JIRSS (2007)
Vol. 6, No. 1, pp 77-87

Testing Exponentiality Based on Record Values

A. Habibi Rad, F. Yousefzadeh, M. Amini, N. R. Arghami

Department of statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Iran. (habibi@math.um.ac.ir, yousefzadeh_2004@yahoo.com, mort.amini@gmail.com, arghami@math.um.ac.ir)

Abstract. We introduce a goodness of fit test for exponentiality based on record values. The critical points and powers for some alternatives are obtained by simulation.

1 Introduction

Suppose a random variable X has cumulative distribution function (cdf) $F(x)$ and a continuous probability density function (pdf) $f(x)$. The entropy $H(f)$ of the random variable X is defined in [10] to be

$$H(f) = -\int_{-\infty}^{\infty} f(x) \log f(x) dx. \tag{1}$$

The Kullback-Leibler (K-L) information of $f(x)$ against $f_0(x)$ is defined in [7] to be

$$I(f; f_0) = \int_{-\infty}^{\infty} f(x) \log \frac{f(x)}{f_0(x)} dx. \tag{2}$$

Key words and phrases: Entropy, estimation, Kullback-Leibler information, maximum likelihood estimator.

Since $I(f; f_0)$ has the property that $I(f; f_0) \geq 0$, and the equality holds only if $f = f_0$, the estimate of the K-L information has also been considered as a goodness of fit test statistic by some authors including [2] and [5]. It has been shown in the aforementioned papers that the test statistics based on the K-L information perform very well for testing exponentiality [5] as compared, in terms of power, with some leading test statistics. In this paper we consider using K-L information for testing exponentiality based on record values.
Some nonparametric estimates of (1) have been proposed in [4], [1] and [12]. In [12], entropy in (1) has been expressed in the form

$$H = \int_0^1 \log\left(\frac{dF^{-1}(p)}{dp}\right) dp. \tag{3}$$

An estimate of (3) can be constructed by replacing the distribution function F by the empirical distribution F_n. The derivative of $F^{-1}(i/n)$ is estimated by $(x_{i+w:n} - x_{i-w:n})n/(2w)$. The estimate of H is then

$$H(w,n) = \frac{1}{n}\sum_{i=1}^{n} \log\left(\frac{n}{2w}(x_{i+w:n} - x_{i-w:n})\right), \tag{4}$$

where the window size w is a positive integer, which is less than $n/2$, and $x_{i:n} = x_{1:n}$ for $i < 1$, and $x_{i:n} = x_{n:n}$ for $i > n$.

Let $X_i,\ i \geq 1$, be a sequence of iid continuous random variables. An observation X_j will be called an upper record value if its value is greater than that of all previous observations. Thus X_j is an upper record value if $X_j > X_i$ for all $i < j$. By convention, X_1 is the first upper record value. There is a similar definition for lower record values by considering the observations that fall below all previous observations.

The times at which upper record values appear are given by the random variables T_j which are called record times and are defined by $T_1 = 1$ and, for $j \geq 2$, $T_j = \mathrm{Min}\{i : \ X_i > X_{T_{j-1}}\}$. The waiting time between the i^{th} upper record value and the $(i+1)^{th}$ upper record value is called the inter-record time (IRT), and is denoted by $\Delta_i = T_{i+1} - T_i,\ i = 1, 2, \ldots$. Record times and inter-record times for lower record values are defined analogously.

Let $U_1, U_2, ..., U_n$ be the first n upper record values from a distribution with the cdf $F(x;\theta)$ and the pdf $f(x;\theta)$, where θ is an unknown parameter (possibly a vector). Then the pdf of the joint distribution of the first n upper record values is given by

$$q(\underline{u};\theta) = \prod_{i=1}^{n-1} \frac{f(u_i;\theta)}{\bar{F}(u_i;\theta)} f(u_n;\theta), \tag{5}$$

where $\bar{F}(x) = 1 - F(x)$. Also the marginal density of U_i (the i^{th} upper record value, $i \geq 1$) is given by

$$q_i(u_i;\theta) = \frac{[-\log(\bar{F}(u_i;\theta))]^{i-1}}{(i-1)!} f(u_i;\theta). \tag{6}$$

The joint distribution of upper record values and their IRT's has density

$$q(\underline{u}, \underline{\Delta};\theta) = \prod_{i=1}^{n} f(u_i;\theta)[F(u_i;\theta)]^{\Delta_i - 1}$$

and the joint density of U_i and Δ_i is

$$q_i(u_i, \Delta_i;\theta) = \frac{[-\log(\bar{F}(u_i;\theta))]^{i-1}}{(i-1)!} f(u_i;\theta)(\bar{F}(u_i;\theta))[F(u_i;\theta)]^{\Delta_i - 1}.$$

See [3] for more details.

The most important use of record values arises in experiments in which a specified characteristic's measurements of a unit are made sequentially and only values that exceed or fall below the current extreme value are recorded. So the only available observations are record values. Such situations often occur in industrial stress, life time experiments, sporting matches, weather data recording and some other experimental fields. The other important application is in life testing problems in which full testing of an item is destructive and costly. If the items are expensive, one can set up the experiment so that only the "low life" units are destroyed. As an example, one may consider the example of testing the breaking strength of wooden beams cited in Glick (1978), in which beams are replaced as soon as the pressure reaches the minimum previously observed breaking pressure. In other words, only the lower record values are observed.

In all the situations mentioned above any statistical inference must be done using record values. In this paper we study the goodness of fit test based on the K-L information using record values.

We introduce a piece-wise linear MLE of cdf based on record values in Section 2. In Section 3 we derive the joint entropy of record values and its estimator, which is used to define our test statistic in section 4. Section 5 contains the critical values and powers against some alternatives obtained by simulation.

2 Maximum likelihood estimation of distribution function based on inter-record times

In this section we use the constrained non-parametric maximum likelihood estimation method, proposed by Yousefzadeh and Arghami, 2007, to estimate the cdf based on upper record values and their inter-record times. The case of lower record values is similar.

Let $f(u_i) = w_i$ and $F(u_i) = \sum_{j=1}^{i} w_j \quad i = 1, \ldots, k$. We maximize the likelihood function subject to $\sum_{i=1}^{k} w_i = 1$. For this purpose, we write the lagrangian as

$$L = \sum_{i=1}^{k} \left(\log w_i + (\Delta_i - 1) \log \sum_{j=1}^{i} w_j \right) - \lambda (\sum_{i=1}^{k} w_i - 1),$$

Solving the equations derived from the above leads us to

$$w_1 = \Delta_1 w_2, \quad w_k = \frac{1}{\sum_{i=1}^{k-1} \Delta_i + 1},$$

and

$$w_{i-1} = \frac{\sum_{t=1}^{i-1} \Delta_t}{\sum_{t=1}^{i-2} \Delta_t + 1} w_i, \quad i = 1, \ldots, k.$$

So a maximum likelihood estimate for $F(u_i)$ is $p_i = \sum_{j=1}^{i} w_j, \quad i = 1, \ldots, k$.

Inter-record times will not be used in the procedure that we propose in section 4.

3 Joint entropy of upper record values and Kullback-Leibler information

The joint entropy of $U_1, U_2, \ldots, U_k$ (the first k upper record values), defined in the literature [9], is

$$H_{1 \cdots k} = - \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{u_2} q(\underline{u}; \theta) \, \log q(\underline{u}; \theta) \, du_1 \cdots du_k, \qquad (7)$$

where $q(\underline{u};\theta)$ is the joint pdf of all k upper record values, which is defined in (5). The following theorem states that the above multiple integral can be simplified to a single integral.

Theorem 3.1.

$$H_{1\dots k} = k - \frac{k(k+1)}{2} - \sum_{i=1}^{k} \int_{-\infty}^{\infty} f(x) \frac{[-\log \bar{F}(x)]^{i-1}}{(i-1)!} \log f(x) dx \quad (8)$$

Proof. By the decomposition property of the entropy measure in [8], we have

$$H_{1\cdots k} = H_1 + H_{2|1} + \cdots + H_{r|r-1} + \cdots + H_{k|k-1}.$$

In [11], another expression of (1) is presented in terms of the hazard function, $h(x) = \frac{f(x)}{\bar{F}(x)}$, as

$$H_1 = 1 - \int_{-\infty}^{\infty} f(x) \log h(x) dx. \quad (9)$$

From (9) the conditional entropy of U_r given $U_{r-1} = u_{r-1}$ is

$$\begin{aligned} H_{r|r-1}(u_{r-1}) &= -\int_{u_{r-1}}^{\infty} \frac{f(x)}{\bar{F}(u_{r-1})} \log \left[\frac{f(x)}{\bar{F}(u_{r-1})} \right] dx \\ &= 1 - \left[\bar{F}(u_{r-1}) \right]^{-1} \int_{u_{r-1}}^{\infty} f(x) \log h(x)\, dx, \end{aligned}$$

Hence

$$\begin{aligned} H_{r|r-1} &= \mathrm{E}\big(H_{r|r-1}(U_{r-1})\big) \\ &= \int_{-\infty}^{\infty} H_{r|r-1}(y) f(y) \frac{[-\log \bar{F}(y)]^{r-2}}{(r-2)!} dy \\ &= 1 - \int_{-\infty}^{\infty} f_{U_r}(x) \log h(x) dx \\ &= 1 - \int_{-\infty}^{\infty} f_{U_r}(x) \log f(x) dx - r. \end{aligned}$$

The required result then follows. □

Using (3) we can write (8) as,

$$H_{1\dots k} = \frac{k(1-k)}{2} + \sum_{i=1}^{k} \int_{0}^{1} \frac{[-\log(1-p)]^{i-1}}{(i-1)!} \log \left(\frac{dF^{-1}(p)}{dp} \right) dp.$$

For the null density function $f_0(x;\theta)$, the K-L information for the first k upper record values is given by

$$I_{1\cdots k}(f:f_0) \quad = \quad \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{u_2} q(\underline{u};\theta)\ \log\frac{q(\underline{u};\theta)}{q_0(\underline{u};\theta)}dx_1\cdots dx_k.$$

So

$$I_{1\cdots k}(f:f_0) = -H_{1\ldots k} - \mathrm{E}(\log q_0(\underline{u};\theta)),$$

For $f_0(x) = \lambda e^{-\lambda x}$, we have

$$I_{1\cdots k}(f:f_0) = -H_{1\ldots k} - k\log\lambda + \lambda\mathrm{E}(U_k).$$

4 Non-parametric information estimate and scale invariant test statistic

In order to obtain a test statistic for testing exponentiality, first we have to derive a nonparametric estimate of the joint entropy of upper record values. This is done by estimating the integral in Theorem 1, which gives the estimator

$$\hat{H}_{1\ldots k} = \frac{k(1-k)}{2} + \sum_{j=1}^{k-1}(p_{j+1}-p_j)\left(\frac{g_j+g_{j+1}}{2}\right),$$

where

$$g_j = \log\left(\frac{U_{j+1}-U_{j-1}}{p_{j+1}-p_{j-1}}\right)\sum_{i=1}^{k}\frac{[-\log(1-p_j)]^{i-1}}{(i-1)!},\quad j=1,\ldots,k-1$$

and

$$p_0 = 0,\ U_0 = U_1,\ U_{k+1} = U_k.$$

The second step is to derive a non-parametric estimator for the mean of the population. Since

$$\lambda = \left[\int_0^1 F^{-1}(p)\ dp\right]^{-1},$$

we have

$$\hat{\lambda} = \left[\sum_{i=1}^{k}(p_{i+1}-p_i)\left(\frac{U_i+U_{i+1}}{2}\right)\right]^{-1}. \tag{10}$$

The non-parametric estimator of the joint K-L information of upper record values is then

$$\hat{I}_{1\ldots k} = -\hat{H}_{1\ldots k} - k\log\hat{\lambda} + \hat{\lambda}U_k. \tag{11}$$

But this statistic is not invariant under the scale group of transformations since

$$\hat{I}^{cX}_{1\ldots k} = \hat{I}^{X}_{1\ldots k} - \hat{I}_k \log c + k\log c,$$

where

$$\hat{I}_k = \sum_{j=1}^{k}(p_{j+1} - p_j)\left(\frac{\Psi_j + \Psi_{j+1}}{2}\right)$$

and

$$\Psi_j = \sum_{i=1}^{k}\frac{[-\log(1-p_j)]^{i-1}}{(i-1)!},\ j = 1, \ldots, k-1,\ \Psi_0 = 0.$$

To get around this problem we replace k in (11) with its equal quantity

$$I_k = \int_0^1 \sum_{i=1}^{k}\frac{[-\log(1-p)]^{i-1}}{(i-1)!}\, dp.$$

Indeed we have

$$I_{1\ldots k} = -H_{1\ldots k} - I_k\log\lambda + \lambda \mathrm{E}(U_k).$$

So

$$\hat{I}_{1\ldots k} = -\hat{H}_{1\ldots k} - \hat{I}_k\log\hat{\lambda} + \hat{\lambda}U_k. \tag{12}$$

Since we have

$$\hat{H}^{cX}_{1\ldots k} = \hat{H}^{X}_{1\ldots k} + \hat{I}_k\log c$$

and consequently

$$\hat{I}^{cX}_{1\ldots k} = \hat{I}^{X}_{1\ldots k},$$

(12) is scale invariant. Therefore

$$T_U = \frac{k(k-1)}{2} - \sum_{j=1}^{k}(p_{j+1} - p_j)\left(\frac{g_j + g_{j+1}}{2}\right) - \hat{I}_k\log\hat{\lambda} + \hat{\lambda}U_k$$

is a scale invariant test statistic for testing exponentiality.

The same procedure can be used to derive the test statistic based on lower records. In that case we have

$$\mathrm{E}(\log f_0(L_1,\ldots,L_k)) = -k\log\lambda + \lambda\sum_{i=1}^{k}\mathrm{E}(L_i) + \sum_{i=1}^{k-1}\mathrm{E}(\log(1-e^{-\lambda L_i})).$$

So a scale invariant test statistic in the case of lower record values is

$$\begin{aligned} T_L &= \frac{k(k-1)}{2} - \sum_{j=1}^{k}(p_{j+1}-p_j)\left(\frac{g_j+g_{j+1}}{2}\right) \\ &= -\hat{I}_k\log\hat{\lambda} + \hat{\lambda}\sum_{i=1}^{k}L_i + \sum_{i=1}^{k-1}\log(1-e^{-\hat{\lambda}L_i}), \end{aligned}$$

where $\hat{\lambda}$ is as in (10) with U_i replaced by L_i and the corresponding ps,

$$g_j = \log\left(\frac{L_{j+1}-L_{j-1}}{p_{j+1}-p_{j-1}}\right)\sum_{i=1}^{k}\frac{(-\log(p_j))^{i-1}}{(i-1)!},\quad j=1,\ldots,k-1$$

and

$$\Psi_j = \sum_{i=1}^{k}\frac{(-\log(p_j))^{i-1}}{(i-1)!},\ j=1,\ldots,k-1,\ \Psi_0=0.$$

5 Critical values and powers of the test

The two test statistics derived in the previous sections are too complicated to allow deriving their exact distributions under the null hypothesis analytically. The critical values of the tests are obtained by a simulation using 10,000 samples and are tabulated for $\alpha = 0.05, 0.1$ and $k = 3, 4, 5$ in tables 1 and 3. Since the record values for larger values of k are rare, we limited k to 5, the usual number of record values in the application, in our simulation study. Another reason for not considering values of k in excess of 5 is that the values of corresponding test statistics become unbounded due to rounding errors. The powers of the test for $\alpha = 0.1$ and different alternatives are also tabulated in tables 2 and 4.

Table 1. Critical values for different values of k and α (upper record values)

	α	
k	0.05	0.1
3	6.19	5.48
4	8.12	7.38
5	11.00	10.18

Table 2. Powers of the test (upper record values)

	Alternative Distribution				
k	$\chi^2(3)$	$\chi^2(5)$	$N(5,1)$	$\Gamma(shape=3)$	Γ(shape=5)
3	0.11	0.12	0.44	0.11	0.14
4	0.12	0.14	0.62	0.15	0.22
5	0.13	0.16	0.64	0.18	0.24

	Alternative Distribution			
k	Weibull(shape=3)	$LN(0,2)$	$\beta(1,2)$	$\beta(2,1)$
3	0.21	0.32	0.12	0.38
4	0.35	0.42	0.14	0.65
5	0.37	0.52	0.17	0.82

Table 3. Critical values for different values of k and α (lower record values)

	α	
k	0.05	0.1
3	8.75	6.66
4	8.94	7.32
5	12.77	9.23

Table 4. Powers of the test (lower record values)

	Alternative Distribution			
k	$\chi^2(3)$	$\chi^2(5)$	$N(5,1)$	$\Gamma(shape=3)$
3	0.12	0.19	0.54	0.20
4	0.18	0.26	0.80	0.33
5	0.20	0.33	0.82	0.38

	Alternative Distribution			
k	Γ(shape=5)	Weibull(shape=3)	$\beta(1,2)$	$\beta(2,1)$
3	0.28	0.29	0.11	0.32
4	0.51	0.48	0.13	0.38
5	0.54	0.51	0.15	0.39

Acknowledgment

The authors would like to thank the refers for their attention to the topic and useful comments and suggestions.

References

[**1**] Ahmad, I. A. and Lin, P. E. (1976), A nonparametric estimation of the entropy for absolutely continuous distributions. IEEE Trans. Inf. Theory, **22**, 372–375.

[**2**] Arizono, I. and Ohta, H. (1989), A test for normality based on Kullback-Leibler information. The American Statistician, **43**, 20–23.

[**3**] Arnold, B. C., Balakrishnan, N., and Nagaraja, H. N. (1998), Records. New York: John Wiley.

[**4**] Dmitriev, Y. G. and Tarasenko, F. P. (1973), On the estimation of functional of the probability density and its derivatives. Theory probab. Applic., **18**, 628–633.

[**5**] Ebrahimi, N. and Habibullah, M. (1992), Testing exponentiality based on Kullback-Leibler information. J. Royal Statist. Soc. B, **54**, 739–748.

[**6**] Glick, N. (1978), Breaking record and breaking boards. Amer. Math. Monthly, **85**, 2–26.

[**7**] Kullback, S. (1959), Information Theory and Statistics. New York: John Wiley.

[**8**] Park, S. (1995), The entropy of consecutive order statistics. IEEE Trans. Inform. Theory, **41**, 2003–2007.

[**9**] Park, S. (2005), Testing exponentiality based on the Kullback-Leibler information with the type II censored data. IEEE Trans. on Rel., **54**, 22–26.

[**10**] Shannon, C. E. (1948), A mathematical theory of communications. Bell System Tech. J., **27**, 379–423.

[**11**] Teitler, S., Rajagopal, A. K., and Ngai, K. L. (1986), Maximum entropy and reliability distributions. IEEE Trans. on Rel., **35**, 391–395.

[**12**] Vasicek, O. (1976), A test for normality based on sample entropy. J. Royal Statist. Soc. B, **38**, 730–737.

[**13**] Yousefzadeh, F. and Arghami, N. R. (2007), A piece-wise linear cdf estimator and some applications. submitted.