Yugoslav Journal of Operations Research
9 (1999), Number 1, 141-146

HOW TO DISTINGUISH BETWEEN RECORD TYPES*

Slobodanka JANKOVIĆ

Mathematical Institute
Kneza Mihaila 35, 11001 Belgrade, Yugoslavia

Abstract: Let G be a probability distribution having one of three possible record types. Suppose the values of distribution G on set S are known. We give the criterion to determine the record type of G from its values on set S and also the minimal cardinality that S must have in order that there is only one distribution of record type taking given values on S.

Keywords: Records, record moments, distribution type, normal distribution.

1. INTRODUCTION

Mathematical theory of records is a part of the extreme value theory and it has an enormous variety of applications. In various situations only the record-breaking values are observed or are interesting. Such data arise in life testing and stress testing and in industrial quality-control experiments, also in meteorology, climatology, atmospheric sciences as well as in sport. A common feature of these situations is that decisions are made on the basis of the record-breaking observations. For instance when constructing the bridge on the river the record water height is always taken into account. Methodology drawn from the fields of probability and statistics plays an important role in studying (predicting, forecasting, modeling) the records.

Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed random variables having continuous probability distribution F. Denote by M_n the maximum of the first n members of the above sequence, $M_n = \max_{1 \le i \le n} X_i$. Records occur at those values of the index for which the value of the maximum changes. These indices, called record moments, are defined in the following way:

---

* AMS (1991) Classification: 60G70, 60E05. Research supported by the Science Fund of Serbia through the Mathematical Institute of SANU, grant number 04M03/G.

$$L(1) = 1\,,$$

and, for $j \geq 1$,

$$L(j+1) = \min\{i > L(j) : X_i > X_{L(j)}\}\,.$$

The values $\{M_{L(j)}, j \geq 1\} = \{X_{L(j)}, j \geq 1\}$ are called record values, each of them being bigger than the preceding one. Among the first problems studied in connection with records were how often the records occur and what is the limit behavior of records. Mathematical theory of records has begun with the work of Chandler [1] and some important early results are those of Renyi [6], Tata [12], Shorrock [11] and Resnick [8]. Various problems concerning records have been investigated by many authors and these references can be found in monographs [3] and [8] and in the survey [4], see also [2] and [5].

In the sequel we shall need the following well known theorem determining the class of limit distributions for records.

Theorem 1. *Suppose there exist constants* $\alpha_n > 0$, $\beta_n \in R$ *and a non-degenerate distribution G such that*

$$P\{(X_{L(n)} - \beta_n)/\alpha_n \leq x\} \to G(x)\,,$$

holds as $n \to \infty$ *in every continuity point of G. Then the class of limit distributions G consists of the following three types*

(i) $N(x)$

(ii) $N_{1,\alpha}(x) = \begin{cases} 0 & x < 0 \\ N(\alpha \log x) & x \geq 0 \end{cases}$

(iii) $N_{2,\alpha}(x) = \begin{cases} N(-\alpha \log(-x)) & x < 0 \\ 1 & x \geq 0 \end{cases}$

where $\alpha > 0$ *and* $N(x)$ *is the standard normal distribution.*

The distribution type is defined in the usual way: Two distributions F and G are of the same type if they can be obtained from each other by linear transformation of the argument, i.e. if there exist $a > 0$ and $b \in R$ such that $F(x) = G(ax + b)$.

Some twenty years ago a new condition in the Central Limit Theorem appeared - the condition of restricted convergence. The condition was that the convergence of the sequence of sums of independent and identically distributed random variables was supposed to hold on the half line. From that condition it followed that the Central Limit Theorem holds, i.e. the sequence of sums converges on the whole real line to the random variable having normal distribution. The two main problems considered in that theory were the compactness problem and the continuation problem.

The continuation problem deals with the uniqueness of the continuation of a probability distribution from the set $S \subset R$ to the whole real line, within the given class of distributions. These problems were intensively studied by German mathematicians and the results concerning the sums of random variables are collected in monograph [9] and survey paper [10].

In this article we investigate the problem of the uniqueness of the continuation of probability distributions belonging to the class of limit distributions for records, from some bounded set S to the whole real line. We shall determine the smallest cardinality of set S, so that probability distributions belonging to the class of limit distributions for records can be continued in a unique way from set S to the whole real line. As will be seen record laws are completely determined by their values in three points and it is possible to characterize the type of the record law by the value of certain expression depending on the values of the record law in three arbitrary points. This means that three quantiles completely determine the record distribution. This can be applied to construct a statistical test for estimating the type of the record law.

2. CHARACTERIZATION OF THE RECORD TYPES

The results can be summarized in the following two theorems:

Theorem 2. *Two different probability distributions from the class of limit distributions for records can coincide in at most two points in which these distributions take values different from* 0 *and* 1.

From Theorem 2 we can conclude that, given three pairs of numbers (x_i, y_i), $i = 1, 2, 3$, where $x_1 < x_2 < x_3$ and $0 < y_1 < y_2 < y_3 < 1$, there exists at most one probability distribution belonging to the class of limit distributions for records which at given points x_i takes values $G(x_i) = y_i$, $i = 1, 2, 3$. Very important questions in connection with this are: Given three pairs of numbers (x_i, y_i), $i = 1, 2, 3$, where $x_1 < x_2 < x_3$ and $0 < y_1 < y_2 < y_3 < 1$, under what conditions does there exist a distribution from the class of limit distributions for records, which takes the given values at the given points? And if such distribution exists, to which of the three possible types (i), (ii), (iii) does it belong? The following theorem gives answers to these questions.

Theorem 3. *Suppose three pairs of real numbers (x_i, y_i), $i = 1, 2, 3$ are given, satisfying $x_1 < x_2 < x_3$ and $0 < y_1 < y_2 < y_3 < 1$. Then there exists exactly one probability distribution G belonging to the class of limit distributions for records, which satisfies $G(x_i) = y_i$, $i = 1, 2, 3$. Distribution G is of the types (i), (ii), or (iii) if the value of the expression*

$$(x_3 - x_2)N^{-1}(y_1) + (x_1 - x_3)N^{-1}(y_2) + (x_2 - x_1)N^{-1}(y_3) \tag{1}$$

is, respectively, zero, negative or positive (where N^{-1} is the inverse of the function N).

Proof of Theorem 2: There are six cases to consider: the coincidence of distributions of the same type and the coincidence of distributions of different types. We shall consider two of these cases, the others being similar.

Suppose two distributions of types (ii) and (iii) coincide on some set $S \subset R$, i.e.

$$N_{1,\alpha}(ax+b) = N_{2,\beta}(cx+d),$$

$ax+b>0,\ cx+d<0,\ x \in S$, where $\alpha, \beta, a, c > 0,\ b, d \in R$. Since function N is strictly monotone and continuous, it follows that

$$\alpha \log(ax+b) = -\beta \log(-(cx+d))$$

i.e.

$$ax+b = (-(cx+d))^{-\beta/\alpha}.$$

On the lefthand side we have an affine function and on the righthand side a convex function. These two functions can coincide in at most two points.

Suppose now that two distributions of type (iii), not identically equal to each other, coincide on some set $S \subset R$, i.e.

$$N_{2,\alpha}(ax+b) = N_{2,\beta}(cx+d),$$

where $ax+b<0,\ cx+d<0,\ x \in S,\ \alpha, \beta, a, c > 0,\ b, d \in R$. It follows that

$$-\alpha \log(-(ax+b)) = -\beta \log(-(cx+d)),$$

whence

$$-(ax+b) = (-(cx+d))^{\beta/\alpha}.$$

Analogously to the previous case we deduce that these two functions can coincide in at most two points, i. e. set S consists of two points.

Proof of Theorem 3: We shall consider the case when expression (1) is negative and we shall prove that in this case there exists a probability distribution of type (ii) which takes the given values at the given points. The uniqueness of this distribution follows from Theorem 2.

We have to prove that there exists a triple $(\alpha, A, B),\ \alpha, A > 0,\ B \in R$ such that

$$N(\alpha \log(Ax_i + B)) = y_i, \qquad i = 1,2,3,$$

or equivalently, that $\alpha \log(Ax_i + B)) = N^{-1}(y_i)$, $i = 1,2,3$ holds. We have

$$Ax_i + B = \exp(\alpha^{-1} N^{-1}(y_i)), \qquad i = 1,2,3. \tag{2}$$

Put $N^{-1}(y_i) = z_i$, $i = 1,2,3$. Then (2) is a system of three equations with two unknown variables A and B, both depending on the parameter α. The matrix of that system is

$$\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \end{bmatrix}$$

and its rank equals 2. In order for the system above to have a solution it is necessary and sufficient that the rank of the augmented matrix is also equal to 2, i.e. that

$$\det \begin{bmatrix} x_1 & 1 & \exp(\alpha^{-1} z_1) \\ x_2 & 1 & \exp(\alpha^{-1} z_2) \\ x_3 & 1 & \exp(\alpha^{-1} z_3) \end{bmatrix} = 0,$$

which is equivalent to

$$(x_3 - x_2)\exp(\alpha^{-1} z_1) + (x_1 - x_3)\exp(\alpha^{-1} z_2) + (x_2 - x_1)\exp(\alpha^{-1} z_3) = 0.$$

We have to find out whether there exists some $\alpha > 0$ which is the solution of the preceding equation. That equation has the solution $\alpha > 0$ if and only if the function

$$f(t) = (x_3 - x_2)\exp(tz_1) + (x_1 - x_3)\exp(tz_2) + (x_2 - x_1)\exp(tz_3)$$

has one positive zero. The function f and the function

$$g(t) = f(t)\exp(-tz_2) = (x_3 - x_2)\exp(t(z_1 - z_2)) + (x_1 - x_3) + (x_2 - x_1)\exp(t(z_3 - z_2))$$

have the same set of zeros. The properties of function g are:

1) $g(0) = 0$;
2) g is convex;
3) $g(+\infty) = +\infty$, $g(-\infty) = +\infty$;
4) $g'(0) = (x_3 - x_2)z_1 + (x_1 - x_3)z_2 + (x_2 - x_1)z_3$.

Obviously, when $g'(0)$ is negative, then g will have two zeros: the other zero will be positive, which means that α is positive.

There remains to prove that $A > 0$. From (2) it follows that

$$Ax_1 + B = \exp(\alpha^{-1} z_1)$$

$$Ax_2 + B = \exp(\alpha^{-1} z_2)$$

If we subtract the first equation from the second, we obtain:

$$A(x_2 - x_1) = \exp(\alpha^{-1} z_2) - \exp(\alpha^{-1} z_1),$$

and, since $x_2 > x_1$ the righthand side is positive whence A is positive too. The remaining two cases are obtained in an analogous way. The proof is completed.

REFERENCES

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

[2] Deheuvels, P., Nevzorov, V. B., "Limit laws for *k*-record times", *J. Stat. Plann. Inference*, 38 (3) (1994) 279-307.

[3] Galambos, J., *The Asymptotic Theory of Extreme Order Statistics*, John Wiley & Sons, NY, 1978.

[4] Невзоров, В.Б., "Рекорды", *Теория Вероятностей и её Применения*, 32(2) (1987) 219-251.

[5] Nevzorov, V., "Asymptotic distributions of records in nonstationary schemes", *J. Stat. Plann. Inference*, 45(1-2) (1995) 261-273.

[6] Rényi, A., "On outstanding values of a sequence of observations", in: *Selected Papers of A. Rényi, vol 3.*, Akadémia Kiado, Budapest, 1976, 50-65.

[7] Resnick, S., "Limit laws for record values", *Stochastic Processes Appl.*, 1 (1973) 67-82.

[8] Resnick, S., *Extreme Values, Regular Variations and Point Processes*, Springer-Verlag, NY, 1987.

[9] Rossberg, H. -J., Jesiak, B., and Siegel, B., *Analytic Methods in Probability Theory*, Akademie-Verlag, Berlin, 1985.

[10] Россберг, Х. -Й., "Предельные теоремы, включающие ограниченную сходимость", *Теория Вероятностей и её Применения*, 39(4) (1994) 334-356.

[11] Shorrock, R.W., "On record values and record times", *J. Appl. Probability*, 9 (1972) 316-326.

[12] Tata, M. N., "On outstanding values in a sequence of random variables", *Zeitshrift fur Wahrschein. verw. Gebiete*, 12 (1969) 9-20.