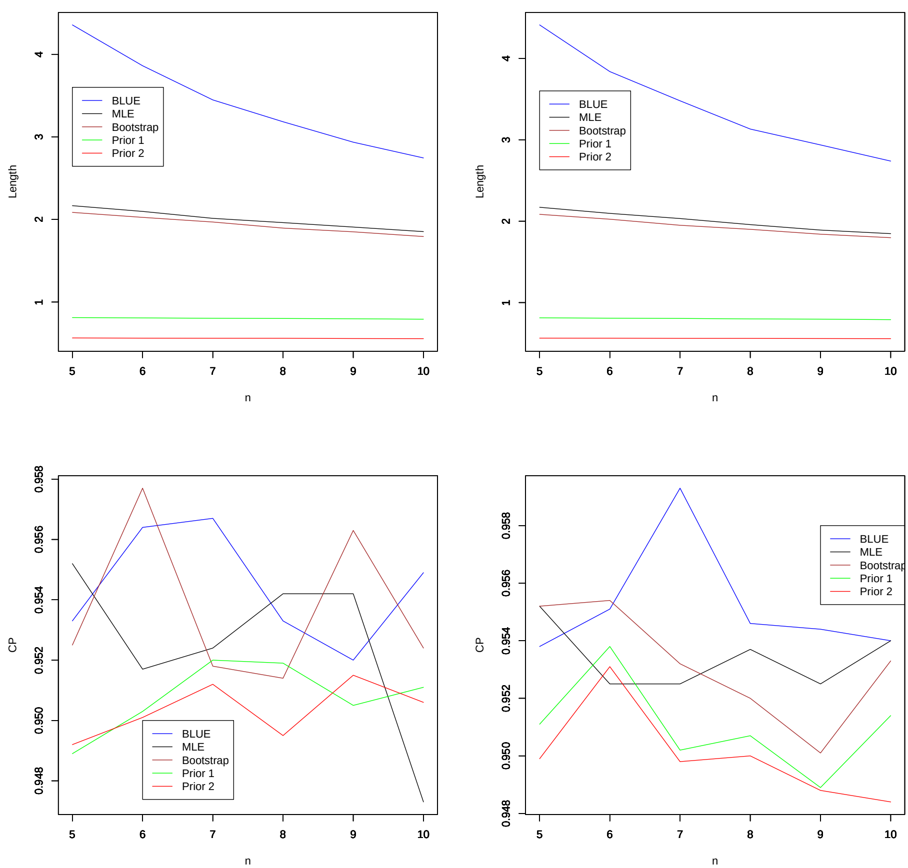

Figure 4 – Confidence lengths and coverage probabilities versus n: coverage length for μ (top left), coverage length for σ (top right), coverage probability for μ (bottom left) and coverage probability for σ (bottom right).

In the context of computational complexities, the Bayesian point estimators are easy to compute. They do not involve solving of non-linear equations. The MLEs involve solving of non-linear equations through some iterative processes. Also BLUEs require the construction of special tables based on means, variances, and covariances of record statistics.

To see how the Bayes predictors based on priors 1 and 2 compare to each other, we also carried out a Monte Carlo simulation. We simulated a sample of ten upper record values from the standard normal distribution and used the first four to predict the sth record for $s = 5, 6, \ldots, 10$. We computed the Bayesian point estimators based on the squared error and Linex loss functions. We also computed 95 percent Bayesian PIs based on the squared error and Linex loss functions. The biases, MSEs, confidence lengths and coverage probabilities for every s were computed over ten thousand iterations as described before. Figure 5 plots the biases and MSEs of the predictors versus $s = 5, 6, \ldots, 10$. Figure 6 plots the confidence lengths and coverage probabilities of the Bayesian PIs versus $s = 5, 6, \ldots, 10$.

We can observe the following from Figure 5: the biases generally increase with increasing s; the biases appear larger when the squared error loss is used and smaller when the Linex loss is used; the use of prior 2 leads to smaller biases; the MSEs generally increase with increasing s; the MSEs appear larger when the squared error loss is used and smaller when the Linex loss is used; the use of prior 2 leads to smaller MSEs.

We can observe the following from Figure 6: confidence length generally increases with increasing s; they appear larger for prior 1 and smaller for prior 2; coverage probabilities appear further from the nominal level for prior 1 and closer to the nominal level for prior 2.

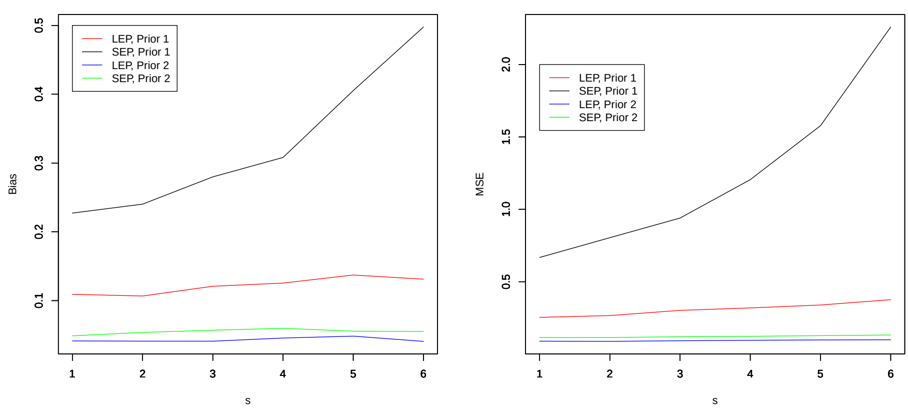

Figure 5 – Biases and MSEs versus s: bias (left) and MSE (right).

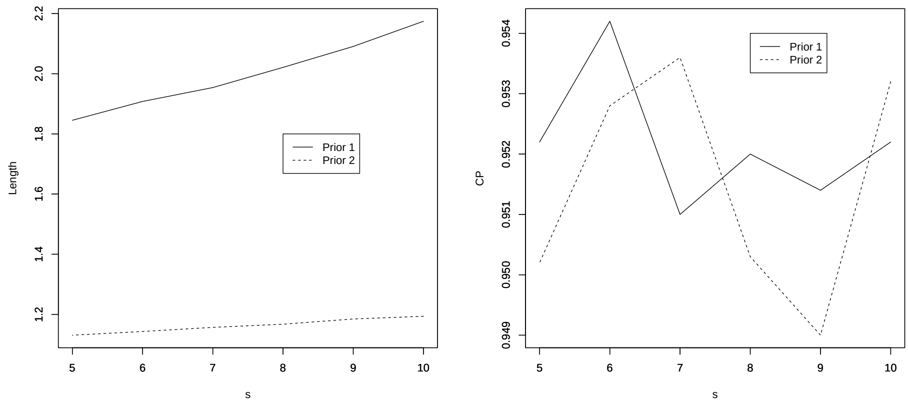

Figure 6 – Coverage lengths and coverage probabilities versus s: coverage length (top left) and coverage probability (right).

Finally, we repeated the simulations for Figures 3 and 4 by contaminating the simulated samples. Instead of simulating a sample of size n from the standard normal distribution, we simulated a sample of size $(n-1)$ from the standard normal distribution and a sample size 1 from the Student's t distribution with one degree of freedom. We then computed the biases, MSEs, confidence lengths and coverage probabilities as before. Plots of them versus n showed a similar pattern to Figures 3 and 4: excluding the BLUEs, the MLEs had the largest biases, the Bayes estimators based on the squared error loss had the second largest biases and the Bayes estimators based on the Linex loss had the smallest biases; the BLUEs had the largest MSEs, the MLEs had the second largest MSEs, the Bayes estimators based on the squared error loss had the third largest MSEs and the Bayes estimators based on the Linex loss had the smallest MSEs; the use of prior 2 led to smaller biases and smaller MSEs; confidence lengths appeared largest for the BLUE CIs and smallest for the HPD intervals based on prior 2; coverage probabilities appeared furthest from the nominal level for the BLUE CIs and closest for the HPD intervals based on prior 2; and so on. But the magnitude of the biases, MSEs and confidence lengths were larger when compared to Figures 3 and 4. Also the coverage probabilities were further away from the nominal level when compared to Figure 4.

7. Conclusions

We have considered Bayesian estimation and prediction for normal distribution based on upper record values. Under the squared error and Linex loss functions, we have derived Bayes estimators of the location and scale parameters using Monte Carlo simulations. It is observed that the Bayes estimators have clear advantages over the MLEs and BLUEs and that the Linex loss is superior to the squared error loss. We also used Monte Carlo simulations to compute Bayesian predictors of future records. Once again the Linex loss gave better predictions than the squared error loss.

Acknowledgements

The authors would like to thank the Editor and the referee for careful reading and for their comments which greatly improved the paper.

References

J. Ahmadi, M. Doostparast, A. Parsian (2005). *Estimation and prediction in a two-parameter exponential distribution based on k-record values under LINEX loss function*. Communications in Statistics-Theory and Methods, 34, pp. 795–805.

M. Ahsanullah (1995). *Record Statistics*. Nova Science Publishers, Commack, New York.

B. C. Arnold, N. Balakrishnan, H. N. Nagaraja (1998). *Records*. John Wiley and Sons, New York.

A. Asgharzadeh, A. Fallah (2011). *Estimation and prediction for exponentiated family of distributions based on records*. Communications in Statistics-Theory and Methods, 40, pp. 68–83.

A. Asgharzadeh, R. Valiollahi, D. Kundu (2015). *Prediction for future failures in Weibull distribution under hybrid censoring*. Journal of Statistical Computation and Simulation, 85, pp. 824–838.

N. Balakrishnan, P. S. Chan (1998). *On the normal record values and associated inference*. Statistics and Probability Letters, 39, pp. 73–80.

N. Balakrishnan, A. C. Cohen (1991). *Order Statistics and Inference: Estimation Methods*. Academic Press, San Diego.

A. P. Basu, N. Ebrahimi (1991). *Bayesian approach to life testing and reliability estimation using a symmetric loss function*. Journal of Statistical Planning and Inference, 29, pp. 21–31.

M. Chacko, M. Mary (2013). *Estimation and prediction based on k-record values from normal distribution*. Statistica, 73, pp. 505–516.

K. N. CHANDLER (1952). *The distribution and frequency of record values*. Journal of the Royal Statistical Society: Series B, 14, pp. 220–228.

M. H. CHEN, Q. M. SHAO (1999). *Monte Carlo estimation of Bayesian credible and HPD intervals*. Journal of Computational and Graphical Statistics, 8, pp. 69–92.

R. P. FEYNMAN (1987). *Mr. Feynman goes to Washington*. Engineering and Science, 51, no. 1, pp. 6–22.

D. KUNDU, H. HOWLADER (2010). *Bayesian inference and prediction of the inverse Weibull distribution for Type-II censored data*. Computational Statistics and Data Analysis, 54, pp. 1547–1558.

D. KUNDU, B. PRADHAN (2009). *Estimating the parameters of the generalized exponential distribution in presence of hybrid censoring*. Communications in Statistics-Theory and Methods, 38, pp. 2030–2041.

D. KUNDU, M. Z. RAQAB (2012). *Bayesian inference and prediction of order statistics for a Type-II censored Weibull distribution*. Journal of Statistical Planning and Inference, 142, pp. 41–47.

V. B. NEVZOROV (2000). *Records: Mathematical Theory*. American Mathematical Society, Providence, Rhode Island.

R DEVELOPMENT CORE TEAM (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

M. Z. RAQAB, J. AHMADI, M. DOOSTPARAST (2007). *Statistical inference based on record data from Pareto model*. Statistics, 41, pp. 105–118.

M. Z. RAQAB, M. T. MADI (2002). *Bayesian prediction of the total time on test using doubly censored Rayleigh data*. Journal of Statistical Computation and Simulation, 72, pp. 781–789.

C. REN, D. SUN, D. K. DEY (2006). *Bayesian and frequentist estimations and prediction for exponential distributions*. Journal of Statistical Planning and Inference, 13, pp. 2873–2897.

N. K. SAJEEVKUMAR, M. R. IRSHAD (2014). *Estimation of the location parameter of distributions with known coefficient of variation by record values*. Statistica, 74, no. 3, pp. 335–349.

A. A. SOLIMAN, F. M. AL-ABOUD (2008). *Bayesian inference using record values from Rayleigh model with application*. European Journal of Operational Research, 185, pp. 659–672.

H. R. VARRIAN (1975). *A Bayesian approach to real estate assessment*. In S. E. FEINBERGE, A. ZELLNER (eds.), *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage*, North Holland, Amsterdam, pp. 195–208.

S. J. WU, D. H. CHEN, S. T. CHEN (2006). *Bayesian inference for Rayleigh distribution under progressive censored sample*. Applied Stochastic Models in Business and Industry, 22, pp. 269–279.

A. ZELLNER (1986). *Bayesian estimation and prediction using asymmetric loss function*. Journal of the American Statistical Association, 81, pp. 446–451.

SUMMARY

Based on record data, the estimation and prediction problems for normal distribution have been investigated by several authors in the frequentist set up. However, these problems have not been discussed in the literature in the Bayesian context. The aim of this paper is to consider a Bayesian analysis in the context of record data from a normal distribution. We obtain Bayes estimators based on squared error and linear-exponential (Linex) loss functions. It is observed that the Bayes estimators can not be obtained in closed forms. We propose using an importance sampling method to obtain Bayes estimators. Further, the importance sampling method is also used to compute Bayesian predictors of future records. Finally, a real data analysis is presented for illustrative purposes and Monte Carlo simulations are performed to compare the performances of the proposed methods. It is shown that Bayes estimators and predictors are superior than frequentist estimators and predictors.

Keywords: Bayesian prediction; Best linear unbiased estimators; Maximum likelihood estimators; Record data.