Journal of the Korean Statistical Society (2006), **35**: *2, pp 167–177*

STATISTICAL EVIDENCE METHODOLOGY FOR MODEL ACCEPTANCE BASED ON RECORD VALUES

M. DOOSTPARAST[1] AND M. EMADI[2]

ABSTRACT

An important role of statistical analysis in science is interpreting observed data as evidence, that is "what do the data say?". Although standard statistical methods (hypothesis testing, estimation, confidence intervals) are routinely used for this purpose, the theory behind those methods contains no defined concept of evidence and no answer to the basic question "when is it correct to say that a given body of data represent evidence supporting one statistical hypothesis against another?" (Royall, 1997).

In this article, we use likelihood ratios to measure evidence provided by record values in favor of a hypothesis and against an alternative. This hypothesis is concerned on mean of an exponential model and prediction of future record values.

AMS 2000 subject classifications. Primary 62G10; Secondary 62G30.

Keywords. Exponential model, hypothesis testing, likelihood ratio, prediction, record values, statistical evidence.

1. INTRODUCTION

Let $\{X_i, i \geq 1\}$ be a sequence of independent and identically distributed continuous random variables having the same distribution as the (population) random variable X. An observation X_j will be called an upper record value if its value exceeds that of all previous observations. Thus X_j is an (upper) record value if $X_j > X_i$ for every $i < j$. Let us assume that X_j is observed at time j. Then the (upper) record time $\{T_n, n \geq 1\}$ sequence is defined in the following manner: $T_1 = 1$, with probability 1 and, for $n \geq 2$, $T_n = \min\{j > T_{n-1} : X_j > X_{T_{n-1}}\}$.

Received August 2005; accepted April 2006.

[1]Corresponding author. Department of Statistics, Ferdowsi University of Mashhad, P. O. Box 91775-1159, Mashhad, Iran (e-mail: doostparast@wali.um.ac.ir)

[2]Department of Statistics, Ferdowsi University of Mashhad, P. O. Box 91775-1159, Mashhad, Iran (e-mail: emadi@math.um.ac.ir)

The record value sequence is defined by $R_n = X_{T_n}$, $n = 1, 2, 3, \ldots$. Suppose we observe the first m record values $R_1 = r_1, R_2 = r_2, \ldots, R_m = r_m$ from the cumulative distribution function (cdf) $F(x)$ and the probability density function (pdf) $f(x)$. Then, the joint distribution of the first m record value (for more details see Arnold *et al.*, 1998) is given by

$$f(\boldsymbol{r}) = f(r_m) \prod_{i=1}^{m-1} h(r_i), \; r_1 < r_2 < \cdots < r_m, \tag{1.1}$$

where $\boldsymbol{r} = (r_1, r_2, \ldots, r_m)$ and $h(r_i) = f(r_i)/\{1 - F(r_i)\}$.

There has been general interest in record values for centuries, particularly for sporting events like the Olympic Games. But its interest as a probabilistic or statistical subject is of much more recent vintage. Motivated by the reported frequency of record weather conditions, Chandler (1952) began studying the distributions of record data for independent and identically distributed sequence of random variables. Since then, a substantial stream of papers on record values has flowed into the literature. Interested readers may refer to Glick (1978), Ahsanullah (1995), Arnold *et al.* (1998) and Nevzorov (2001) for a review of developments in this area of research. There are also some papers done on statistical inference based on record values. See for instance Samaniego and Whitaker (1986), Gulati and Padegett (2003), Feuerverger and Hall (1998), Berred (1998), Ahmadi (2000), Ahmadi *et al.* (2005) and Ahmadi and Doostparast (2006).

A random variable X is said to have an exponential distribution, denoted by $X \sim Exp(\sigma)$, if its cdf is

$$F(x;\sigma) = 1 - e^{-\frac{x}{\sigma}}, \; x \geq 0, \; \sigma > 0, \tag{1.2}$$

and hence the pdf is given by

$$f(x;\sigma) = \frac{1}{\sigma} e^{-\frac{x}{\sigma}}, \; x \geq 0, \; \sigma > 0. \tag{1.3}$$

The exponential distribution is applied in a wide variety of statistical procedures, especially in life testing problems. Data for survival and reliability analysis, as well as for biomedical and life testing studies have been modeled extensively by exponential model. Based on records, a good review of classical inference may be found in Ahsanullah (1980).

Through this paper, we assume that the data available for study are record values. Such data may be rewritten as

$$R_1, R_2, \ldots, R_m, \tag{1.4}$$

where R_i is the ith record value or new maximum.

The remainder of this article is organized as follows. Section 2 contains a short review of statistical evidence. In section 3, we present likelihood paradigm based on record values. Section 4 deals with prediction problem of future record values. Section 5 includes some remarks and future works.

2. STATISTICAL EVIDENCE

Standard statistical methods regularly lead to the misinterpretation of results of data. The errors are usually quantitative, as when statistical evidence is judged to be stronger or weaker than it really is. But sometimes they are qualitative-sometimes evidence is judged to support one hypothesis over another when the opposite is true. These misinterpretations are not a consequence of scientists misusing statistics. They reflect instead a critical defect in current theories of statistics.

These problems exist because the discipline of statistics has neglected a key question for which it is responsible: when does a given set of observations support one statistical hypothesis over another? That is, when is it right to say that the observations are evidence in favor of one hypothesis *vis-a-vis* another? The answer to this fundamental question has been known for at least century. However, neither the question nor its simple answer is to be found in most modern statistics text books. The reason is that the decision-making paradigms since the work of Neyman and Pearson in the 1930s, have been formulated not in terms of interpreting data as evidence, but in terms of choosing between alternative course of action. This lead to the current state of affairs in which the dominant (Neyman-Pearson) *theory* view common statistical procedures as decision-making tools, while much of statistical *practice* consists of using the same procedures for a different purpose, namely, interpreting data as evidence. For example, in Neyman-Pearson statistical theory, a test of two hypothesis H_1 and H_2 is represented as a procedure for choosing between them. But in applications, when an optimal test chooses H_2, it is often taken to mean that data are evidence favoring H_2 over H_1. This interpretation can be quite wrong. Interested readers may refer to Royall (1997).

Let η (> 0) be any measure of support of H_1 against H_2. Large (Small) values of η are interpreted as evidence given by data in favor of H_1 (H_2). Emadi and Arghami (2003) have studied some measures of support for statistical hypotheses. The probabilities of observing strong misleading evidence under H_1 and H_2 are

$$M_1 := P\left(\eta < \frac{1}{k}|H_1 \text{ is true}\right), \tag{2.1}$$

and

$$M_2 := P\left(\eta > k|H_2 \text{ is true}\right), \tag{2.2}$$

respectively. Also, the probabilities of weak evidence under H_1 and H_2 are

$$W_1 := P\left(\frac{1}{k} < \eta < k|H_1 \text{ is true}\right), \tag{2.3}$$

and

$$W_2 := P\left(\frac{1}{k} < \eta < k|H_2 \text{ is true}\right), \tag{2.4}$$

respectively (Royall, 2000).

3. Evidential Model Testing

In this section, we shall be concerned with statistical evidence for exponential model based on record values. Suppose, we can observe the sequence of record values $R_1 = r_1, R_2 = r_2, \ldots, R_m = r_m$ from exponential distribution with cdf and pdf given by (1.2) and (1.3), respectively. Let λ be the likelihood ratio for the competing hypotheses H_1 and H_2, *i.e.*

$$\lambda = \frac{L_1}{L_2}, \tag{3.1}$$

where L_i is likelihood function under H_i. We will use λ as a measure of support H_1 against H_2.

Notice that by (1.1), (1.2) and (1.3) it is easy to verify that the likelihood function is given by

$$L(\sigma; \boldsymbol{r}) = \left(\frac{1}{\sigma}\right)^m e^{-r_m/\sigma}, \ \sigma > 0. \tag{3.2}$$

It can be shown that $2R_m/\sigma$ has chisquare distribution with $2m$ degree of freedom. Two hypothesis $H_1 : \sigma = \sigma_1$ and $H_2 : \sigma = \sigma_2$ are under consideration where $0 < \sigma_1 < \sigma_2$. By (3.1) and (3.2) the likelihood ratio for the competing hypothesis H_1 and H_2 is given by

$$\lambda = \left(\frac{\sigma_2}{\sigma_1}\right)^m e^{r_m\left(\frac{1}{\sigma_2}-\frac{1}{\sigma_1}\right)}. \tag{3.3}$$

PROPOSITION 3.1. *Misleading evidences are given by*

$$M_1 = 1 - F_{\chi^2_{(2m)}}\left\{2\frac{\ln(k) + m\ln(\sigma_2/\sigma_1)}{1-\frac{\sigma_1}{\sigma_2}}\right\}, \tag{3.4}$$

$$M_2 = F_{\chi^2_{(2m)}}\left\{2\frac{-\ln(k) + m\ln(\sigma_2/\sigma_1)}{\frac{\sigma_2}{\sigma_1}-1}\right\}. \tag{3.5}$$

where $F_{\chi^2_{(v)}}(\cdot)$ is cdf of a chisquare distribution with v degree of freedom.

PROOF. By (2.1) and (3.3) we have

$$\begin{aligned}
M_1 &= P\left(\lambda < \frac{1}{k} \;\middle|\; H_1 \text{ is true}\right)\\
&= P\left\{\left(\frac{\sigma_2}{\sigma_1}\right)^m e^{R_m\left(\frac{1}{\sigma_2}-\frac{1}{\sigma_1}\right)} < \frac{1}{k} \;\middle|\; \sigma = \sigma_1\right\}\\
&= P\left\{R_m > \frac{-\ln(k) - m\ln(\sigma_2/\sigma_1)}{\frac{1}{\sigma_2}-\frac{1}{\sigma_1}} \;\middle|\; \sigma = \sigma_1\right\}\\
&= 1 - F_{\chi^2_{(2m)}}\left\{2\frac{\ln(k) + m\ln(\sigma_2/\sigma_1)}{1-\frac{\sigma_1}{\sigma_2}}\right\},
\end{aligned}$$

and similarly proceeding

$$\begin{aligned}
M_2 &= P\left(\lambda > k \mid H_2 \text{ is true}\right)\\
&= P\left\{\left(\frac{\sigma_2}{\sigma_1}\right)^m e^{R_m\left(\frac{1}{\sigma_2}-\frac{1}{\sigma_1}\right)} > k \;\middle|\; \sigma = \sigma_2\right\}\\
&= P\left\{R_m < \frac{\ln(k) - m\ln(\sigma_2/\sigma_1)}{\frac{1}{\sigma_2}-\frac{1}{\sigma_1}} \;\middle|\; \sigma = \sigma_2\right\}\\
&= F_{\chi^2_{(2m)}}\left\{2\frac{-\ln(k) + m\ln(\sigma_2/\sigma_1)}{\frac{\sigma_2}{\sigma_1}-1}\right\}.
\end{aligned}$$

These prove (3.4) and (3.5). □

COROLLARY 3.1. *By proposition 3.1 we have*

(i) $\lim_{\sigma_2 \to +\infty} M_1 = \lim_{\sigma_2 \to +\infty} M_2 = 0.$

(ii) $\lim_{\sigma_2 \to \sigma_1^+} M_1 = \lim_{\sigma_2 \to \sigma_1^+} M_2 = 0.$

(iii) *The point of global maximum of M_1 and M_2 can be obtained as a solution of the following non-linear equations*

$$m - m\ln(\sigma_1) + \ln(k) = m\frac{\sigma_2}{\sigma_1} - m\ln(\sigma_2), \tag{3.6}$$

and

$$m + \ln(k) + m\ln(\sigma_1) = m\frac{\sigma_1}{\sigma_2} + m\ln(\sigma_2), \tag{3.7}$$

respectively.

(iv) *For $\sigma_2 < \sigma_1 e^{\sqrt[m]{k}}$, we have $M_2 = 0$.*

PROOF. (i) and (ii) can be prove by L'Hopital rule. The (3.6) and (3.7) are consequence of following equations

$$\partial M_1/\partial\sigma_2 = 0, \ \partial M_2/\partial\sigma_2 = 0.$$

(iv) can be easily shown from (3.5). □

Based on (3.4) and (3.5), the values of M_1 and M_2 are shown in Figure 3.1.

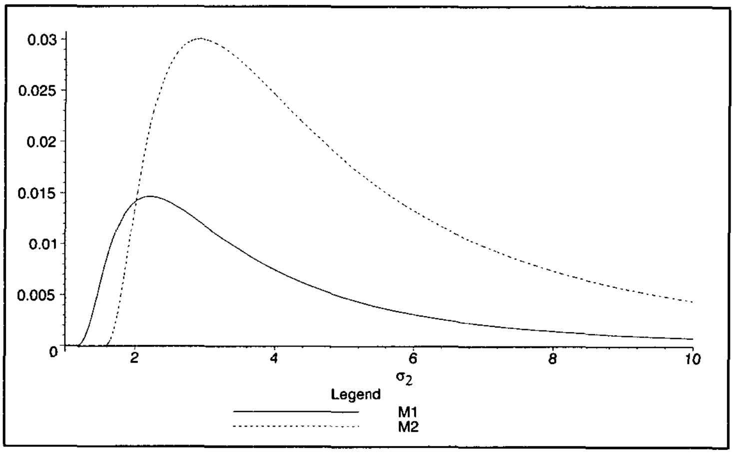

FIGURE 3.1 *The values of M_1 and M_2 for $k = 8$, $m = 5$ and $\sigma_1 = 1$.*

REMARK 3.1. It may be noticed that when σ_2 tends to infinity, the distance between populations will increase as much as possible. Hence the probability of misleading tend to zero. Also, when σ_2 tends to σ_1, the distance between two populations will decrease as much as possible. So the M_1 and M_2 will be mixed with W_1 and W_2 and they tend to zero. As we will see in corollary 3.2, in this case, for determination of true hypothesis we will therefore need more record values or more data (thus generating more record values).

PROPOSITION 3.2. *Weak evidences are given by*

$$W_1 = F_{\chi^2_{(2m)}}\left\{2\frac{\ln(k)+m\ln(\sigma_2/\sigma_1)}{1-\frac{\sigma_1}{\sigma_2}}\right\} - F_{\chi^2_{(2m)}}\left\{2\frac{-\ln(k)+m\ln(\sigma_2/\sigma_1)}{1-\frac{\sigma_1}{\sigma_2}}\right\}, \quad (3.8)$$

$$W_2 = F_{\chi^2_{(2m)}}\left\{2\frac{\ln(k)+m\ln(\sigma_2/\sigma_1)}{\frac{\sigma_2}{\sigma_1}-1}\right\} - F_{\chi^2_{(2m)}}\left\{2\frac{-\ln(k)+m\ln(\sigma_2/\sigma_1)}{\frac{\sigma_2}{\sigma_1}-1}\right\}. \quad (3.9)$$

PROOF. By (2.3) and (3.3) we have

$$\begin{aligned}
W_1 &= P\left(\frac{1}{k} < \lambda < k \mid H_1 \text{ is true}\right) \\
&= P\left\{\left(\frac{\sigma_2}{\sigma_1}\right)^m e^{R_m(\frac{1}{\sigma_2}-\frac{1}{\sigma_1})} < k \;\middle|\; \sigma = \sigma_1\right\} \\
&\quad -P\left\{\left(\frac{\sigma_2}{\sigma_1}\right)^m e^{R_m(\frac{1}{\sigma_2}-\frac{1}{\sigma_1})} < \frac{1}{k} \;\middle|\; \sigma = \sigma_1\right\} \\
&= P\left\{2\frac{-\ln(k)+m\ln(\sigma_2/\sigma_1)}{1-\frac{\sigma_1}{\sigma_2}} < \frac{2R_m}{\sigma_1} < 2\frac{\ln(k)+m\ln(\sigma_2/\sigma_1)}{1-\frac{\sigma_1}{\sigma_2}} \;\middle|\; \sigma = \sigma_1\right\} \\
&= F_{\chi^2_{(2m)}}\left\{2\frac{\ln(k)+m\ln(\sigma_2/\sigma_1)}{1-\frac{\sigma_1}{\sigma_2}}\right\} - F_{\chi^2_{(2m)}}\left\{2\frac{-\ln(k)+m\ln(\sigma_2/\sigma_1)}{1-\frac{\sigma_1}{\sigma_2}}\right\},
\end{aligned}$$

and similarly proceeding

$$\begin{aligned}
W_2 &= P\left(\frac{1}{k} < \lambda < k \mid H_2 \text{ is true}\right) \\
&= P\left\{\frac{2R_m}{\sigma_2} > 2\frac{\ln(k)-m\ln(\sigma_2/\sigma_1)}{1-\frac{\sigma_2}{\sigma_1}} \;\middle|\; \sigma = \sigma_2\right\} \\
&\quad -P\left\{\frac{2R_m}{\sigma_2} > 2\frac{\ln(1/k)-m\ln(\sigma_2/\sigma_1)}{1-\frac{\sigma_2}{\sigma_1}} \;\middle|\; \sigma = \sigma_2\right\}
\end{aligned}$$

$$= P\left\{2\frac{-\ln(k)+m\ln(\sigma_2/\sigma_1)}{\frac{\sigma_2}{\sigma_1}-1} < \frac{2R_m}{\sigma_2} < 2\frac{\ln(k)+m\ln(\sigma_2/\sigma_1)}{\frac{\sigma_2}{\sigma_1}-1} \;\Big|\; \sigma=\sigma_2\right\}$$

$$= F_{\chi^2_{(2m)}}\left\{2\frac{\ln(k)+m\ln(\sigma_2/\sigma_1)}{\frac{\sigma_2}{\sigma_1}-1}\right\} - F_{\chi^2_{(2m)}}\left\{2\frac{-\ln(k)+m\ln(\sigma_2/\sigma_1)}{\frac{\sigma_2}{\sigma_1}-1}\right\}.$$

These prove (3.8) and (3.9). □

COROLLARY 3.2. *By proposition 3.2 we have*

(i) $\lim_{\sigma_2\to+\infty} W_1 = \lim_{\sigma_2\to+\infty} W_2 = 0$.

(ii) $\lim_{\sigma_2\to\sigma_1^+} W_1 = \lim_{\sigma_2\to\sigma_1^+} W_2 = 1$.

(iii) *The point of global maximum of W_1 and W_2 can be obtained as a solution of the following non-linear equations*

$$\left[\frac{m\left(\frac{\sigma_2}{\sigma_1}-1\right)-\left\{\ln(k)+m\ln\left(\frac{\sigma_2}{\sigma_1}\right)\right\}}{m\left(\frac{\sigma_2}{\sigma_1}-1\right)-\left\{-\ln(k)+m\ln\left(\frac{\sigma_2}{\sigma_1}\right)\right\}}\right] = \left\{\frac{-\ln(k)+m\ln\left(\frac{\sigma_2}{\sigma_1}\right)}{\ln(k)+m\ln\left(\frac{\sigma_2}{\sigma_1}\right)}\right\}^{\frac{\frac{\sigma_2}{\sigma_1}(m-1)}{\left(\frac{\sigma_2}{\sigma_1}-1\right)}}, \tag{3.10}$$

and

$$\left[\frac{m\left(1-\frac{\sigma_1}{\sigma_2}\right)-\left\{\ln(k)+m\ln\left(\frac{\sigma_2}{\sigma_1}\right)\right\}}{m\left(1-\frac{\sigma_1}{\sigma_2}\right)-\left\{-\ln(k)+m\ln\left(\frac{\sigma_2}{\sigma_1}\right)\right\}}\right] = \left\{\frac{-\ln(k)+m\ln\left(\frac{\sigma_2}{\sigma_1}\right)}{\ln(k)+m\ln\left(\frac{\sigma_2}{\sigma_1}\right)}\right\}^{\frac{m-1}{\left(\frac{\sigma_2}{\sigma_1}-1\right)}}, \tag{3.11}$$

respectively.

PROOF. (i) and (ii) can be prove by L'Hopital rule. The (3.10) and (3.11) are consequence of following equations

$$\partial W_1/\partial\sigma_2 = 0, \quad \partial W_2/\partial\sigma_2 = 0.$$

□

Based on (3.8) and (3.9), the values of W_1 and W_2 are shown in Figure 3.2.

REMARK 3.2. When σ_2 tends to infinity, the distance between populations will increase as much as possible. So, even with few data we can make the decision about true hypothesis. Also, when σ_2 tends to σ_1, the distance between two populations will decrease as much as possible. Hence we have a few record values to determine true hypothesis, so we need more data.

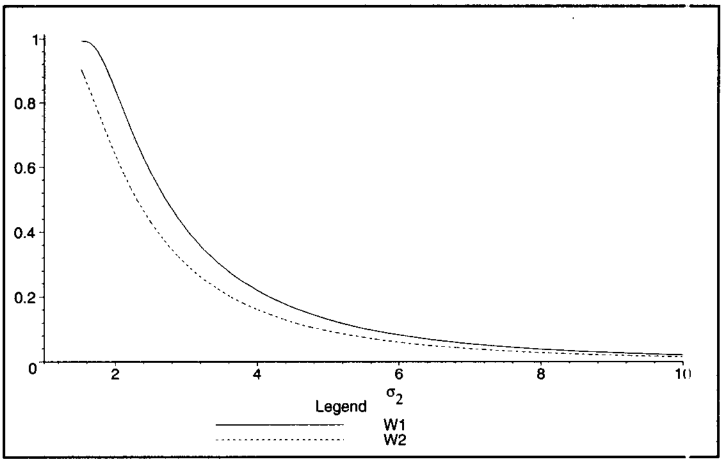

FIGURE 3.2 *The values of* W_1 *and* W_2 *for* $k = 8$, $m = 5$ *and* $\sigma_1 = 1$.

4. PREDICTION OF FUTURE RECORD

Let L^2 be the vector space of random variables Z with $E(Z^2) < \infty$. Let $X \in L^2$ be an unobservable random variable, whose value we wish to predict from observation of other random variables $Y_1, \ldots, Y_n$. In order to use the knowledge of $Y_1, \ldots, Y_n$ to predict X, the predictor must be a function of $Y_1, \ldots, Y_n$, and, in particular, is a random variable. To solve such a prediction problem, we must identify the "best" predictor according to some optimality criterion. In this article, our criterion is *mean squared error* (MSE). The mean squared error of a predictor Z is

$$MSE(Z) := E(Z - X)^2.$$

Assuming it exists, the optimal predictor,

$$\hat{X} := \arg\min_{Z} MSE(Z),$$

the *minimum mean squared error* (MMSE) predictor of X, then satisfies

$$E(\hat{X} - X)^2 \leq E(Z - X)^2,$$

for all Z. It can be show that the conditional expectation of a random variable X given $Y_1, \ldots, Y_n$ is the MMSE predictor of X among functions of $Y_1, \ldots, Y_n$.

Assume that we have the first m record values $R_1 = r_1, ..., R_m = r_m$ from the $Exp(\sigma)$. Based on such a sample, prediction is needed for s-th record $(1 \leq m < s)$. Now let $Y = R_s$ be the s-th record value. The conditional pdf of Y given the m record values had been observed is given by (Arnold *et al.*, 1998)

$$f(y|\boldsymbol{r}) = \frac{\{H(y) - H(r_m)\}^{s-m-1}}{\Gamma(s-m)} \frac{f(y)}{1 - F(r_m)}, \quad r_m < y < \infty, \tag{4.1}$$

where $H(\cdot) = -\ln\{1 - F(\cdot)\}$. So, by (1.2), (1.3) and (4.1) we have

$$f(y|r_m, \sigma) = \frac{(y - r_m)^{s-m-1}}{\sigma^{s-m}\Gamma(s-m)} e^{-\frac{1}{\sigma}(y-r_m)}, \quad y \geq r_m. \tag{4.2}$$

By (4.2), we have

$$E(R_s|\boldsymbol{r}) = r_m + \sigma(s-m). \tag{4.3}$$

Hence, hypothesis $H_i' : E(R_s|\boldsymbol{r}) = a_i$ $(i = 1, 2)$ are equivalence to $H_i : \sigma = (a_i - r_m)/(s-m)$. So, statistical evidence can be captured by mentioned procedures in section 3.

5. Further Research

1. We may consider one side hypothesis $H_1 : \sigma \leq \sigma_1$ against $H_2 : \sigma > \sigma_1$ or more generally $H_1 : \sigma_1 \in \Omega_1$ and $H_2 : \sigma_2 \in \Omega_2$ where $\Omega_1 \bigcap \Omega_2 = \phi$, Ω_1 and Ω_2 are subsets of $(0, +\infty)$. This situation needs a new measure of support for one hypothesis against another.

2. Exponential distribution has been demonstrated to provide good approximations to many lifetime distributions, the inference procedures based on the model are widely used. Some scientists, for example, points out that the time between failures of machine approximately follows exponential model. Owing to technology development and other causes, the failure of machine may be changed from previous exponential distribution. Therefore, the assumption of Weibull model is better than that of exponential distribution. So, the results of this paper may be extended to Weibull model and some other lifetime distributions.

Acknowledgments

The authors thank the referees for their valuable suggestions which improved the presentation of this paper. The authors also thank Professor N. Balakrishnan of the McMaster University, Canada for his comments.

REFERENCES

AHMADI, J. (2000). *Record values, theory and applications*, Ph.D. Dissertation, Ferdowsi University of Mashhad, Mashhad, Iran.

AHMADI, J. AND DOOSTPARAST, M. (2006). "Bayesian estimation and prediction for some life distributions based on record values", *Statistical Papers*, to appear.

AHMADI, J., DOOSTPARAST, M. AND PARSIAN, A. (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**, 795–805.

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

AHSANULLAH, M. (1980). "Linear prediction of record values for the two-parameter exponential distribution", *Annals of the Institute of Statistical Mathematics*, **32**, 363–368.

ARNOLD, B. C., BALAKRISHNAN, N. AND NAGARAJA, H. N. (1998). *Records*, John Wiley & Sons, New York.

BERRED, A. M. (1998). "Prediction of record values", *Communications in Statistics–Theory and Methods*, **27**, 2221–2240.

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

EMADI, M. AND ARGHAMI, N. R. (2003). "Some measures of support for statistical hypotheses", *Journal of Statistical Theory and Applications*, **2**, 165–176.

FEUERVERGER, A. AND HALL, P. (1998). "On statistical inference based on record values", *Extremes*, **1**, 169–190.

GLICK, N. (1978). "Breaking record and breaking boards", *The American Mathematical Monthly*, **85**, 2–26.

GULATI, S. AND PADGETT, W. J. (2003). *Parametric and Nonparametric Inference from Record-breaking Data*, Springer-Verlag, New York.

NEVZOROV, V. B. (2001). *Records: Mathematical Theory. Translated from the Russian Manuscript by D. M. Chibisov*, American Mathematical Society, Providence, RI.

ROYALL, R. M. (1997). *Statistical Evidence. A Likelihood Paradigm*, Chapman and Hall, London.

ROYALL, R. M (2000). "On the probability of observing misleading statistical evidence", *Journal of the American Statistical Association*, **95**, 760–780.

SAMANIEGO, F. J. AND WHITAKER, L. R. (1986). "On estimating population characteristics from record-breaking observations. I. Parametric results.", *Naval Research Logistics Quarterly*, **33**, 531–543.