Ann. Inst. Statist. Math.
35 (1983), Part A, 267-277

THE FUTURE OCCURRENCE OF RECORDS

IAN R. DUNSMORE

(Received Apr. 26, 1982; revised Dec. 20, 1982)

Summary

The occurrence of future record values based on data from a sequence of independent, identically distributed random variables is considered. Two situations are analysed, namely (i) where only the first m record observations have been noted, and (ii) where all the observations have been noted up to the mth record. Tolerance regions and Bayesian predictive distributions are derived for the increase in size of the $(m+r)$th record value over the observed mth record value for two exponential models. Predictive distributions are also given for the additional number of observations required after the mth record value until the $(m+1)$th record value occurs.

1. Introduction

Suppose that $x_1, x_2, \cdots$ is a sequence of independent, continuous random variables, each with distribution function $F(x)$ and probability density function $p(x)$. If $\{N(n)\}$ is defined by

$$N(n)=\min\{j:\ j>N(n-1),\ x_j>x_{N(n-1)}\}$$

for $n=2, 3, \cdots$, with $N(1)=1$, then $x_{N(1)}, x_{N(2)}, \cdots$ provides a sequence of (maximal) record values. We refer to $N(1), N(2), \cdots$ as the record times. We will be concerned here only with maximal record values but similar results can be derived for minimal record values where we are searching for the smallest values.

Galambos ([7], pp. 309–310) provides a useful survey of the literature on records from the early papers of Chandler [5] and Foster and Stuart [6] on record times through the work, for example, of Renyi [8], Tata [13], Resnick [9], [10] and Shorrock [11], [12] on record values as well. The main results are presented in Galambos ([7], Sections 6.3 and 6.4). Many asymptotic $(n\rightarrow\infty)$ results have been derived. Whilst

Key words: Bayesian prediction intervals; exponential model; predictive distribution; record values; tolerance regions; two parameter exponential model.

these are of interest from the mathematical point of view, they are generally of little value to the applied scientist since $x_{N(n)}$ is rarely observed with large n.

Here we consider two prediction problems associated with the theory of records. Firstly we investigate the prediction of the $(m+r)$th record value on the basis of observing the first m record values $x_{N(1)}, x_{N(2)}, \cdots, x_{N(m)}$ $(r=1, 2, \cdots)$. Ahsanullah ([2], [3]) has discussed such predictions for the exponential and two-parameter exponential distributions and we confine attention to such distributions here. He only suggests point predictions however. In Section 2 we provide tolerance region predictions within the classical framework, whilst in Section 3 we derive predictive distributions within a Bayesian model. In some situations besides a knowledge of the record values we may in fact observe $x_1, x_2, \cdots, x_{N(m)}$, that is, all the values of the underlying series up to and including the mth record. In Section 4 we compare the predictions obtained from such data with those based on the records only.

The second problem, which is discussed in Section 5, involves the prediction of the record time $N(m+1)$ in each of the two situations mentioned earlier. For the case of all the data up to the mth record being available this prediction process proves to be independent of the underlying distribution $F(x)$.

An example to illustrate the predictions obtained is provided in Section 6.

2. Tolerance region predictions: records only

We look first at the problem of predicting the record value $x_{N(m+r)}$ based on the knowledge of the first m record values $x_{N(1)}, x_{N(2)}, \cdots, x_{N(m)}$. Note that we do not assume knowledge of the record times $N(1), N(2), \cdots, N(m)$. We derive mean coverage and guaranteed coverage tolerance regions for $x_{N(m+r)}$.

2.1. *Exponential model*

Suppose that $x_1, x_2, \cdots$ are independently, identically distributed $Ex(\theta)$ random variables, that is, $p(x_i \mid \theta)=\theta \exp(-\theta x_i)$ $(x_i>0;\ \theta>0)$ for $i=1, 2, \cdots$. From a characterization theorem of Tata [13], which Resnick [9] presents more generally, we know that $x_{N(1)}, x_{N(2)}-x_{N(1)}, \cdots, x_{N(m)}-x_{N(m-1)}$ are independent, identically distributed $Ex(\theta)$ random variables. Thus $x_{N(m)}$ is a sufficient statistic for θ based on the data with a $Ga(m, \theta)$ distribution, that is

$$p(x_{N(m)} \mid \theta)=\frac{\theta^m x_{N(m)}^{m-1} \exp(-\theta x_{N(m)})}{\Gamma(m)} \qquad (x_{N(m)}>0)\,.$$

We present predictions for $x_{N(m+r)}$ in terms of the random variable $y = x_{N(m+r)} - x_{N(m)}$ which measures the increase in the $(m+r)$th record over the mth record value. From the same theorem as above we see that y has a $Ga(r, \theta)$ distribution independently of $x_{N(m)}$.

From the framework provided in Aitchison and Dunsmore ([4], Section 5.7)

$$(1) \quad \left\{\frac{1-Be(m, r; c')}{Be(m, r; c')} x_{N(m)}, \infty\right\} \quad \text{and} \quad \left\{0, \frac{1-Be(m, r; 1-c')}{Be(m, r; 1-c')} x_{N(m)}\right\}$$

are both tolerance regions for the increase y with similar mean coverage c'. Here $Be(k, K; c')$ represents the c'th quantile of a $Be(k, K)$ distribution.

Similarly from Aitchison and Dunsmore ([4], Section 6.4) the tolerance regions

$$(2) \quad \left\{\frac{Ga(r, 1; 1-c')}{Ga(m, 1; g')} x_{N(m)}, \infty\right\} \quad \text{and} \quad \left\{0, \frac{Ga(r, 1; c')}{Ga(m, 1; 1-g')} x_{N(m)}\right\}$$

for y provide coverage c' with guarantee g', where $Ga(k, 1; c')$ represents the c'th quantile of a $Ga(k, 1)$ distribution.

2.2. *Two-parameter exponential model*

Suppose now that $x_1, x_2, \cdots$ are independent, identically distributed $Er(\mu, \tau)$ random variables, that is

$$p(x_i \mid \mu, \tau) = \tau \exp\{-\tau(x_i - \mu)\} \qquad (x_i > \mu;\ \tau > 0)$$

for $i = 1, 2 \cdots$. Ahsanullah [3] shows that $x_{N(1)}, x_{N(2)} - x_{N(1)}, \cdots, x_{N(m)} - x_{N(m-1)}$ are again independently distributed random variables, where $x_{N(1)}$ is $Er(\mu, \tau)$ and $x_{N(k)} - x_{N(k-1)}$ is $Ex(\tau)$ $(k = 2, 3, \cdots, m)$. Thus $x_{N(1)}$ and $x_{N(m)} - x_{N(1)}$ are jointly sufficient for $\theta = (\mu, \tau)$ and have independently distributed $Er(\mu, \tau)$ and $Ga(m-1, \tau)$ distributions respectively. If we consider again the increase $y = x_{N(m+r)} - x_{N(m)}$, then y is a $Ga(r, \tau)$ random variable which is independent of $x_{N(1)}$ and $x_{N(m)}$. It follows for example that

$$(3) \quad \begin{aligned} &\left\{\frac{1-Be(m-1, r; c')}{Be(m-1, r; c')} (x_{N(m)} - x_{N(1)}), \infty\right\} \\ \text{and} \quad &\left\{0, \frac{1-Be(m-1, r; 1-c')}{Be(m-1, r; 1-c')} (x_{N(m)} - x_{N(1)})\right\} \end{aligned}$$

are both similar mean coverage tolerance regions for the increase y; whilst

$$\left\{\frac{Ga(r, 1; 1-c')}{Ga(m-1, 1; g')}(x_{N(m)}-x_{N(1)}), \infty\right\} \qquad \text{and}$$

$$\left\{0, \frac{Ga(r, 1; c')}{Ga(m-1, 1; 1-g')}(x_{N(m)}-x_{N(1)})\right\}$$

are both similar (c', g')-guaranteed coverage tolerance regions.

3. Predictive approach

We now provide a Bayesian approach centred on the predictive distribution. Suppose that the data can be summarized by a sufficient statistic x for the unknown parameter θ with an underlying model $p(x|\theta)$. A prior $p(\theta)$ on the parameter space Θ can then be updated to a posterior $p(\theta|x)$ based on the data, and the information about the future value y can be obtained from the predictive density function

$$p(y|x)=\int_{\Theta} p(y|\theta)p(\theta|x)d\theta .$$

3.1. *Exponential model*

We have seen in Section 2.1 that $x=x_{N(m)}$ is sufficient for θ in the exponential model. If we assume that $p(\theta)$ is of conjugate gamma form $Ga(g, h)$, then the predictive distribution for the increase $y=x_{N(m+r)}-x_{N(m)}$ is of the form

$$p(y|x)=\frac{H^G y^{r-1}}{B(r, G)(y+H)^{G+r}} \qquad (y>0) ,$$

which we denote by $InBe(r, G, H)$, with $G=g+m$, $H=h+x_{N(m)}$.

3.2. *Two-parameter exponential model*

In Section 2.2 we saw that $x_{N(1)}$ and $x_{N(m)}-x_{N(1)}$ are jointly sufficient for $\theta=(\mu, \tau)$ in the two-parameter exponential model. We assume a conjugate exponential-gamma prior distribution $ElGa(b, c, g, h)$ for θ, that is

$$p(\mu, \tau)=c\tau \exp\{-c\tau(b-\mu)\}\frac{h^g\tau^{g-1}\exp(-h\tau)}{\Gamma(g)} \qquad (\mu<b, \ \tau>0) .$$

From Aitchison and Dunsmore ([4], Table 2.3) the predictive distribution for the increase $y=x_{N(m+r)}-x_{N(m)}$ is $InBe(r, G, H)$, where

$$G=g+m-1+\delta(c) , \tag{4}$$
$$H=\begin{cases} h+x_{N(m)}-x_{N(1)}-c(x_{N(1)}-b) & \text{if } x_{N(1)}<b , \\ h+x_{N(m)}-x_{N(1)}+\ (x_{N(1)}-b) & \text{if } x_{N(1)}\geqq b , \end{cases}$$

and

$$\delta(c)=\begin{cases} 0 & \text{if } c=0\,, \\ 1 & \text{if } c>0\,. \end{cases}$$

3.3. *Point and interval predictions*

The predictive distribution provides the complete up-to-date picture of our ideas about the increase y. Point and interval predictions can be made to summarize this information.

For example with the two-parameter exponential model point predictions are provided by the mode, $\{(r-1)H\}/(G+1)$, or the mean, $rH/(G-1)$. For vague prior knowledge these correspond to $\{(r-1)\cdot(x_{N(m)}-x_{N(1)})\}/m$ and $\{r(x_{N(m)}-x_{N(1)})\}/(m-2)$ respectively. Ahsanullah [3] suggests point estimates $\{r(x_{N(m)}-x_{N(1)})\}/(m-1)$ and $\{r(x_{N(m)}-x_{N(1)})\}/m$ for the increase, but provides no structure for their errors.

Similarly the intervals

$$\left\{\frac{1-Be(G, r; \kappa)}{Be(G, r; \kappa)}H,\ \infty\right\} \quad \text{and} \quad \left\{0,\ \frac{1-Be(G, r; 1-\kappa)}{Be(G, r; 1-\kappa)}H\right\}$$

provide Bayesian cover κ. For vague prior knowledge these correspond to the mean coverage tolerance predictors (1) or (3) with $\kappa=c'$.

4. Comparison with 'all data' case

In the analysis so far we have assumed that only the record values $x_{N(1)}, x_{N(2)}, \cdots, x_{N(m)}$ up to the mth record are available. Situations can be envisaged in which all the data $x_1, x_2, \cdots, x_{N(m)}$ up to the mth record will be available, and it is of interest to compare the information available for prediction purposes. For the 'all data' case we know that $N(m)=n$, say. Renyi [8] shows that the marginal distribution of $N(m)$ is independent of the underlying $F(\cdot)$ and is given by

$$\mathrm{P}\,(N(m)=n)=\frac{|S_{n-1}^{m-1}|}{n!} \qquad (n=m, m+1, \cdots)\,,$$

where S_n^m are the Stirling numbers of the first kind, defined by

$$x(x-1)\cdots(x-n+1)=\sum_{m\geq 0} S_n^m x^m\,.$$

We see that the joint distribution of the data is specified by

$$\mathrm{P}\,(N(m)=n)\prod_{i=1}^{n} p(x_i\,|\,\theta)\,. \tag{5}$$

4.1. *Tolerance regions*

The tolerance regions are now less straightforward to obtain. For example, consider a mean coverage tolerance region for the increase y. For the exponential model a reasonable form would be $\left\{0, (k-1)\sum_{i=1}^{N(m)} x_i\right\}$, whilst for the two-parameter case $\left\{0, (k-1)\left(\sum_{i=1}^{N(m)} x_i - N(m)x_{(1)}\right)\right\}$ is more appropriate. Here $x_{(1)}=\min(x_1, x_2, \cdots, x_{N(m)})$. For mean coverage c' we require the solution k of

$$\sum_{n=m}^{\infty} \mathrm{P}(N(m)=n)\int_{1/k}^{1} \frac{s^{\nu(n)-1}(1-s)^{r-1}}{B(\nu(n), r)}\,ds = c' , \tag{6}$$

where $\nu(n)=n$ for the exponential model and $(n-1)$ for the two-parameter case. No analytical solution readily presents itself. For the case $r=1$ of predicting the increase to the next record we set

$$\sum_{n=m}^{\infty} \mathrm{P}(N(m)=n)\left(\frac{1}{k}\right)^{\nu(n)} = 1-c' . \tag{7}$$

For regions of the appropriate form $(*, \infty)$ we simply replace c' by $1-c'$ in (6) or (7).

Similarly for a (c', g') guaranteed coverage tolerance region for the increase y of the same forms as above we select $q=k-1$ such that

$$\sum_{n=m}^{\infty} \mathrm{P}(N(m)=n)\int_{Ga(r,1;\,c')/q}^{\infty} \frac{v^{\nu(n)-1}e^{-v}}{\Gamma(\nu(n))}\,dv = g' .$$

Chandler [5] provides some tables for the distribution of $N(m)$. For computational purposes the evaluation of $\mathrm{P}(N(m)=n)$ can perhaps best be achieved by use of a recurrence relation connecting Stirling's numbers of the first kind as given by Abramowitz and Stegun ([1], p. 824, II.A). If we write

$$Q(m, n) = \frac{S_{n-1}^{m-1}}{n!} \qquad (n \geqq m \geqq 1) ,$$

so that $\mathrm{P}(N(m)=n)=|Q(m, n)|$ $(n=m, m+1, \cdots)$, we have that

$$Q(m, n) = \sum_{k=m-2}^{n-2} \frac{(k+1)(-1)^{n-k-2}}{(n-k-1)(m-1)n} Q(m-1, k+1) \qquad (n \geqq m \geqq 2) .$$

With starting values

$$Q(1,n) = \begin{cases} 1 & \text{if } n=1 , \\ 0 & \text{if } n=2, 3, \cdots , \end{cases}$$

we are able to derive $\mathrm{P}(N(m)=n)$ easily.

4.2. *Predictive distributions*

The predictive distributions can be derived in a straightforward manner since from (5) $\mathrm{P}(N(m)=n)$ plays no part in the updating of $p(\theta)$ to $p(\theta\,|\,\text{data})$.

Thus for the two exponential models the predictive distribution for the increase $y=x_{N(m+r)}-x_{N(m)}$ is $InBe(r, G', H')$, where $G'=g+n$, $H'=h+\sum_{i=1}^{n} x_i$ for the exponential model, and

$$G'=g+n-1+\delta(c)\ ,$$

$$H'=\begin{cases} h+\left(\sum_{i=1}^{n} x_i-nx_{(1)}\right)-c(x_{(1)}-b) & \text{if } x_{(1)}<b\ , \\ h+\left(\sum_{i=1}^{n} x_i-nx_{(1)}\right)+n(x_{(1)}-b) & \text{if } x_{(1)}\geqq b\ , \end{cases}$$

for the two-parameter model. For vague prior knowledge we therefore have the following predictive distributions for the increase y:

	Exponential	Two-parameter exponential
Records only:	$InBe(r, m, x_{N(m)})$	$InBe(r, m-1, x_{N(m)}-x_{N(1)})$
All data:	$InBe\left(r, N(m), \sum_{i=1}^{N(m)} x_i\right)$	$InBe\left(r, N(m)-1, \sum_{i=1}^{N(m)} x_i-N(m)x_{(1)}\right)$

5. Prediction of record times

We consider now the prediction of future record times and confine attention to $N(m+1)$, that is, the next record. Again we investigate the two situations in which (i) only the first m records have been observed, and (ii) all the observations up to the mth record are available.

5.1. *Records only*

Suppose that we have observed the data $\boldsymbol{x}_R=(x_{N(1)}, x_{N(2)},\cdots, x_{N(m)})$ from an underlying sequence of identically distributed random variables, each with distribution function F. We do not know the value of $N(m)$ and so provide a predictive distribution for the additional number of observations to the next record, that is, for $n_m=N(m+1)-N(m)$. Shorrock [11] showed that

$$\mathrm{P}(n_m>k\,|\,\tau_1, \tau_2,\cdots, \tau_m)=(1-e^{-\tau_m})^k\ ,$$

where $\tau_j=-\log\{1-F(x_{N(j)})\}$ $(j=1, 2,\cdots, m)$. It follows that

$$\mathrm{P}(n_m=k\,|\,\tau_1, \tau_2,\cdots, \tau_m)=\{F(x_{N(m)})\}^{k-1}\{1-F(x_{N(m)})\}\ . \tag{8}$$

Consider first the exponential model with $F(x)=1-\exp(-\theta x)$. Expression (8) is a function of θ and $\boldsymbol{x}_R$. Hence, since the conjugate $p(\theta\,|\,\boldsymbol{x}_R)$ is $Ga(G=g+m,\ H=h+x_{N(m)})$, we have the predictive function

$$\mathrm{P}\,(n_m=k\,|\,\text{records})=\sum_{i=0}^{k-1}\binom{k-1}{i}(-1)^i\frac{H^G}{(H+(i+1)x_{N(m)})^G} \qquad (k=1,2,\cdots)\,. \tag{9}$$

Note that for a vague prior with $g\to 0$, $h\to 0$ the predictive function (9) is independent of $x_{N(1)}, x_{N(2)},\cdots, x_{N(m)}$ and given by

$$\mathrm{P}\,(n_m=k\,|\,\text{records})=\sum_{i=0}^{k-1}\binom{k-1}{i}(-1)^i(i+2)^{-m}\,. \tag{10}$$

This form is in fact independent of the parent populatian and is identical to that derived by Chandler [5] for the additional number to the next record.

Similarly for the two-parameter exponential model with $F(x)=1-\exp\{-\tau(x-\mu)\}$ and with a conjugate $ElGa(b, c, g, h)$ prior for (μ, τ), the predictive function $\mathrm{P}\,(n_m=k\,|\,\text{records})$ is given by

$$\sum_{i=0}^{k-1}\binom{k-1}{i}(-1)^i\frac{C}{(C+i+1)}\,\frac{H^G}{\{H+(i+1)(x_{N(m)}-B)\}^G} \qquad (k=1,2,\cdots)\,, \tag{11}$$

where $B=\min(b, x_{N(1)})$, $C=c+1$, and G and H are as defined in (4). Again the predictive function (11) is independent of $x_{N(1)}, x_{N(2)},\cdots, x_{N(m)}$ for the case of a vague prior, and is indeed of the same form as (10) above.

5.2. *All data to mth record*

Here we observe $x_1, x_2,\cdots, x_{N(m)}$ and so know the value of $N(m)$. We require the predictive function $\mathrm{P}\,\{N(m+1)=j\,|\,N(m)=n,\ x_1, x_2,\cdots, x_{N(m)}\}$. Simple conditional arguments reveal that this is identical to $\mathrm{P}\,\{N(m+1)=j\,|\,N(m)=n\}$, irrespective both of $x_1, x_2,\cdots, x_{N(m)}$ and of the form of the underlying distribution function F. It follows therefore, as for example in Galambos ([7], p. 292), that

$$\mathrm{P}\,\{N(m+1)=j\,|\,N(m)=n,\ x_1, x_2,\cdots, x_{N(m)}\}=\begin{cases}\dfrac{n}{j(j-1)} & \text{if } j>n\geqq m\geqq 1\,,\\[2mm] 0 & \text{otherwise}\,,\end{cases}$$

for both the models considered here. Equivalently

$$\mathrm{P}\,(n_m=k\,|\,N(m)=n,\ x_1, x_2,\cdots, x_{N(m)})=\frac{n}{(n+k)(n+k-1)} \qquad (k=1,2,\cdots)$$

(12) or

$$P(n_m \leqq k \mid N(m)=n,\ x_1, x_2, \cdots, x_{N(m)}) = \frac{k}{(n+k)} \qquad (k=1, 2, \cdots)$$

describe the distribution attached to the additional number of observations n_m from the mth to the $(m+1)$th record. From (12) a prediction interval for n_m of the form $(0, K)$ with cover κ is readily obtained by setting

$$K = \frac{n\kappa}{1-\kappa}\ . \tag{13}$$

6. Example

A rock crushing machine has to be reset if, at any operation, the size of rock being crushed is larger than any that has been crushed before. Suppose for illustration that the sizes (in suitable units) of rocks to be crushed can be represented by independent $Ex(\theta)$ random variables. The data below are the sizes dealt with up to the third time that the machine has been reset.

$$9.3,\ 0.6,\ 24.4,\ 18.1,\ 6.6,\ 9.0,\ 14.3,\ 6.6,\ 13.0,\ 2.4,\ 5.6,\ 33.8\,.$$

If only the sizes at the operations when resetting was necessary had been observed we would have simply noted the record values

$$9.3, \quad 24.4, \quad 33.8\,.$$

Here then for the record data analysis we have $m=3$, $x_{N(3)}=33.8$, whilst for the full data $m=3$, $n=N(3)=12$ and $\sum_{i=1}^{12} x_i = 143.7$.

Our interest lies in when the machine will next need to be reset and the size of rock which will necessitate the resetting.

Consider first the analysis based on the records only. From (1) with $r=1$ we see that $(0, 57.9)$ provides a 95% mean coverage tolerance region for the increase in size, so that $(33.8, 91.7)$ provides an interval for the size of the rock. Similarly from (2) the larger interval $(0, 123.8)$ satisfies the more stringent requirements of a guaranteed coverage tolerance region for the increase with cover 0.95 and guarantee 0.95.

The predictive distribution for the increase in size from 33.8 to the next resetting is of the form $InBe(1, 3, 33.8)$ for a vague prior. This yields a 95% Bayesian prediction interval of the form $(0, 57.9)$, as expected.

For the full data we find from (7) that $(0, 113.6)$ provides a 95% mean coverage tolerance region for the increase in size; whilst the Bayesian predictive distribution is of the form $InBe(1, 12, 143.7)$, which yields a 95% Bayesian prediction interval of the form (0, 40.7). Here these two intervals are not the same, since the cover of the Bayesian prediction interval of the form $\left(0, (q-1) \sum_{i=1}^{N(m)} x_i\right)$ is no longer a constant but depends on the random variable $N(m)$. Notice also that in any particular application we do not necessarily obtain a shorter mean coverage tolerance region when we have the full data available. On average however over repeated applications we would expect a better prediction for the full data case.

For predicting the additional number of operations from $N(3)=12$ to when the next resetting will be required we can use (9) and (12) for the two situations. Table 1 provides the critical values K for prediction intervals of the form $(0, K)$ for various values of the cover κ. In the records only case we have assumed vague prior knowledge in (9). The evaluation of the critical values in that case is most easily derived from the fact that

$$\mathrm{P}\,(n_3 \geqq k) = \frac{1}{k} \sum_{t=1}^{k} \frac{1}{t} \sum_{i=1}^{t} \frac{1}{i}\ ,$$

a result given in Chandler [5]. As might be expected, shorter intervals result in the case in which all the observations are available.

Table 1. Critical values K for prediction intervals of the form $(0, K)$ for various values of the cover κ

Cover κ :	0.50	0.60	0.70	0.80	0.85	0.90	0.95	0.99
Records only :	10	16	28	58	92	172	469	4022
All data :	12	18	28	48	68	108	228	1188

I would like to acknowledge gratefully some assistance from David J. Robson, University of Sheffield, with some of the computational problems encountered in Section 6.

UNIVERSITY OF SHEFFIELD

REFERENCES

[1] Abramowitz, M. and Stegun, I. (1972). *Handbook of Mathematical Functions*, Dover, New York.

[2] Ahsanullah, M. (1978). Record values and the exponential distribution, *Ann. Inst. Statist. Math.*, **30**, 429-433.

[3] Ahsanullah, M. (1980). Linear prediction of record values for the two parameter exponential distribution, *Ann. Inst. Statist. Math.*, **32**, 363-368.

[4] Aitchison, J. and Dunsmore, I. R. (1975). *Statistical Prediction Analysis*, Cambridge University Press, Cambridge.

[5] Chandler, K. N. (1952). The distribution and frequency of record values, *J. R. Statist. Soc.*, B, **14**, 220-228.

[6] Foster, F. C. and Stuart, A. (1954). Distribution-free tests in time-series based on the breaking of records, *J. R. Statist. Soc.*, B, **16**, 1-22.

[7] Galambos, J. (1978). *The Asymptotic Theory of Extreme Order Statistics*, Wiley, New York.

[8] Renyi, A. (1962). Théorie des éléments saillants d'une suite d'observations, *Colloq. Combinatorial Meth. Prob. Theory*, Aarhus University, 104-115.

[9] Resnick, S. I. (1973a). Limit laws for record values, *J. Stochastic Processes and their Appl.*, **1**, 67-82.

[10] Resnick, S. I. (1973b). Record values and maxima, *Ann. Prob.*, **1**, 650-662.

[11] Shorrock, R. W. (1972a). A limit theorem for inter-record times, *J. Appl. Prob.*, **9**, 219-223.

[12] Shorrock, R. W. (1972b). On record values and record times, *J. Appl. Prob.*, **9**, 316-326.

[13] Tata, M. N. (1969). On outstanding values in a sequence of random variables, *Zeit. Wahrscheinlichkeitsth.*, **12**, 9-20.