Nonlinear Analysis: Modelling and Control, Vilnius, IMI, 1998, No 2

EXTREME VALUE STOCHASTIC MODELS

Algimantas Aksomaitis
Kaunas University of Technology, Studentų 50, 3028 Kaunas, Lithuania

Abstract

The investigation of extreme values (maxima and minima) and their asymptotics is presented. Attention is paid to the classical theory (independent observations of identical distributions). The generalizations and applications of the theory are discussed and problems are revealed.

1. INTRODUCTION

The problems related with extreme values from some collection of random variables are met in various fields of science and applications. In probability theory, there are not so many formulas both exact and convenient for calculation, if stochastic model involves a great number of random factors. Therefore the approximation of distributions by suitable functions is an urgent problem. This problem is characteristic for stochastic extremes, as well and is solved by asymptotic theory of extremes. There are important theoretical and applicational cases where extremes are formed using a collection of random number of random variables. Such a situation, for instance, arises in investigation of the record duration and record properties.

Extreme value problem (maxima and minima)

$$Z_n = \bigvee_{j=1}^{n} , \qquad W_n = \bigvee_{j=1}^{n} X_j$$

from some collection of random variables $X = (X_1, \ldots, X_n)$ can serve as exact or at least approximate model for great number of real objects. Notice, that nonlinear semigroup operations $\vee$ or $\wedge$ (there is no inverse element) are commutative and associative. Since $W_n = -\bigvee_{j=1}^{n}\left(-X_j\right)$, attention will be paid to investigation of the structure Z_n .

The age of classical extreme theory (the simple random samples) exceeds 75 years. Development of the theory was influenced significantly by works of von Mises (1923-1936), Frechet (1927), Fisher and Tippet (1928), Gumbel (since 1935), Gnedenko (since 1941). Exaustive representation of this theory is given in monographs [1], [5] and [6].

Requirements of classical theory (random variables X_j are independent and their distributions are identical) in applications are often not fulfiled. Therefore the generalizations of the theory are needed. A great deal of the results, obtained in this direction, are represented in the monographs mentioned above and also in the review paper [2] and its plentiful reference list.

Though a new results are presented in this paper as well, the main aim of it is both a brief review and applying possibilities of extreme theory.

2. EXACT EXTREME MODELS

Let $X = (X_1, X_2, \ldots, X_n)$ be a simple random sample from the distribution F. A volume n of the sample can be a random number as well (then we will denote it as N).

The distribution functions of the extremes Z_n and W_n are

$$P(Z_n < x) = F^n(x), \quad P(W_n < x) = 1 - (1 - F(x))^n. \tag{1}$$

In case volume N is random number and does not depend on any of X_j, $j \geq 1$, we get

$$P(Z_n < x) = F^n(x), \quad P(W_n < x) = 1 - (1 - F(x))^n. \tag{2}$$

Here

$$g_n(z) = \sum_j z^j P(N = j), |z| \leq 1$$

is a generating function of the random variable N.

The general distribution function is given by

$$P(Z_n < x, W_n < y) = \begin{cases} F^n(x), & y \geq x, \\ F^n(x) - (F(x) - F(y))^n, & y < x. \end{cases} \tag{3}$$

Let we have one more simple sample $Y = (Y_{1,}, Y_2, \ldots, Y_n)$ that does not depend on X. If $P(Y_j < y) = G(y)$ for all $j \geq 1$, then

$$\begin{aligned} &P\left(\overset{n}{\underset{j=1}{\vee}} X_j < x, \overset{n}{\underset{j=1}{\vee}} Y_j < y\right) = (F(x)G(x))^n, \\ &P\left(\overset{n}{\underset{j=1}{\vee}} X_j < x, \overset{n}{\underset{j=1}{\wedge}} Y_j < y\right) = F^n(x)\left(1 - (1 - G(y))^n\right). \end{aligned} \tag{4}$$

In case the variables X_j are independent with $F_j(x) = P(X_j < x)$, we get

$$P(Z_n < x) = \prod_{j=1}^{n} F_j(x), P(W_n < x) = 1 - \prod_{j=1}^{n} (1 - F_j(x)). \tag{5}$$

In case of dependent variables X_j, $j \geq 1$, we have

$$P(Z_n < x) = \prod_{j=1}^{n} F_j(x). \tag{6}$$

Here

$$F_1(x) = P(X_1 < x), F_j(x) = P\left(X_j < x \middle| X_1 < x, \ldots, X_{j-1<x}\right), \; j > 1.$$

If, for instance, variables X_j, $j \geq 1$ make a stacionary Markov sequence, then

$$P(Z_n < x) = F_1(x) F_2^{n-1}(x), \tag{7}$$

here

$$F_1(x) = P(X_1 < x), F_2(x) = P(X_2 < x | X_1 < x).$$

3. CLASSICAL ASYMPTOTICAL PROBLEM

Calculation by (1)–(7) formulas is rather complicated, if volume n is large. Therefore in both theoretical and practical aspects it is urgent to find the limit (as $n \to \infty$) of distributions of these statistics.

The problem is as follows: what conditions must the distribution function F satisfy for

$$P\left(\frac{Z_n - a_n}{b_n} < x\right) \Rightarrow H(x), \tag{8}$$

as $n \to \infty$? Here the sign "$\Rightarrow$" means weak convergence, and $H(x)$ is non–singular distribution function. Complete solution of this problem was published in the fundamental work of Gnedenko [3]. Here the necessary and sufficient conditions of convergence are formulated and algorithm for calculation of the normalizing constants a_n and $b_n > 0$ is given. The limit distribution function can only be of the following type

$$H(x) = H_c(x) = \exp\left\{-(1+cx)^{-1/c}\right\}, \quad 1+cx > 0, x \in \boldsymbol{R}. \tag{9}$$

Noticé, that there are a few functions F that does not belong to the attraction domain of H_c. For example,

$$F(x) = 1 - \frac{1}{\ln x}, x \ge e .$$

But in the case of nonlinear normalization we get

$$P\left(\frac{\ln Z_n}{n} < x\right) \Rightarrow e^{-e^{-x}}, x \in \boldsymbol{R},$$

i.e. the limit distribution is $H_0(x)$.

Due to the mentioned work of Gnedenko [3] there arose a new problem of estimation of the rate of convergence of (8):

- non-uniform estimation: $\left|P(Z_n < xb_n + a_n) - H_c(x)\right| \le \Delta_n(x)$,
- uniform estimation: $\left|P(Z_n < xb_n + a_n) - H_c(x)\right| \le \Delta_n(x)$.

Various forms of estimates of the rate of convergence are presented in the monographs [1] and [5].

Gnedenko in 1984 published an extension of the classical scheme in case the volume $N = N_0$ of the simple sample is random and does not depend on X_j, $j \ge 1$ [4]. Here we give the formulation of this transference theorem: let (8) holds and

$$\lim_{n\to\infty} P\left(\frac{N_n}{n} < x\right) = A(x).$$

Then

$$\lim_{n\to\infty} P\left(\frac{Z_{N_n} - a_n}{b_n} < x\right) = \psi(x) = \int_0^{\infty} H^z(x) dA(z).$$

Notice that one can find such variables X_j and N_n for which direct calculation of the distribution function $P\left(Z_{N_n} < xb_n + a_n\right)$ with respect to the rate of convergence gives more precise results than the transference theorem. This will take place if

$$F(x) = \frac{1}{1+e^{-x}},\ x \in \boldsymbol{R},$$

$$P(N_n = j) = \frac{1}{n}\left(1 - \frac{1}{n}\right)^{j-1},\ j \geq 1.$$

Non-uniform estimate of the rate of convergence

$$\left|P\left(Z_{N_n} < xb_n + a_n\right) - \psi(x)\right| \leq \Delta_{N_n}(x),$$

was obtained in 1987 [8].

4. EXTENSION OF THE CLASSICAL ASIMPTOTICAL PROBLEM

Neglecting the limitations of the classical scheme, one can get different generalizations:

- diversely distributed variables,
- nonlinear normalization,
- random vectors,
- stochastic processes.

Let us discuss briefly an asymptotical problem of a diversely distributed random variables.

Suppose, that for any x

$$\lim_{n\to\infty} \max P\left(X_j \geq xb_n + a_n,\ j = \overline{1,n}\right) = 0.$$

Then the equality

$$\lim_{n\to\infty} \sum_{j=1}^{n} P\left(X_j \geq xb_n + a_n\right) = u(x)$$

comprise the necessary and sufficient condition for

$$\lim_{n\to\infty} P\left(Z_n < xb_n + a_n\right) = e^{-u(x)}.$$

The proof of this theorem is presented in [7]. Non-uniform estimate of convergence rate is given in paper [9]. In the case of diversly distributed random variables the transference theorem is proved in [10]. The asymptotics of the general distributions of normalized extremes of random variables is detailly investigated in [12] and [13]. Analysis of general distributions of extremes of random vectors $\{X_j = (X_{1,j},\dots,X_{mj}) j \geq 1\}$ is presented in [14] and [15].

5. APPLICATORY PROBLEMS

The methods and results of extreme value theory are applied in various branches of science, technology, economics, etc. The greatest part of applicatory problems are related with a technique of the "weakest link".

Suppose that X_j is the force, breaking off the j-th link. Then the force breaking off the entire chain of n links is $W_n = \overset{n}{\underset{j=1}{\wedge}} X_j$. Considering another example, suppose that we estimate the durability of the system composed of n elements. If one element of the system is operating while the others are in "hot" regime, then the durability of a such system $Z_n = \overset{n}{\underset{j=1}{\vee}} X_j$.

Similar problems are characteristic for various statistical problems on reliability of complicated systems, queueing theory problems, as well as problems of meteorology, ecology, oceanology, finance, insurance, and even of the probabilistic number theory. All about this one can read in [3].

Some concrete models

1. *Strength of a tape.* Let the tape of length l is divided into n parts and X_l is the strength of the tape. What does the classic extreme theory require and render?
Requirements:

- $F_{l_1} = F_{l_2} = \dots = F_{l_n} = F, \quad l_k = \frac{l}{n}, k = \overline{1,n},$
- $x_l = \min(x_{l_1},\dots,x_{l_n}),$
- X_{l_k} – independent, $k = \overline{1,n},$
- $F_l(x) = F(xb_l + a_l)$.

Then

$$F_l(x) = 1 - \exp\left\{-\left(\frac{x-a}{b}\right)^{\gamma}\right\},$$

or

$$F_l(x) = 1 - \exp\left\{-e^{\frac{x-c}{d}}\right\}. \tag{10}$$

Here $a, b > 0,\ c, d > 0$ are numbers.

2. *Electricity insulating covers.* Consider cylindrical sample parts of wire of length l and let h be a thickness of insulating cover ($h \ll l$). If X_j is a voltage across the insulating cover of j-th part, then the voltage across the insulator of entire electric wire is described by the following structure:

$$W_n = \min(X_1, \ldots, X_n).$$

If X_j are normal and n is large, then W_n is described well by distribution (10).

3. *Maximal velocity of the wind.* Suppose, that

$$X(t) = \{X_1(t), X_2(t)\}$$

is a velocity vector of wind in the horizontal plane. If $X_1(t)$ ir $X_2(t)$ are independent and normal, then the distribution of the structure

$$Z_T = \max(|X(t)|, t \in T)$$

is well approximated by $H_0(x)$ given by (9) in case T is large.

6. PROBLEMS TO BE SOLVED

We shall indicate a few problems of the classical scheme and extensions of it.

1) Theorems of large deviation: asymptotical analysis of the structure of the ratio

$$\frac{P(Z_n \geq xb_n + a_n)}{1 - H(x)},$$

as $n \to \infty$ and $x \to \infty$.

2) Asymptotical analysis of general distribution of normalized extremes from different samples.

3) Asymptotical analysis of distributions of products of normalized extremes.

4) Analysis of the rate of convergence in the transference theorem, in case variables X_j have different distributions.

5) Transference theorem and rate of convergence for densities.

All these problems can further be solved for extensions of classical scheme: in the case of dependent variables, variables with different distributions, etc.

REFERENCES

1. Galambos J. *The asymptotic theory of extreme order statistics,* New York : Wiley (1978)

2. Galambos J. *The development of the mathematical theory of extremes in the past half century,* Teorija veroj. i ejo prim.- 39, p. 272-293 (1994)
3. Gnedenko B.V. *Sur la distribution limite de terme maximum d'une série aleatoire,* Ann.Math.- 44, p. 423-453 (1943)
4. Gnedenko B.V. and Gnedenko D.B. *О распределениях Лапласа и логистическом как предельных в теории вероятностей,* Serdika - 8, p. 229-234 (1984)
5. Leadbetter M.R., Lindgren G. and Roptzen H. *Extremes and related properties of random sequences and processes,* New York Heidelberg Berlin (1983)
6. Resnick S.I. *Extreme values, regular variation and point processes,* Springer-Verlag. New York (1987)
7. Zolotariov V.M. *Современная теория суммирования независимых случайных величин,* М.: Наука (1986)
8. Aksomaitis A. *Non-uniform estimate of the rate of convergence in the transfer theorem for the extremal values,* Liet. mat. rinkinys. 27, p.219-223 (1987)
9. Aksomaitis A. *Non-uniform estimate of the rate of convergence in a limit theorem for max-scheme,* Liet. mat. rinkinys. 28, p. 211-215 (1988)
10. Aksomaitis A. *The transfers theorems for max-scheme,* Liet. mat. rinkinys. 29, p. 207-211 (1989)
11. Aksomaitis A. *On convergence of the maximum of dependent random variables,* Lith. Math. J. 32, p. 1-4 (1992)
12. Aksomaitis A. *Rate of convergence in the limit transfer theorem of the maximum and minimum,* Probability theory and mathematical statistics. Proceedings. VSP/TEV, p. 35-42 (1993)
13. Aksomaitis A. and Jokimaitis A. *An asymptotic of the distribution of extrema of independent random variables,* Lith. Math. J. 35, p. 1-10 (1995)
14. Aksomaitis A. and Jokimaitis A. *Asymptotical analysis of extreme values of multidimensional sample,* Tartu conference of computational statistics. Proceedings, p. 2-4 (1996)
15. Jokimaitis A. *Die Konvergenzgeschwindigkeit der Verteilung des maximums der Zufallsvektoren,* Lith. Math. J. 32, p. 229-236 (1992)