SORT 40 (1) January-June 2016, 89-112

Point and interval estimation for the logistic distribution based on record data

A. Asgharzadeh[1,*], R. Valiollahi[2] and M. Abdi[3]

Abstract

In this paper, based on record data from the two-parameter logistic distribution, the maximum likelihood and Bayes estimators for the two unknown parameters are derived. The maximum likelihood estimators and Bayes estimators can not be obtained in explicit forms. We present a simple method of deriving explicit maximum likelihood estimators by approximating the likelihood function. Also, an approximation based on the Gibbs sampling procedure is used to obtain the Bayes estimators. Asymptotic confidence intervals, bootstrap confidence intervals and credible intervals are also proposed. Monte Carlo simulations are performed to compare the performances of the different proposed methods. Finally, one real data set has been analysed for illustrative purposes.

MSC: 62G30, 62F10, 62F15, 62E15.

Keywords: Logistic distribution, record data, maximum likelihood estimator, Bayes estimator, Gibbs sampling.

1. Background and statistical context

Let $\{Y_i, i \geq 1\}$ be a sequence of independent and identically distributed (iid) random variables with cumulative distribution function (cdf) $G(y;\theta)$ and probability density function (pdf) $g(y;\theta)$, where θ is a vector of parameters. An observation Y_j is called an upper record value if $Y_j > Y_i$ for all $i = 1, 2, \ldots, j-1$. An analogous definition can be given for lower record values. Generally, if $\{U(n),\, n \geq 1\}$ is defined by

$$U(1) = 1, \qquad U(n) = \min\{j : j > U(n-1),\, Y_j > Y_{U(n-1)}\},$$

* Corresponding author: a.asgharzadeh@umz.ac.ir

[1] Department of Statistics, Faculty of Mathematical Sciences, University of Mazandaran, Babolsar, Iran.

[2] Department of Mathematics, Statistics and Computer Science, Semnan University, Semnan, Iran.

[3] Department of Mathematics and Soft Computing, Higher Education Complex of Bam, Bam, Iran.

Received: February 2014

Accepted: April 2016

for $n \geq 2$, then the sequence $\{Y_{U(n)},\ n \geq 1\}$ provides a sequence of upper record statistics. The sequence $\{U(n),\ n \geq 1\}$ represents the record times.

Suppose we observe the first m upper record values $Y_{U(1)} = y_1, Y_{U(2)} = y_2, \cdots, Y_{U(m)} = y_m$ from the cdf $G(y;\theta)$ and pdf $g(y;\theta)$. Then, the joint pdf of the first m upper record values is given (see Ahsanullah, 1995) by

$$h(\mathbf{y};\theta) = g(y_m;\theta) \prod_{i=1}^{m-1} \frac{g(y_i;\theta)}{1-G(y_i;\theta)}, \quad -\infty < y_1 < y_2 < \ldots < y_m < \infty, \tag{1.1}$$

where $\mathbf{y} = (y_1,\ldots,y_m)$. The marginal pdf of the nth record $Y_{U(n)}$ is

$$h_n(y;\theta) = \frac{[-\ln(1-G(y;\theta))]^{n-1}}{(n-1)!} g(y;\theta).$$

The definition of record statistics was formulated by Chandler (1952). These statistics are of interest and important in many real life problems involving weather, economics, sports data and life testing studies. In reliability and life testing experiments, many products fail under stress. For example, an electronic component ceases to function in an environment of too high temperature, a wooden beam breaks when sufficient perpendicular force is applied to it, and a battery dies under the stress of time. Hence, in such experiments, measurements may be made sequentially and only the record values (lower or upper) are observed. For more details and applications of record values, one may refer to Arnold et al. (1998) and Nevzorov (2001).

The logistic distribution has been used for growth models in the biological sciences, and is used in a certain type of regression known as the logistic regression. It has many applications in technological problems including reliability, studies on income, graduation of mortality statistics, modeling agriculture production data, and analysis of categorical data. The shape of the logistic distribution is very similar to that of the normal distribution, but it is more peaked in the center and has heavier tails than the normal distribution. Because of the similarity of the two distributions, the logistic model has often been selected as a substitute for the normal model. For more details and other applications, see Balakrishnan (1992) and Johnson et al. (1995).

Although extensive work has been done on inferential procedures for logistic distribution based on complete and censored data, but not much attention has been paid on inference based on record data. In this article, we consider the point and interval estimation of the unknown parameters of the logistic distribution based on record data. We first consider the maximum likelihood estimators (MLEs) of the unknown parameters. It is observed the MLEs can not be obtained in explicit forms. We present a simple method of deriving explicit MLEs by approximating the likelihood function. We further consider the Bayes estimators of the unknown parameters and it is observed the Bayes estimators and the corresponding credible intervals can not be obtained in explicit forms. We use an approximation based on the Gibbs sampling procedure to compute the Bayes estimators and the corresponding credible intervals.

The rest of the paper is organized as follows. In Section 2, we discuss the MLEs of the unknown parameters of the logistic distribution. In Section 3, we provide the approximate maximum likelihood estimators (AMLEs). Bayes estimators and the corresponding credible intervals are provided in Section 4. The Fisher information and different confidence confidence intervals are presented in Section 5. Finally, in Section 4, one numerical example and a Monte Carlo simulation study are given to illustrate the results.

2. Maximum likelihood estimation

Let the failure time distribution be a logistic distribution with probability density function (pdf)

$$g(y;\mu,\sigma) = \frac{e^{-(y-\mu)/\sigma}}{\sigma(1+e^{-(y-\mu)/\sigma})^2}\ , \quad -\infty < y < \infty, \quad \mu \in R, \quad \sigma > 0, \qquad (2.1)$$

and cumulative distribution function (cdf)

$$G(y;\mu,\sigma) = \frac{1}{1+e^{-(y-\mu)/\sigma}}\ , \quad -\infty < y < \infty, \quad \mu \in R, \quad \sigma > 0. \qquad (2.2)$$

Consider the random variable $X = (Y-\mu)/\sigma$. Then, X has the standard logistic distribution with pdf and cdf as

$$f(x) = \frac{e^{-x}}{(1+e^{-x})^2}\ , \quad -\infty < x < \infty, \qquad (2.3)$$

and

$$F(x) = \frac{1}{1+e^{-x}}\ , \quad -\infty < x < \infty, \qquad (2.4)$$

respectively. Note that $g(y;\mu,\sigma) = \frac{1}{\sigma}f((y-\mu)/\sigma)$ and $G(y;\mu,\sigma) = F((y-\mu)/\sigma)$. It should also be noted that $f(x)$ and $F(x)$ satisfy the following relationships:

$$f(x) = F(x)[1-F(x)], \qquad f'(x) = f(x)[1-2F(x)]. \qquad (2.5)$$

Suppose we observe the first m upper record values $Y_{U(1)} = y_1, Y_{U(2)} = y_2, \cdots, Y_{U(m)} = y_m$ from the logistic distribution with pdf (2.1) and cdf (2.2). The likelihood function is given by

$$L(\mu,\sigma) = g(y_m,\mu,\sigma)\prod_{i=1}^{m-1}\frac{g(y_i;\mu,\sigma)}{1-G(y_i;\mu,\sigma)}. \tag{2.6}$$

By using Eqs. (2.3), (2.4) and (2.5), the likelihood function may be rewritten as

$$L(\mu,\sigma) = \sigma^{-m} f(x_m)\prod_{i=1}^{m-1} F(x_i)\ , \tag{2.7}$$

where $x_i = (y_i-\mu)/\sigma$. Subsequently, the log-likelihood function is

$$\ln L(\mu,\sigma) = -m\ln\sigma + \ln f(x_m) + \sum_{i=1}^{m-1}\ln F(x_i). \tag{2.8}$$

Again, by using Eq. (2.5), we derive the likelihood equations for μ and σ from (2.8), as

$$\frac{\partial \ln L(\mu,\sigma)}{\partial\mu} = -\frac{1}{\sigma}\left[m - F(x_m) - \sum_{i=1}^{m} F(x_i)\right] = 0, \tag{2.9}$$

$$\frac{\partial \ln L(\mu,\sigma)}{\partial\sigma} = -\frac{1}{\sigma}\left[m + \sum_{i=1}^{m} x_i - x_m F(x_m) - \sum_{i=1}^{m} x_i F(x_i)\right] = 0. \tag{2.10}$$

The MLES $\widehat{\mu}$ and $\widehat{\sigma}$, respectively of μ and σ, are solution of the system of Eqs. (2.9) and (2.10). They can not be obtained in closed forms and so some iterative methods such as Newton's method are required to compute these estimators.

3. Approximate maximum likelihood estimation

It is observed that the likelihood equations (2.9) and (2.10) do not yield explicit estimators for the MLEs, because of the presence of the term $F(x_i)$, $i = 1,\ldots,m$, and they have to be solved by some iterative methods. However, as mentioned by Tiku and Akkaya (2004), solving the likelihood equations by iterative methods can be problematic for reasons of (i) multiple roots, (ii) nonconvergence of iterations, or (iii) convergence to wrong values. Moreover, these methods are usually very sensitive to their initiate values. Here, we present a simple method to derive approximate MLEs for μ and σ by linearizing the term $F(x_i)$ using Taylor series expansion. Approximate solutions for MLEs have been discussed in the book by Tiku and Akkaya (2004) for several specific distributions.

Balakrishnan and Aggarwala (2000), Balakrishnan and Kannan (2000), Balakrishnan and Asgharzadeh (2005), Agharzadeh (2006), Raqab et al. (2010) and Asgharzadeh et al. (2011) used approximate solutions for the MLEs, when the data are progressively censored.

We approximate the term $F(x_i)$ by expanding it in a Taylor series around $E(X_i) = \delta_i$. From Arnold et al. (1998), it is known that

$$F(X_i) \stackrel{d}{=} U_i,$$

where U_i is the i-th record statistic from the uniform $U(0,1)$ distribution. We then have

$$X_i \stackrel{d}{=} F^{-1}(U_i),$$

and hence

$$\delta_i = E(X_i) \approx F^{-1}(E(U_i)).$$

From Arnold et al. (1998), it is known that

$$E(U_i) = 1 - \left(\frac{1}{2}\right)^{i+1}, \ \ i = 1, \ldots, m.$$

Since, for the standard logistic distribution, we have

$$F^{-1}(u) = \ln\left(\frac{u}{1-u}\right),$$

we can approximate δ_i by $F^{-1}[1-(\frac{1}{2})^{i+1}] = \ln(2^{i+1}-1)$.

Now, by expanding the function $F(x_i)$ around the point δ_i and keeping only the first two terms, we have the following approximation

$$\begin{aligned} F(x_i) &\simeq F(\delta_i) + (x_i - \delta_i) f(\delta_i) \\ &= \alpha_i + \beta_i x_i, \end{aligned} \tag{3.1}$$

where

$$\alpha_i = F(\delta_i) - \delta_i f(\delta_i),$$

and

$$\beta_i = f(\delta_i) \geq 0,$$

for $i = 1, \cdots, m$.

Using the expression in (3.1), we approximate the likelihood equations in (2.9) and (2.10) by

$$\frac{\partial \ln L^*(\mu,\sigma)}{\partial \mu} = -\frac{1}{\sigma}\left[m - (\alpha_m + \beta_m x_m) - \sum_{i=1}^{m}(\alpha_i + \beta_i x_i)\right] = 0, \tag{3.2}$$

$$\frac{\partial \ln L^*(\mu,\sigma)}{\partial \sigma} = -\frac{1}{\sigma}\left[m + \sum_{i=1}^{m} x_i - x_m(\alpha_m + \beta_m x_m) - \sum_{i=1}^{m} x_i(\alpha_i + \beta_i x_i)\right] = 0, \tag{3.3}$$

which can be rewritten as

$$\left[m - \alpha_m - \sum_{i=1}^{m}\alpha_i\right] - \frac{1}{\sigma}\left[\beta_m y_m + \sum_{i=1}^{m}\beta_i y_i\right] + \frac{1}{\sigma}\left[\beta_m + \sum_{i=1}^{m}\beta_i\right]\mu = 0, \tag{3.4}$$

$$m + \frac{1}{\sigma}\left[\left(\sum_{i=1}^{m} y_i - \alpha_m y_m - \sum_{i=1}^{m}\alpha_i y_i\right) + \frac{(\beta_m y_m + \sum_{i=1}^{m}\beta_i y_i)(\alpha_m + \sum_{i=1}^{m}\alpha_i - m)}{\beta_m + \sum_{i=1}^{m}\beta_i}\right]$$

$$+ \frac{1}{\sigma^2}\left[-\left(\beta_m y_m^2 + \sum_{i=1}^{m}\beta_i y_i^2\right) + \frac{(\beta_m y_m + \sum_{i=1}^{m}\beta_i y_i)^2}{\beta_m + \sum_{i=1}^{m}\beta_i}\right] = 0, \tag{3.5}$$

respectively. By solving the quadratic equation in (3.5) for σ, we obtain the approximate MLE of σ as

$$\widetilde{\sigma} = \frac{-A + \sqrt{A^2 - 4mB}}{2m}, \tag{3.6}$$

where

$$A = \left(\sum_{i=1}^{m} y_i - \alpha_m y_m - \sum_{i=1}^{m}\alpha_i y_i\right) + \frac{(\beta_m y_m + \sum_{i=1}^{m}\beta_i y_i)(\alpha_m + \sum_{i=1}^{m}\alpha_i - m)}{\beta_m + \sum_{i=1}^{m}\beta_i}, \tag{3.7}$$

$$B = -\left(\beta_m y_m^2 + \sum_{i=1}^{m}\beta_i y_i^2\right) + \frac{(\beta_m y_m + \sum_{i=1}^{m}\beta_i y_i)^2}{\beta_m + \sum_{i=1}^{m}\beta_i}. \tag{3.8}$$

Now, by using (3.4), we obtain the approximate MLE of μ as

$$\tilde{\mu} = C + D\tilde{\sigma}, \tag{3.9}$$

where

$$C = \frac{\beta_m y_m + \sum_{i=1}^m \beta_i y_i}{\beta_m + \sum_{i=1}^m \beta_i}, \qquad D = \frac{\alpha_m + \sum_{i=1}^m \alpha_i - m}{\beta_m + \sum_{i=1}^m \beta_i}. \tag{3.10}$$

Note that Eq. (3.5) has two roots but since $B \leq 0$, only one root in (3.6) is admissible. The proof of $B \leq 0$ is given in Appendix A.

Note that, the AMLE method has an advantage over the MLE method as the former provides explicit estimators. The AMLEs in (3.6) and (3.9) can be used as good starting values for the iterative solution of the likelihood equations (2.9) and (2.10) to obtain the MLEs. As mentioned in Tiku and Akkaya (2004), the AMLEs of the location an scale parameters μ and σ are asymptotically equivalent to the corresponding MLEs for any location-scale distribution. This is due to the asymptotic equivalence of the approximate likelihood and the likelihood equations. The approximate MLEs have all desirable asymptotic properties of MLEs. They are asymptotically unbiased and efficient. They have also robustness properties for all the three types distributions: skew, short-tailed symmetric and long-tailed symmetric distributions. For more details, see Tiku and Akkaya (2004).

4. Bayesian estimation and credible intervals

In this section, the Bayes estimators of the unknown parameters μ and σ are derived under the squared error loss function. Further, the corresponding credible intervals of μ and σ are also obtained. It is assumed that joint prior distribution for μ and σ is in the form

$$\pi(\mu, \sigma) = \pi_1(\mu|\sigma)\pi_2(\sigma),$$

where σ has an inverse gamma prior $IG(a, b)$, with the pdf

$$\pi_2(\sigma) \propto e^{-\frac{b}{\sigma}} \sigma^{-(a+1)}, \quad \sigma > 0, \; a, b > 0,$$

and μ given σ has the logistic prior with parameters μ_0 and σ

$$\pi_1(\mu|\sigma) = \frac{e^{-\frac{\mu-\mu_0}{\sigma}}}{\sigma \left[1 + e^{-\frac{\mu-\mu_0}{\sigma}}\right]^2},$$

This joint prior is suitable for deriving the posterior distribution in a location and scale parameter estimation.

From (2.6), for the logistic distribution, the likelihood function of μ and σ for the given record sample $\mathbf{y} = (y_1, y_2, \ldots, y_m)$ is given by

$$L(\mu,\sigma|\mathbf{y}) = e^{-\frac{y_m-\mu}{\sigma}} \sigma^{-m} \frac{\prod_{i=1}^{m}(1+e^{-\frac{y_i-\mu}{\sigma}})^{-1}}{1+e^{-\frac{y_m-\mu}{\sigma}}}. \tag{4.1}$$

By combining the likelihood function in (4.1) and the joint prior distribution, we obtain the joint posterior distribution of μ and σ as

$$\pi(\mu,\sigma|\mathbf{y}) \propto e^{-\frac{b+y_m-\mu_0}{\sigma}} \sigma^{-(m+a+2)} \frac{\prod_{i=1}^{m}(1+e^{-\frac{y_i-\mu}{\sigma}})^{-1}}{\left[1+e^{-\frac{y_m-\mu}{\sigma}}\right]\left[1+e^{-\frac{\mu-\mu_0}{\sigma}}\right]^2}. \tag{4.2}$$

Therefore, the Bayes estimators of μ and σ are respectively obtained as

$$\widehat{\mu}_{BS} = E(\mu|\mathbf{y}) = k \int_{-\infty}^{\infty}\int_{0}^{\infty} \mu\, e^{-\frac{b+y_m-\mu_0}{\sigma}} \sigma^{-(m+a+2)} \frac{\prod_{i=1}^{m}(1+e^{-\frac{y_i-\mu}{\sigma}})^{-1}}{\left[1+e^{-\frac{y_m-\mu}{\sigma}}\right]\left[1+e^{-\frac{\mu-\mu_0}{\sigma}}\right]^2} d\sigma d\mu,$$

and

$$\widehat{\sigma}_{BS} = E(\sigma|\mathbf{y}) = k \int_{-\infty}^{\infty}\int_{0}^{\infty} e^{-\frac{b+y_m-\mu_0}{\sigma}} \sigma^{-(m+a+1)} \frac{\prod_{i=1}^{m}(1+e^{-\frac{y_i-\mu}{\sigma}})^{-1}}{\left[1+e^{-\frac{y_m-\mu}{\sigma}}\right]\left[1+e^{-\frac{\mu-\mu_0}{\sigma}}\right]^2} d\sigma d\mu,$$

where k is the normalizing constant.

It is seen that the Bayes estimators can not be obtained in closed forms. In what follows, similarly as in Kundu (2007, 2008), we provide the approximate Bayes estimators using a rejection-sampling within the Gibbs sampling procedure. Note that the joint posterior distribution of μ and σ given $\mathbf{y}$ in (4.2), can be written as

$$\pi(\mu,\sigma|\mathbf{y}) \propto g_1(\sigma|\mathbf{y}) g_2(\mu|\sigma,\mathbf{y}). \tag{4.3}$$

Here $g_1(\sigma|\mathbf{y})$ is an inverse gamma density function with the shape and scale parameters as $m+a+1$ and $b+y_m-\mu_0$, respectively, and $g_2(\mu|\sigma,\mathbf{y})$ is a proper density function given by

$$g_2(\mu|\sigma,\mathbf{y}) \propto \frac{\prod_{i=1}^{m}(1+e^{-\frac{y_i-\mu}{\sigma}})^{-1}}{\left[1+e^{-\frac{y_m-\mu}{\sigma}}\right]\left[1+e^{-\frac{\mu-\mu_0}{\sigma}}\right]^2}. \tag{4.4}$$

To obtain the Bayes estimates using the Gibbs sampling procedure, we need the following result.

Theorem 1. *The conditional distribution of μ given σ and $\mathbf{y}$, $g_2(\mu|\sigma,\mathbf{y})$, is log-concave.*

Proof: See the Appendix B.

Thus, the samples of μ can be generated from (4.4) using the method proposed by Devroye (1984). Now, using Theorem 1, and adopting the method of Devroye (1984), we can generate the samples (μ,σ) from the posterior density function (4.3), using the Gibbs sampling procedure as follows:

1. Generate σ_1 from $g_1(.|\mathbf{y})$.
2. Generate μ_1 from $g_2(.|\sigma_1,\mathbf{y})$ using the method developed by Devroye (1984).
3. Repeat steps 1 and 2 N times and obtain $(\mu_1,\sigma_1),\cdots,(\mu_N,\sigma_N)$.

Note that in step 2, we use the Devroye algorithm as follows:

i) Compute $c=g_2(m|\sigma,\mathbf{y})$. (m is the mode of $g_2(.|\sigma,\mathbf{y})$).

ii) Generate U uniform on [0,2], and V uniform on [0,1].

iii) If $U\leq 1$ then $\mu=U$ and $T=V$, else $\mu=1-\ln(U-1)$ and $T=V(U-1)$.

iv) Let $\mu=m+\frac{\mu}{c}$.

v) If $T\leq\frac{g_2(\mu|\sigma,\mathbf{y})}{c}$, then μ is a sample from $g_2(.|\sigma,\mathbf{y})$, else go to Step (ii).

Now, the Bayesian estimators of μ and σ under the squared error loss function are obtained as

$$\widehat{\mu}_{BS}=\frac{\sum_{j=1}^{N}\mu_j}{N},\qquad \widehat{\sigma}_{BS}=\frac{\sum_{j=1}^{N}\sigma_j}{N}. \tag{4.5}$$

Now we obtain the credible intervals of μ and σ using the idea of Chen and Shao (1999). To compute the credible intervals of μ and σ, we generate $\mu_1,\ldots,\mu_N$ and $\sigma_1,\ldots,\sigma_N$ as described above. We then order $\mu_1,\ldots,\mu_N$ and $\sigma_1,\ldots,\sigma_N$ as $\mu_{(1)},\ldots,\mu_{(N)}$ and $\sigma_{(1)},\ldots,\sigma_{(N)}$. Then, the $100(1-\gamma)\%$ credible intervals μ and σ can be constructed as

$$\left(\mu_{(\frac{\gamma}{2}N)},\ \mu_{((1-\frac{\gamma}{2})N)}\right),\qquad \left(\sigma_{(\frac{\gamma}{2}N)},\ \sigma_{((1-\frac{\gamma}{2})N)}\right). \tag{4.6}$$

5. Fisher information and different confidence intervals

In this section, we derive the Fisher information matrix based on the likelihood as well as the approximate likelihood functions. Using the Fisher information matrix and based on the asymptotic distribution of MLEs, we can obtain the asymptotic confidence intervals of μ and σ. We further, propose two confidence intervals based on the bootstrap method.

5.1. Fisher information

From (2.9) and (2.10), the expected Fisher information matrix of $\boldsymbol{\theta} = (\mu, \sigma)$ is

$$I(\boldsymbol{\theta}) = -\begin{pmatrix} E(\frac{\partial^2 \ln L(\mu,\sigma)}{\partial \mu^2}) & E(\frac{\partial^2 \ln L(\mu,\sigma)}{\partial \mu \partial \sigma}) \\ E(\frac{\partial^2 \ln L(\mu,\sigma)}{\partial \sigma \partial \mu}) & E(\frac{\partial^2 \ln L(\mu,\sigma)}{\partial \sigma^2}) \end{pmatrix} = \begin{pmatrix} I_{11} & I_{12} \\ I_{12} & I_{22} \end{pmatrix}, \tag{5.1}$$

where

$$\begin{aligned} I_{11} &= \frac{1}{\sigma^2}\left[E[f(X_m)] + \sum_{i=1}^{m} E[f(X_i)]\right], \\ I_{12} &= \frac{1}{\sigma^2}\left[m - E[F(X_m)] - \sum_{i=1}^{m} E[F(X_i)] - E[X_m f(X_m)] - \sum_{i=1}^{m} E[X_i f(X_i)]\right], \\ I_{22} &= -\frac{1}{\sigma^2}\left[m + 2\sum_{i=1}^{m} E[X_i(1 - F(X_i))] - 2E[X_m F(X_m)]\right. \\ &\qquad \left. - E[X_m^2 f(X_m)] - \sum_{i=1}^{m} E[X_i^2 f(X_i)]\right]. \end{aligned}$$

Similarly, the expected approximate Fisher information matrix of $\boldsymbol{\theta} = (\mu, \sigma)$ is obtained to be

$$I^*(\boldsymbol{\theta}) = -\begin{pmatrix} E(\frac{\partial^2 \ln L^*(\mu,\sigma)}{\partial \mu^2}) & E(\frac{\partial^2 \ln L^*(\mu,\sigma)}{\partial \mu \partial \sigma}) \\ E(\frac{\partial^2 \ln L^*(\mu,\sigma)}{\partial \sigma \partial \mu}) & E(\frac{\partial^2 \ln L^*(\mu,\sigma)}{\partial \sigma^2}) \end{pmatrix} = \begin{pmatrix} I_{11}^* & I_{12}^* \\ I_{12}^* & I_{22}^* \end{pmatrix}, \tag{5.2}$$

where

$$\begin{aligned} I_{11}^* &= \frac{1}{\sigma^2}\left[\beta_m + \sum_{i=1}^{m} \beta_i\right], \\ I_{12}^* &= -\frac{1}{\sigma^2}\left[m - \alpha_m - \sum_{i=1}^{m} \alpha_i - 2\beta_m E[X_m] - 2\sum_{i=1}^{m} \beta_i E[X_i]\right], \end{aligned}$$

$$I_{22}^{*} = -\frac{1}{\sigma^2}\left[m + 2\sum_{i=1}^{m}(1-\alpha_i)E[X_i] - 2\alpha_m E[X_m] - 3\beta_m E[X_m^2] - 3\sum_{i=1}^{m}\beta_i E[X_i^2]\right].$$

From Ahsanullah (1995), since

$$E[X_1] = 0, \quad E[X_i] = \sum_{l=2}^{i}\zeta(l), \quad i \geq 2,$$

and

$$E[X_i^2] = 2i\sum_{l=2}^{i+1}\zeta(l) - i(i+1) + \sum_{l=2}^{\infty}\frac{B_l}{(l+1)^i},$$

where $\zeta(.)$ is Riemann zeta function $\zeta(n) = \sum_{k=1}^{\infty} k^{-n}$ and for $n \geq 2$

$$B_n = \frac{1}{n}(1 + \frac{1}{2} + \cdots + \frac{1}{n-1}),$$

we can derive the elements of Fisher information matrix in (5.2). Now, to derive the elements of Fisher information matrix in (5.1), we need to calculate the expectations $E[f(X_i)]$, $E[F(X_i)]$, $E[X_i(1-F(X_i))]$, $E[X_i f(X_i)]$, $E[X_i F(X_i)]$ and $E[X_i^2 f(X_i)]$. We use the following lemma to compute these expectations.

Lemma 1. *Let $X_1 < X_2 < \cdots < X_m$ is the first m upper record values from the standard logistics distribution with pdf (2.3). Then we have*

$$E[f(X_i)] = \frac{1}{2^i} - \frac{1}{3^i}, \tag{5.3}$$

$$E[F(X_i)] = 1 - \frac{1}{2^i}, \tag{5.4}$$

$$E[X_i f(X_i)] = \sum_{l=1}^{\infty}\left[\frac{1}{l(l+3)^i} - \frac{1}{l(l+2)^i}\right] + i\left[\frac{1}{2^i} - \frac{1}{3^i}\right], \tag{5.5}$$

$$E\left[X_i(1-F(X_i))\right] = \frac{i}{2^{i+1}} - \sum_{l=1}^{\infty}\frac{1}{l(l+2)^i}, \tag{5.6}$$

and

$$\begin{aligned}
E\left[X_i^2 f(X_i)\right] &= \sum_{l=1}^{\infty}\left[\frac{1}{l^2(2l+2)^i}-\frac{1}{l^2(2l+3)^i}\right] \\
&+2\sum\sum_{1\le l<k<\infty}\left[\frac{1}{lk(l+k+2)^i}-\frac{1}{lk(l+k+3)^i}\right] \\
&+2i\sum_{l=1}^{\infty}\left[\frac{1}{l(l+3)^{i+1}}-\frac{1}{l(l+2)^{i+1}}\right] \\
&+i(i+1)\left[\frac{1}{2^{i+2}}-\frac{1}{3^{i+2}}\right]. \qquad (5.7)
\end{aligned}$$

Proof. See the Appendix C.

Moreover, $E[X_iF(X_i)]$ can be obtained from the expression

$$E[X_iF(X_i)] = E[X_i] - E[X_i(1-F(X_i))].$$

It should be mentioned here that the loss of information due to using record data instead of the complete logistic data can be discussed by comparing the Fisher information contained in record data with that of the Fisher information contained in the complete data. Since $\boldsymbol{\theta} = (\mu,\sigma)$ is a vector parameter, the comparison is not a trivial issue. One method is that to compare the Fisher information matrices for the two data using their traces. Based on a given data, the trace of Fisher information matrix of $\boldsymbol{\theta} = (\mu,\sigma)$ is the sum of the Fisher information measures of μ, when σ is known, and σ, when μ is known. For the logistic distribution, the Fisher information matrix of $\boldsymbol{\theta} = (\mu,\sigma)$ based on the first m record observations can be obtained from (5.1). On the other hand, the Fisher information matrix based on the m complete logistic observations is (see Nadarajah (2004))

$$J(\boldsymbol{\theta}) = \begin{pmatrix} J_{11} & J_{12} \\ J_{12} & J_{22} \end{pmatrix},$$

where

$$\begin{aligned}
J_{11} &= \frac{m}{3\sigma^2}\left(\frac{\pi^2}{3}+1\right), \\
J_{12} &= J_{12} = -\frac{m}{\sigma^2}, \\
J_{22} &= \frac{m}{3\sigma^2}.
\end{aligned}$$

Table 1: *The trace of the Fisher information matrix based on complete and record observations for different values of m.*

	Complete observations	Record observations
$m=2$	3.526	3.149
$m=3$	5.289	4.502
$m=5$	8.816	6.916
$m=10$	17.633	12.131
$m=15$	26.450	19.175
$m=20$	35.265	27.917

We have computed the traces of the corresponding Fisher information matrices for both data and the results are reported in Table 1. From Table 1, as expected, we see that the Fisher information contained in the m complete observations is greater than that the Fisher information contained in the m record observations.

5.2. Different confidence intervals

Now, the variances of the MLEs $\widehat{\mu}$ and $\widehat{\sigma}$, can be approximated by inverting the Fisher information matrix in (5.1), *i.e*,

$$\begin{pmatrix} Var(\widehat{\mu}) & Cov(\widehat{\mu},\widehat{\sigma}) \\ Cov(\widehat{\mu},\widehat{\sigma}) & Var(\widehat{\sigma}) \end{pmatrix} \approx \begin{pmatrix} I_{11} & I_{12} \\ I_{12} & I_{22} \end{pmatrix}^{-1}. \tag{5.8}$$

The approximate asymptotic variance covariance matrices are valid only if asymptotic normality holds. For the asymptotic normality, the certain regularity conditions must be satisfied (see, for example, the conditions in Theorem 4.17 of Shao (2003)). These conditions mainly relate to differentiability of the density and the ability to interchange differentiation and integration. In most reasonable problems, the regularity conditions are often satisfied. Since the logistic distribution satisfies all the the regularity conditions (see Shao (2005), Pages 198-200), we can obtain the approximate $100(1-\gamma)\%$ confidence intervals of μ and σ using the asymptotic normality of MLEs as

$$\left(\widehat{\mu}-z_{1-\gamma/2}\sqrt{\widehat{Var(\widehat{\mu})}}\ ,\ \widehat{\mu}+z_{1-\gamma/2}\sqrt{\widehat{Var(\widehat{\mu})}}\right), \tag{5.9}$$

and

$$\left(\widehat{\sigma}-z_{1-\gamma/2}\sqrt{\widehat{Var(\widehat{\sigma})}}\ ,\ \widehat{\sigma}+z_{1-\gamma/2}\sqrt{\widehat{Var(\widehat{\sigma})}}\right). \tag{5.10}$$

Similarly, the approximate confidence intervals can be obtained based on the AMLEs also, by inverting the approximate Fisher information in (5.2).

Now, we present two confidence intervals based on the parametric bootstrap methods: (i) percentile bootstrap method (we call it Boot-p) based on the idea of Efron (1982), (ii) bootstrap-t method (we refer to it as Boot-t) based on the idea of Hall (1988). The algorithms for these two bootstrap procedures are briefly described as follows.

(i) Boot-p method:

1. Estimate μ and σ, say $\widehat{\mu}$ and $\widehat{\sigma}$, from sample based on the MLE procedure.
2. Generate a bootstrap sample $\{X_1^*,\cdots,X_m^*\}$, using $\widehat{\mu}$ and $\widehat{\sigma}$. Obtain the bootstrap estimates of μ and σ, say $\widehat{\mu}^*$ and $\widehat{\sigma}^*$ using the bootstrap sample.
3. Repeat Step 2 NBOOT times.
4. Order $\widehat{\mu}_1^*,\cdots,\widehat{\mu}_{NBOOT}^*$ as $\widehat{\mu}_{(1)}^*,\cdots,\widehat{\mu}_{(NBOOT)}^*$ and $\widehat{\sigma}_1^*,\cdots,\widehat{\sigma}_{NBOOT}^*$ as $\widehat{\sigma}_{(1)}^*,\cdots,$ $\widehat{\sigma}_{(NBOOT)}^*$. Then, the approximate $100(1-\gamma)\%$ confidence intervals for μ and σ become, respectively, as

$$\Big(\widehat{\mu}^*_{Boot-p}(\frac{\gamma}{2})\ ,\ \widehat{\mu}^*_{Boot-p}(1-\frac{\gamma}{2})\Big),\qquad \Big(\widehat{\sigma}^*_{Boot-p}(\frac{\gamma}{2})\ ,\ \widehat{\sigma}^*_{Boot-p}(1-\frac{\gamma}{2})\Big).\quad (5.11)$$

(ii) Boot-t method:

1. Estimate μ and σ, say $\widehat{\mu}$ and $\widehat{\sigma}$, from sample based on the MLE method.
2. Generate a bootstrap sample $\{X_1^*,\cdots,X_m^*\}$, using $\widehat{\mu}$ and $\widehat{\sigma}$ and obtain the bootstrap estimates of μ and σ, say $\widehat{\mu}^*$ and $\widehat{\sigma}^*$ using the bootstrap sample.
3. Determine

$$T_\mu^* = \frac{(\widehat{\mu}^*-\widehat{\mu})}{\sqrt{\widehat{Var}(\widehat{\mu}^*)}}\ ,\qquad T_\sigma^* = \frac{(\widehat{\sigma}^*-\widehat{\sigma})}{\sqrt{\widehat{Var}(\widehat{\sigma}^*)}},$$

   where $\widehat{Var}(\widehat{\mu}^*)$ and $\widehat{Var}(\widehat{\sigma}^*)$ are obtained using (5.8)
4. Repeat Steps 2 and 3 NBOOT times.
5. Define $\widehat{\mu}^*_{Boot-t}=\widehat{\mu}+\sqrt{\widehat{Var}(\widehat{\mu}^*)}T_\mu^*$ and $\widehat{\sigma}^*_{Boot-t}=\widehat{\sigma}+\sqrt{\widehat{Var}(\widehat{\sigma}^*)}T_\sigma^*$. Order $\widehat{\mu}_1^*,\cdots,$ $\widehat{\mu}_{NBOOT}^*$ as $\widehat{\mu}_{(1)}^*,\cdots,\widehat{\mu}_{(NBOOT)}^*$ and $\widehat{\sigma}_1^*,\cdots,\widehat{\sigma}_{NBOOT}^*$ as $\widehat{\sigma}_{(1)}^*,\cdots,\widehat{\sigma}_{(NBOOT)}^*$. Then, the approximate $100(1-\gamma)\%$ confidence intervals for μ and σ become respectively as

$$\Big(\widehat{\mu}^*_{Boot-t}(\frac{\gamma}{2})\ ,\ \widehat{\mu}^*_{Boot-t}(1-\frac{\gamma}{2})\Big),\qquad \Big(\widehat{\sigma}^*_{Boot-t}(\frac{\gamma}{2})\ ,\ \widehat{\sigma}^*_{Boot-t}(1-\frac{\gamma}{2})\Big).\quad (5.12)$$

6. Data analysis and simulation

In this section, we analyze a real data set to illustrate the estimation methods presented in the preceding sections. Further, a Monte Carlo simulation study is conducted to compare the performance of proposed estimators.

6.1. Data analysis

The following data are the total annual rainfall (in inches) during March recorded at Los Angeles Civic Center from 1973 to 2006 (see the website of Los Angeles Almanac: www.laalman-ac.com/weather/we08aa.htm).

2.70	3.78	4.83	1.81	1.89	8.02	5.85	4.79	4.10	3.54
8.37	0.28	1.29	5.27	0.95	0.26	0.81	0.17	5.92	7.12
2.74	1.86	6.98	2.16	0.00	4.06	1.24	2.82	1.17	0.32
4.31	1.17	2.14	2.87						

The Los Angeles rainfall data have been used earlier by some authors. See for example, Raqab (2006), Madi and Raqab (2007) and Raqab et al. (2010).

We analyzed the above rainfall data by using the logistic distribution with $\mu = 2.905$ and $\sigma = 1.367$. It is observed that the Kolmogorov-Smirnov (KS) distance and the corresponding p-value are respectively

$$\text{KS} = 0.1066, \qquad \text{and} \qquad \text{p-value} = 0.8120.$$

Hence the logistic model (2.1) fits quite well to the above data.

For the above data, we observe the following five upper record values

2.70	3.78	4.83	8.02	8.37

We shall use the above rainfall records to obtain the different estimators discussed in this paper. Here, we have $m = 5, A = -3.644, B = -1.436, C = 4.089$ and $D = -0.742$. From (3.6), we obtain the AMLE of σ as

$$\tilde{\sigma} = \frac{-A + \sqrt{A^2 - 4mB}}{2m} = 1.012.$$

Now, by using (3.9), the AMLE of μ becomes

$$\tilde{\mu} = C + D\tilde{\sigma} = 3.338.$$

The MLEs of μ and σ are then respectively as $\widehat{\mu} = 2.929$ and $\widehat{\sigma} = 0.998$. Note that the MLEs were obtained by solving the nonlinear equations (2.9) and (2.10) using the Maple package, in which the AMLEs were used as starting values for the iterations. To ensure that the solution $(\widehat{\mu} = 2.929, \widehat{\sigma} = 0.998)$ of the likelihood equations (2.9) and (2.10) is indeed a maximum, it must be shown that the matrix of second-order partial derivatives (Hessian matrix)

$$H = \begin{pmatrix} \frac{\partial^2 \ln L(\mu,\sigma)}{\partial \mu^2} & \frac{\partial^2 \ln L(\mu,\sigma)}{\partial \mu \partial \sigma} \\ \frac{\partial^2 \ln L(\mu,\sigma)}{\partial \sigma \partial \mu} & \frac{\partial^2 \ln L(\mu,\sigma)}{\partial \sigma^2} \end{pmatrix},$$

is a negative definite when $\mu = \widehat{\mu}$ and $\sigma = \widehat{\sigma}$. Based on the above rainfall records and for $\widehat{\mu} = 2.929$ and $\widehat{\sigma} = 0.998$, the Hessian matrix is

$$H = \begin{pmatrix} -0.5857 & 0.4156 \\ 0.4156 & -5.0194 \end{pmatrix},$$

which can be shown that is negative definite. Therefore, we have indeed found a maximum. On the other hand, we have also plotted the likelihood function of μ and σ for the given record data in Figure 1. From Figure 1, one can observe that the likelihood surface has curvature in both μ and σ directions. This leads to the interpretation that MLEs of μ and σ are exist and unique.

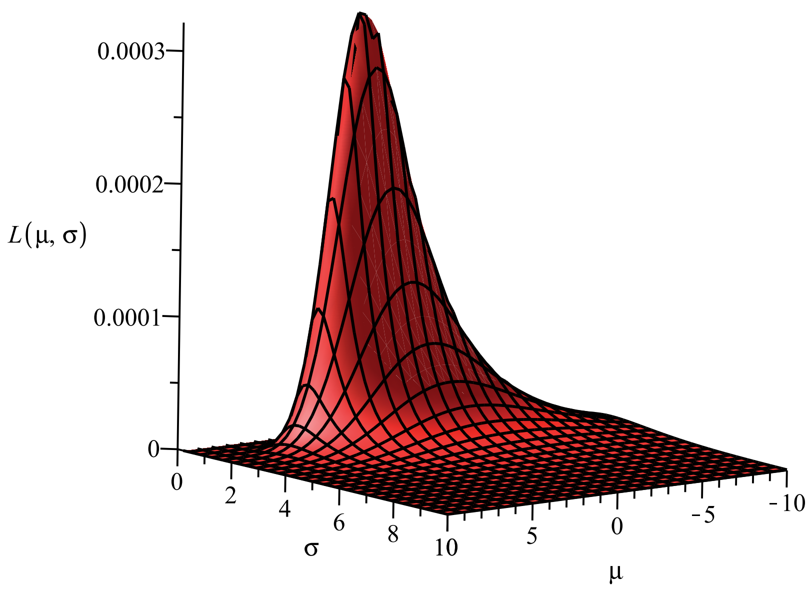

Figure 1: *Likelihood function of μ and σ.*

Table 2: *Point and interval estimators of μ and σ.*

	Point estimators			95% Confidence intervals				
	MLE	AMLE	Bayes	MLE	AMLE	p-boot	t-boot	Bayes
μ	2.929	3.338	3.649	(2.059,3.798)	(1.785,4.889)	(0.643,5.554)	(1.501,3.641)	(2.831,4.239)
σ	0.998	1.012	1.370	(0.625,1.371)	(0.689,1.335)	(0.251,0.936)	(0.719,0.992)	(0.592,1.604)

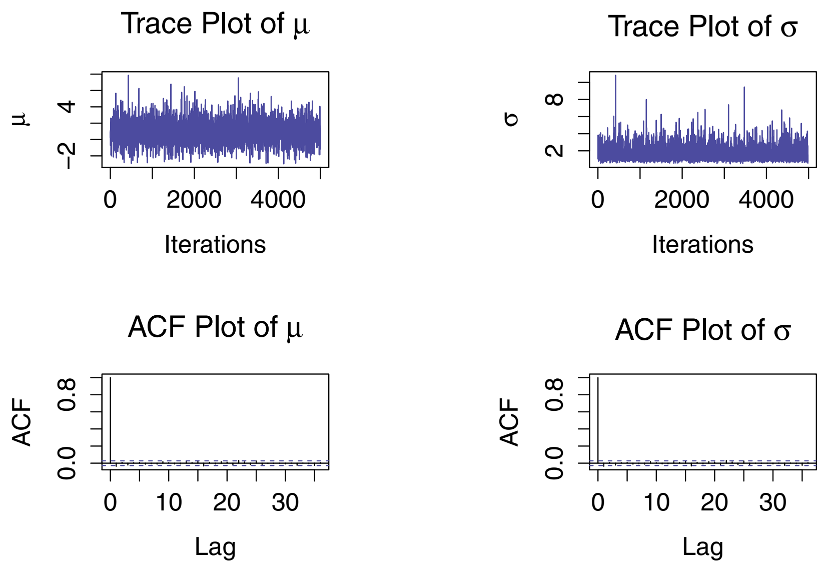

Figure 2: *Trace and autocorrelation plots of μ and σ.*

We also computed the Bayes estimators of μ and σ using Gibbs sampling procedure. To compute the Bayes estimators, since we do not have any prior information, we have used very small (close to zero) values of the hyper-parameters on σ, i.e. $a = b = 0.00001$. In this case, the prior on σ is a proper prior but it is almost improper. Since μ_0 is a location parameter for the logistic prior of μ given σ, without loss of generality, we assumed that $\mu_0 = 0$. For Gibbs sampling procedure we use $N = 5000$ and we have checked the convergence of generated samples of μ and σ. We have used the graphical diagnostics tools like trace plots and autocorrelation function (ACF) plots for this purpose. Figure 2 shows the trace plots and ACF plots for the parameters. The trace plots look like a random scatter and show the fine mixing of the chains for both parameters μ and σ. ACF plots show that chains have very low autocorrelations. Based on these plots, we can fairly conclude that convergence has been attained.

We also computed different confidence intervals namely the approximate confidence intervals based MLEs and AMLEs, p-boot and t-boot confidence intervals and credible intervals. All results are reported in Table 2.

6.2. Simulation study

In this section, a Monte Carlo simulation is conducted to compare the performance of the different estimators. In this simulation, we have randomly generated 1000 upper record sample $X_1, X_2, \ldots, X_m$ from the standard logistic distribution (i.e., $\mu = 0$ and $\sigma = 1$) and then computed the MLEs, AMLEs and Bayes estimators of μ and σ. We then compared the performances of these estimators in terms of biases, and mean square errors (MSEs). For computing Bayes estimators, we take $\mu_0 = 0$. We use both non-informative and informative priors for the scale parameter σ. In case of non-informative prior, we take $a = b = 0$. We call it as Prior 1. For the informative prior, we chose $a = 3$ and $b = 1$. We call it as Prior 2. Clearly Prior 2 is more informative than the non-informative Prior 1.

In Table 3, for different values of m, we reported the average biases, and MSEs of the MLEs, AMLEs and Bayes estimators over 1000 replications. All the computations are performed using Visual Maple (V16) package.

Table 3: *Biases and MSEs of the MLEs, AMLEs and Bayes estimators for different values of m.*

		Estimation of μ				Estimation of σ			
		MLE	AMLE	Bayes		MLE	AMLE	Bayes	
				Prior 1	Prior 2			Prior 1	Prior 2
$m = 2$	Bias	−0.732	−0.749	−0.635	−0.608	0.362	0.386	0.310	0.286
	MSE	2.619	2.654	2.574	2.543	0.510	0.538	0.497	0.467
$m = 3$	Bias	−0.653	−0.681	−0.568	−0.534	0.284	0.297	0.261	0.242
	MSE	2.468	2.492	2.419	2.397	0.451	0.468	0.416	0.402
$m = 5$	Bias	−0.558	−0.579	−0.488	−0.443	0.142	0.166	0.123	0.107
	MSE	2.129	2.171	1.938	1.903	0.109	0.139	0.087	0.073
$m = 10$	Bias	−0.313	−0.366	−0.265	−0.244	0.067	0.084	0.041	0.016
	MSE	1.567	1.636	1.510	1.482	0.059	0.067	0.049	0.041
$m = 15$	Bias	−0.238	−0.250	−0.197	−0.170	0.059	0.063	0.053	0.048
	MSE	1.148	1.176	1.125	1.107	0.043	0.051	0.034	0.027
$m = 20$	Bias	−0.150	−0.177	−0.121	−0.104	0.033	0.045	0.021	0.018
	MSE	0.999	1.024	0.956	0.937	0.024	0.033	0.018	0.011

From Table 3, we observe that the AMLEs and the MLEs are almost identical in terms of both bias and MSEs. The AMLEs are almost as efficient as the MLEs for all sample sizes. Comparing the two Bayes estimators based on two priors 1 and 2, it is observed that the Bayes estimators based on prior 2 perform better than the Bayes estimators based on non-informative prior 1. In addition, the Bayes estimators perform better than the classical estimators MLEs and AMLEs. It is also noted as m increases, the performances of all estimators better in terms of biases and MSEs.

We also computed the 95% confidence/credible intervals for μ and σ based on the asymptotic distributions of the MLEs and AMLEs. We further computed Boot-p, and

Table 4: *Average confidence/credible lengths and coverage probabilities for different values of m.*

			MLE	AMLE	p-boot	t-boot	Bayes	
							Prior 1	Prior 2
Estimation of μ	$m = 2$	Length	1.964	1.972	1.937	1.925	1.916	1.892
		Cov. Prob.	0.937	0.936	0.938	0.939	0.939	0.940
	$m = 3$	Length	1.729	1.741	1.709	1.686	1.681	1.669
		Cov. Prob.	0.938	0.937	0.940	0.941	0.940	0.941
	$m = 5$	Length	1.411	1.424	1.384	1.377	1.358	1.345
		Cov. Prob.	0.939	0.937	0.941	0.942	0.941	0.943
	$m = 10$	Length	1.097	1.110	1.068	1.046	1.028	1.009
		Cov. Prob.	0.941	0.940	0.943	0.943	0.943	0.944
	$m = 15$	Length	0.804	0.811	0.794	0.783	0.752	0.739
		Cov. Prob.	0.943	0.942	0.944	0.945	0.945	0.946
	$m = 20$	Length	0.653	0.673	0.634	0.625	0.605	0.590
		Cov. Prob.	0.945	0.943	0.945	0.947	0.947	0.948
Estimation of σ	$m = 2$	Length	1.310	1.328	1.286	1.279	1.271	1.260
		Cov. Prob.	0.939	0.936	0.939	0.940	0.939	0.941
	$m = 3$	Length	1.186	1.197	1.172	1.164	1.152	1.140
		Cov. Prob.	0.941	0.939	0.941	0.941	0.942	0.943
	$m = 5$	Length	0.924	0.931	0.907	0.902	0.894	0.887
		Cov. Prob.	0.942	0.940	0.943	0.944	0.944	0.945
	$m = 10$	Length	0.716	0.724	0.701	0.694	0.680	0.671
		Cov. Prob.	0.943	0.941	0.943	0.944	0.945	0.946
	$m = 15$	Length	0.543	0.560	0.530	0.522	0.516	0.505
		Cov. Prob.	0.944	0.942	0.944	0.945	0.946	0.948
	$m = 20$	Length	0.375	0.383	0.366	0.359	0.352	0.345
		Cov. Prob.	0.946	0.945	0.945	0.947	0.948	0.949

Boot-t confidence intervals, and the credible intervals. Table 4 presents the average confidence/credible lengths and the corresponding coverage probability over 1000 replications. The nominal level for the confidence intervals is 0.95 in each case.

From Table 4, the length of the 95% confidence interval based on the asymptotic distribution of the MLE, is slightly smaller than the corresponding length of the interval based on the asymptotic distribution of the AMLE. We also observe that the Bayesian credible intervals work slightly better than the bootstrap and asymptotic confidence intervals in terms of both confidence length and coverage probability. Also, Boot-t confidence intervals perform very similarly to the Bayesian credible intervals. The bootstrap confidence intervals work better than the asymptotic confidence intervals. The Boot-t confidence intervals perform better than the Boot-p confidence intervals. Also, it is observed that all the simulated coverage probabilities are very close to the nominal level

95%. Also, for all interval estimators, the confidence lengths and the simulated coverage percentages decrease as m increases.

Overall speaking, from Tables 3 and 4, we would recommend the use of Bayesian method for point and interval estimation, especially when reliable prior information about the logistic parameters is available.

Appendix A

To prove $B \leq 0$, we need to show that

$$\frac{(\beta_m y_m + \sum_{i=1}^{m} \beta_i y_i)^2}{\beta_m + \sum_{i=1}^{m} \beta_i} \leq (\beta_m y_m^2 + \sum_{i=1}^{m} \beta_i y_i^2),$$

or equivalently

$$2\beta_m y_m \sum_{i=1}^{m} \beta_i y_i + \left(\sum_{i=1}^{m} \beta_i y_i\right)^2 \leq \beta_m \sum_{i=1}^{m} \beta_i y_i^2 + \beta_m y_m^2 \sum_{i=1}^{m} \beta_i + \left(\sum_{i=1}^{m} \beta_i\right)\left(\sum_{i=1}^{m} \beta_i y_i^2\right). \tag{A.1}$$

We can rewrite (A.1) as

$$\beta_m \left(\sum_{i=1}^{m} \beta_i [2y_m y_i]\right) + \left(\sum_{i=1}^{m} \beta_i y_i\right)^2 \leq \beta_m \left(\sum_{i=1}^{m} \beta_i [y_i^2 + y_m^2]\right) + \left(\sum_{i=1}^{m} \beta_i\right)\left(\sum_{i=1}^{m} \beta_i y_i^2\right). \tag{A.2}$$

Now, since $y_i^2 + y_j^2 \geq 2y_i y_j$, we have

$$\beta_m \left(\sum_{i=1}^{m} \beta_i [2y_m y_i]\right) \leq \beta_m \left(\sum_{i=1}^{m} \beta_i [y_i^2 + y_m^2]\right), \tag{A.3}$$

and

$$\left(\sum_{i=1}^{m} \beta_i y_i\right)^2 \leq \left(\sum_{i=1}^{m} \beta_i\right)\left(\sum_{i=1}^{m} \beta_i y_i^2\right). \tag{A.4}$$

Now by using (A.3) and (A.4), (A.2) is true and the proof is thus obtained.

Appendix B (Proof of Theorem 1)

The log-likelihood function of $g_2(\mu|\sigma)$ is

$$\ln g_2(\mu|\sigma,\mathbf{y}) \propto -\sum_{i=1}^{m} \ln(1+e^{-(y_i-\mu)/\sigma}) - \ln(1+e^{-(y_m-\mu)/\sigma}) - 2\ln(1+e^{-(\mu-\mu_0)/\sigma}).$$

The second derivative of $\ln g_2(\mu|\sigma,\mathbf{y})$ is obtained as

$$-\frac{1}{\sigma^2}\left[\sum_{i=1}^{m}\frac{e^{-(y_i-\mu)/\sigma}}{(1+e^{-(y_i-\mu)/\sigma})^2}+\frac{e^{-(y_m-\mu)/\sigma}}{(1+e^{-(y_m-\mu)/\sigma})^2}+\frac{2\,e^{-(\mu-\mu_0)/\sigma}}{(1+e^{-(\mu-\mu_0)/\sigma})^2}\right],$$

which is negative. So, the result follows.

Appendix C (Proof of Lemma 1)

Using the relation (2.5), we have

$$\begin{aligned} E\left[f(X_i)\right] &= E\left[F(X_i)(1-F(X_i))\right] \\ &= \int_{-\infty}^{\infty} F(x)[1-F(x)]\frac{\left[-\ln(1-F(x))\right]^{i-1}}{(i-1)!}f(x)dx \\ &= \int_{0}^{1} u(1-u)\frac{[-\ln(1-u)]^{i-1}}{(i-1)!}du \\ &= \int_{0}^{\infty}(1-e^{-t})e^{-2t}\frac{t^{i-1}}{(i-1)!}dt = \frac{1}{2^i}-\frac{1}{3^i}, \end{aligned}$$

and

$$\begin{aligned} E\left[F(X_i)\right] &= \int_{-\infty}^{\infty} F(x)\frac{\left[-\ln(1-F(x))\right]^{i-1}}{(i-1)!}f(x)dx \\ &= \int_{0}^{1} u\frac{[-\ln(1-u)]^{i-1}}{(i-1)!}du \\ &= \int_{0}^{\infty}(1-e^{-t})e^{-t}\frac{t^{i-1}}{(i-1)!}dt = 1-\frac{1}{2^i}. \end{aligned}$$

We also have

$$\begin{aligned} E\left[X_i f(X_i)\right] &= E\left[X_i F(X_i)(1-F(X_i))\right] \\ &= \int_{-\infty}^{\infty} xF(x)[1-F(x)]\frac{[-\ln(1-F(x))]^{i-1}}{(i-1)!}f(x)dx \\ &= \int_0^1 [\ln u - \ln(1-u)]u(1-u)\frac{[-\ln(1-u)]^{i-1}}{(i-1)!}du, \end{aligned}$$

since $F^{-1}(u) = \ln u - \ln(1-u)$. Setting $t = -\ln(1-u)$, we get

$$\begin{aligned} E\left[X_i f(X_i)\right] &= \int_0^{\infty} \ln(1-e^{-t})e^{-2t}(1-e^{-t})\frac{t^{i-1}}{(i-1)!}dt + \int_0^{\infty} e^{-2t}(1-e^{-t})\frac{t^i}{(i-1)!}dt \\ &= \sum_{l=1}^{\infty}\left[\frac{1}{l(l+3)^i} - \frac{1}{l(l+2)^i}\right] + i\left[\frac{1}{2^i} - \frac{1}{3^i}\right]. \end{aligned}$$

The two other expectations $E\left[X_i(1-F(X_i))\right]$ and $E[X_i^2 f(X_i)]$, can be obtained in the same manner using the binomial expansion and writing $\ln(1-e^{-t}) = -\sum_{l=1}^{\infty}\frac{e^{-lt}}{l}$.

Acknowledgements

We would like to thank the editor for his encouragement and the referees for their constructive suggestions and comments that substantially improved the paper.

References

Ahsanullah, M. (1995). *Record Values-Theory and Applications*. New York, University Press of America Inc.

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

Asgharzadeh, A. (2006). Point and interval estimation for a generalized logistic distribution under progressive Type II censoring. *Communications in Statistics-Theory and Methods* 35, 1685–1702.

Asgharzadeh, A., Valiollahi, R. and Raqab, M. Z. (2011). Stress-strength reliability of Weibull distribution based on progressively censored samples. *SORT*, 35, 103–124.

Balakrishnan, N. (1992). *Handbook of Logistic Distribution*. New York, Dekker.

Balakrishnan, N. and Aggarwala, R. (2000). *Progressive Censoring: Theory Methods and Applications*. Birkhäuser, Boston.

Balakrishnan, N. and Asgharzadeh, A. (2005). Inference for the scaled half-logistic distribution based on progressively Type II censored samples. *Communications in Statistics-Theory and Methods,* 34, 73–87.

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

Balakrishnan, N., Kannan, N. (2000). Point and interval estimation for the parameters of the logistic distribution based on progressively Type-2 censored samples. *Handbook of Statistics*, 20, 431–456.

Chandler, K. N. (1952). The distribution and frequency of record values. *Journal of the Royal Statistical Society*, B14, 220–228.

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

Devroye, L. (1984). A simple algorithm for generating random variates with a log-concave density. *Computing*, 33, 247–257.

Efron, B. (1982). The Jackknife, the Bootstrap and Other Re-Sampling Plans. *Philadelphia, PA: SIAM, vol. 38, CBMS-NSF Regional Conference Series in Applied Mathematics.*

Hall, P. (1988). Theoretical comparison of bootstrap confidence intervals, *Annals of Statistics,* 16, 927–953.

Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions, vol. 2*, 2-nd edition. New York, John Wiley and Sons.

Kundu, D. (2007). On hybrid censoring Weibull distribution. *Journal of Statistical Planning and Inference*, 137, 2127–2142.

Kundu, D. (2008). Bayesian inference and life testing plan for the Weibull distribution in presence of progressive censoring. *Technometrics*, 50,144–154.

Madi, M. T. and Raqab, M. Z. (2007). Bayesian prediction of rainfall records using the generalized exponential distribution. *Environmetrics*, 18, 541–549.

Nadarajah, S. (2004). Information matrix for logistic distributions. *Mathematical and Computer Modelling*, 40, 953–958.

Nevzorov, V. (2001). *Records: Mathematical Theory*. In: Translation of Mathematical Monographs, vol. 194. Amer. Math. Soc. Providence, RI, USA.

Raqab, M. Z. (2006). Nonparametric prediction intervals for the future rainfall records. *Environmetrics*, 17, 457–464.

Raqab, M. Z., Asgharzadeh, A. and Valiollahi. R. (2010). Prediction for Pareto distribution based on progressively Type-II censored samples. *Computational Statistics and Data Analysis*, 54, 1732–1743.

Tiku, M. L. and Akkaya A. D. (2004) *Robust Estimation and Hypothesis Testing*. New Delhi, New Age International.

Shao, J. (2003) *Mathematical Statistics*. Second edition, New Work, Springer-Verlag.

Shao, J. (2005) *Mathematical Statistics: Exercises and Solutions*. New York, Springer-Verlag.