Reports of the Academy of Sciences of the USSR

MATHEMATICS

1969

SovietRxiv

Abstract

Full Text

Reports of the Academy of Sciences of the USSR
1969. Volume 188, No. 4

UDC 519.282

MATHEMATICS

Yu. N. BLAGOVESHCHENSKII

ON THE DISTRIBUTION OF THE EXTREME TERMS OF A VARIATION SERIES

(Presented by Academician A. N. Kolmogorov, 20 II 1969)

Let $X_n = (x_1, \dots, x_n)$ be a sample with distribution $F(x)$, i.e., a sequence of independent random variables identically distributed according to the law $F(x)$; we shall denote by x_{kn} the k-th term of the variation series of the sample X_n, and its distribution by $F_{kn}(x)$. For brevity, we shall speak of N-asymptotics instead of asymptotic normality, and of N_L-asymptotics when a local limit theorem holds: for some a_{kn}, s_{kn},

$$\frac{d}{dx} P\left\{\frac{x_{kn} - a_{kn}}{s_{kn}} < x\right\} \to \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{1}$$

uniformly on every finite interval of variation of x.

In papers ($^{1-6}$) the question of the asymptotics of $F_{kn}(x)$ was investigated in sufficient detail when $k = \text{const}$ and $n \to \infty$; in papers ($^{7-9}$) the case was studied when $k/n = \lambda(1 + o(1/\sqrt{n}))$ and $0 < \lambda < 1$ is fixed. In paper (^{10}) the case was investigated when $k/n \to 0$, $k \to \infty$, but $k(n)$ satisfies the limiting equality

$$\lim_{n\to\infty} \left[\sqrt{k(n + \nu n^{1-\alpha/2})} - \sqrt{k(n)}\right] = C\alpha\nu/2, \tag{2}$$

where $0 < \alpha < 1$, $C > 0$, and ν is any real number. It is easy to show: it follows from (2) that $k(n)/n^\alpha \to C^2$, but not conversely.

In the cited papers, local limit theorems were not investigated (with the exception of certain special cases, for example, for the uniform distribution).

In the present paper an attempt is made to find broad conditions on $F(x)$ for the fulfillment of N_L-asymptotics under arbitrary behavior of $k/n \to 0$ and $k \to \infty$. This case differs from those studied earlier by the two-parameter nature $(k/n, k)$ of the asymptotics.

In terms of distributions, the condition obtained in this work for the N_L-asymptotics of x_{kn} under any behavior $k \to \infty$ and $k/n \to 0$ may be described as follows: let $\theta_F = \inf\{x : F(x) > 0\}$ and

$$\rho_F(x) = \begin{cases} x - \theta_F, & \text{if } \theta_F > -\infty, \\ 1/|x|, & \text{if } \theta_F = -\infty. \end{cases}$$

Then (under certain monotonicity conditions on $f(x) = \dfrac{d}{dx}F(x)$) in a neighborhood of θ_F one must have

$$\ln F(x)/\ln \rho_F(x) \geq \delta > 0. \tag{3}$$

Furthermore, for a broad class of distributions not satisfying (3), it is shown that there always exist both sequences $k(n) \to \infty$ $(k(n)/n \to 0)$ such that the N-asymptotics for x_{kn} is violated, and sequences for which N_L-asymptotics holds. In the same class of distributions it is shown that N_L-asymptotics always holds for

$$\nu_{kn} = \underbrace{\ln \ldots \ln}_{r \text{ times}} \frac{1}{\rho_F(x_{kn})}$$

for some $r < \infty$, under any behavior $k \to \infty$ and $k/n \to 0$.

We shall call a distribution $F(x)$ locally regular (at the point θ_F) if $p_F(\lambda) = f(F^{-1}(\lambda))$ is uniquely defined in some neighborhood of $\lambda = 0$ and

$$p_F(\lambda_1)/p_F(\lambda_2) \to 1, \quad \text{when } \lambda_1/\lambda_2 \to 1 \quad (\lambda_1, \lambda_2 \to 0). \tag{4}$$

Here $F^{-1}(\lambda) = x_\lambda$ is determined from the equality $F(x) = \lambda$. We denote the class of locally regular distributions by R_l.

Theorem 1. *If $F(x) \in R_l$, then integral and local $N(a_{kn}, s_{kn})$-asymptotics hold for x_{kn} (or $F_{kn}(x)$), where*

$$a_{kn} = F^{-1}\left(\frac{k}{n+1}\right) \quad \text{and} \quad s_{kn}^2 = k(n-k+1)/(n+1)^2(n+2)\, p_F^2\left(\frac{k}{n+1}\right).$$

Let us note that when $k/n \to 0$ $(k/n \to \lambda,\ 0 < \lambda < 1)$, the normalizing constants a_{kn} and s_{kn} pass into the corresponding normalizing constants of work (9).

For many distributions $F(x)$, the computation of $p_F(\lambda)$ and the verification of the condition $F(x) \in R_l$ are laborious tasks. In this connection, it would be desirable to describe the class of R_l-distributions in terms of $F(x)$ and $f(x)$.

Under the condition of sufficiently monotone behavior of $F(x)$ as $x \to \theta_F$, the class R_l is well described by the following two theorems.

Theorem 2. *Let $\theta_F = -\infty$, $F_1(x) = F(\mathrm{tg}(\pi x - \pi/2))$, and $F_2(x) = F_1(x+\theta)$, where θ is any fixed real number. Then the distributions $F(x)$, $F_1(x)$, and $F_2(x)$ either all belong to R_l, or all do not belong to R_l.*

Theorem 3. *Let $\theta_F = 0$, and let there exist a finite or infinite limit $x^2(\ln f(x))''$ as $x \to 0$. Then, if $\varepsilon_F(x) = \ln F(x)/\ln f(x)$ does not have zero as a limit point, then $F(x) \in R_l$; if $\varepsilon_F(x)$ tends monotonically to zero as $x \to 0$, then $F(x) \in R_l$.*

We point out that, under the conditions of the theorem, $\varepsilon_F(x)$ has only one limit point, and moreover the set of limit points for distributions satisfying the condition of the theorem coincides with the line without the interval $(0,1)$. Finally, in order that $F(x) \in R_l$ under the conditions of Theorem 3, one can also indicate four equivalent* conditions: as $x \to 0$,

$$\lim x^2(\ln f(x))'' < 1; \tag{1}$$

$$\lim x(\ln f(x))' > -1; \tag{2}$$

$$\lim \ln f(x)/\ln x > -1; \tag{3}$$

$$\lim[\ln F(x)/\ln x] > 0. \tag{4}$$

As a consequence of Theorems 2 and 3, we note that all continuous distributions used in statistics (for example, the normal, Cauchy, Weibull, gamma, F-distribution, and others) are locally regular, i.e., belong to the class R_l.

The heuristic ideas of the proof of the integral part of Theorem 2 are extremely simple. Let $\xi_{kn} = F(x_{kn})$; then ξ_{kn} is the k-th term of the order statistic of a sample of size n from the uniform distribution on $(0,1)$. For ξ_{kn}, Theorem 1 is proved by a direct analytic method with $a_{kn} = k/(n+1)$ (instead of a_{kn}) and $\sigma_{kn}^2 = k(n-k+1)/(n+1)^2(n+2)$ (instead of s_{kn}); it is essential that $\sigma_{kn}/a_{kn} \sim 1/\sqrt{k} \to 0$. Next let $\zeta_{kn} = (\xi_{kn} - a_{kn})/\sigma_{kn}$; then heuristically

$$x_{kn} = F^{-1}(a_{kn} + \sigma_{kn}\zeta_{kn}) = F^{-1}(a_{kn}) + \sigma_{kn}\zeta_{kn}/p_F(a_{kn}) + \cdots,$$

since

$$\frac{d}{d\lambda}F^{-1}(\lambda) = \frac{1}{p_F(\lambda)},$$

and it remains only to justify the corresponding expansions, which turns out to be possible for $F(x) \in R_l$.

---

* As was observed by A. N. Kolmogorov, a sufficient condition is also the requirement $|f'(x)F(x)| \leq Cf^2(x)$, $C < \infty$, for $F(x) \leq \delta_0$, $\delta_0 > 0$.

Let now $\Psi_{kn}(x)$ and $\psi_{kn}(x)$ be the distribution and density for ζ_{kn}. Then the distribution $F^*_{kn}(x)$ of $(x_{kn} - a_{kn})/s_{kn}$ coincides, for sufficiently large k and n, with $\Psi_{kn}(y_{kn}(x))$, where

$$y_{kn}(x) = \frac{1}{\sigma_{kn}} \left[F\left(F^{-1}(\alpha_{kn}) + \sigma_{kn}x/p_F(\alpha_{kn})\right) - \alpha_{kn}\right].$$

For the local theorem it is sufficient, obviously, to prove that $y_{kn}(x) \to x$ and $y'_{kn}(x) \to 1$, which is done by purely analytic means for $F(x) \in R_l$. Indeed, from the fact that $\Psi_{kn}(y_{kn}(x)) \to \Phi(x)$ and $\Psi_{kn}(y_{kn}(x)) - \Phi(y_{kn}(x)) \to 0$, it follows, obviously, that $y_{kn}(x) \to x$ (here

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{u^2}{2}\right) du$$

).

Further,

$$y'_{kn}(x) = \frac{f\left(F^{-1}(\alpha_{kn}) + \sigma_{kn}x/p_F(\alpha_{kn})\right)}{p_F(\alpha_{kn})}.$$

We note that from the boundedness of the sequence $y_{kn}(x)$ (since $y_{kn}(x) \to x$) and the definition it follows that

$$F^{-1}(\alpha_{kn}) + \sigma_{kn}x/p_F(\alpha_{kn}) = F^{-1}(\alpha_{kn} + \sigma_{kn}y_{kn}(x)).$$

Hence we obtain that

$$y'_{kn}(x) = p_F(\alpha_{kn} + \sigma_{kn}y_{kn}(x))/p_F(\alpha_{kn}) \to 1$$

for any fixed x, since $\sigma_{kn}y_{kn}(x)/\alpha_{kn} \to 0$ and $F(x) \in R_l$.

Theorem 2 is rather simple, while the idea of Theorem 3 consists in the fact that, when $f(x) \to \infty$ sufficiently rapidly (as $x \to 0$), the distribution $F_{kn}(x)$ may fail to have time to symmetrize ($x_{kn} > 0$, but sufficiently close to zero when $k/n \to 0$), and this leads to asymptotic nonnormality and nonregularity of $F(x)$. A drawback of Theorem 3 is the requirement of some smoothness conditions on

$f(x)$. A broad class of locally regular distributions is singled out by Lemmas 1 and 2.

Lemma 1. If $\theta_F = 0$ and

$$\lim_{x\to 0} \frac{f(x)}{u_m(x)} = 1, \qquad u_m(x) = Cx^{\alpha_0}(\ln(1/x))^{\alpha_1} \cdots (\ln\ln\cdots\ln(1/x))^{\alpha_m},$$

where $\alpha_0 > -1$, and $\alpha_1, \dots, \alpha_m$ and $C > 0$ are otherwise arbitrary real numbers, then $F(x) \in R_l$.

Lemma 2. Let $G(x)$ be concentrated on $[0,1]$ and let $F(x) = \exp[1 - 1/G(x)]$. Then $F(x)$ is again a distribution on $[0,1]$ and $F(x) \in R_l$, if $G(x) \in R_l$.

For example, the distribution $F(x) = \exp\left(1 - \frac{1}{x}\right) \in R_l$, but

$$F(x) = \frac{1}{1 + \ln(1/x)} \notin R_l.$$

This example shows the noninvertibility of Lemma 2.

Let us now consider what the asymptotic distributions of x_{kn} may be as $k \to \infty$ and $k/n \to 0$, when $F(x) \notin R_l$.

This problem is not solved completely in the present paper; one fairly broad class of irregular distributions is investigated to the end (with $\theta_F = 0$, since, using Theorem 2, it is easy to transfer the result also to the case $\theta_F \neq 0$). We note that among sufficiently smooth and monotone functions $f(x)$, the irregular distributions lie in the set of those distributions for which $x^\alpha f(x) \to \infty$ as $x \to 0$ for every $\alpha < 1$. Such distributions $F_G(x)$, for example, are obtained if $G(x) = x^\alpha$, $\alpha > 0$, and

$$F_G(x) = G\left(\frac{1}{\pi}\operatorname{arctg}(\ln x) + \frac{1}{2}\right).$$

We shall retain the notation $F_G(x)$ for the distribution obtained by this formula from an arbitrary distribution $G(x)$ concentrated on $[0,1]$.

Theorem 4. For $F_G(x) \in R_l$, it is necessary and sufficient that $G(x) \in R_l$ and

$$p_G(\lambda)\,[G^{-1}(\lambda)]^2/\lambda \geqslant \delta_0 > 0$$

for all $\lambda \leqslant \lambda_0,\ \lambda_0 > 0$.

It is easy to obtain, as a consequence of Theorem 4, that all $F_G(x)$, when $g(x) = dG(x)/dx = x^\alpha(1 + \varepsilon(x))$, $\alpha > -1$, and $\varepsilon(x) \to 0$ as $x \to 0$, are irregular.

Having obtained a criterion of irregularity and described an entire family of irregular distributions, we would like to describe the possible limiting distributions in it for x_{kn}. This question is resolved in Theorem 5.

Theorem 5. If $F_G(x) \in R_l$ and $G(x) \in R_l$, then for the N-asymptotics of $x_{k_j n_j}$ (both local and integral) it is necessary and sufficient that

$$\lim_{j\to\infty} \sqrt{k_j/n_j}\, p_G(k_j/n_j)\, [G^{-1}(k_j/n_j)]^2 = 0,$$

and, for any $F_G(x) \in R_l$, there exist sequences (k_j, n_j) both satisfying this condition and not satisfying it. In any case, Theorem 1 holds for $\ln x_{kn}$.

Consider the family of F_G-distributions given by the formula

$$H_\alpha(x) = \left(\frac{1}{\pi}\operatorname{arc\,tg}(\ln x) + \frac{1}{2}\right)^\alpha, \qquad \alpha > 0.$$

As was already noted above, $H_\alpha(x) \in R_l$. It is easy, using Theorem 6, to prove the following assertion:

Lemma 3. If the distribution of the population is $H_\alpha(x)$, then for x_{kn}, when $k/n^{2/(2+\alpha)} \to \infty$, local and integral N-asymptotics hold, while when $k/n^{2/(2+\alpha)} \leqslant C$, $k \to \infty$, and $C < \infty$, local and integral N-asymptotics hold for $\ln x_{kn}$ and do not hold for x_{kn}.

Considering distributions $F_{G,r}(x)$, which near $x \to 0$ ($\theta_F = 0$) can be described by applying the operator F_G r times to a regular distribution $G(x)$, one can show that, for them,

$$\underbrace{\ln \ldots \ln}_{r \text{ times}} \frac{1}{x_{kn}}$$

satisfies Theorem 1, and there necessarily exist sequences (k_j, n_j) such that $x_{k_j n_j}$ possesses N_L-asymptotics.

Moreover, one can construct order functions $\gamma_0(n), \gamma_1(n), \ldots, \gamma_r(n)$ such that $\gamma_j(n)/\gamma_{j-1}(n) \to 0$ as $n \to \infty$ for $j = 1, 2, \ldots, r$, and $\gamma_0(n)/n \to 0$; moreover, if $k_n \to \infty$ in such a way that $k_n/n \to 0$ and $k_n/\gamma_0(n) \to \infty$, then N_L-asymptotics will hold, while if $k_n/\gamma_{j-1}(n) \to 0$ and $k_n/\gamma_j(n) \to \infty$, then N-asymptotics and N_L-asymptotics will hold only after taking logarithms of x_{kn} at least j times.

Let us also point out that a change in the sign of $f'(x)$ in an arbitrarily small neighborhood of θ_F may have a substantial influence on the asymptotics of x_{kn}. Let, for example, $f(x) = C(1 + \sin(1/x))$, $x \in (0, 1]$. Then $F(x) = x + x^2(\cos(1/x) + o(x^2))$ as $x \to 0$, i.e. $F(x) = x(1 + o(x))$. A direct calculation proves that $F(x) \in R_l$ and that one can choose $k_n \to \infty$ ($k_n/n \to 0$) so that the distribution of x_{kn}, under no normalization whatsoever, will converge to a nondegenerate limiting distribution.

Moscow State University
named after M. V. Lomonosov

Received
20 I 1969

REFERENCES

1. E. L. Dodd, Trans. Am. Math. Soc., **25**, 525 (1923).
2. P. A. Fisher, L. H. C. Tippett, Proc. Cambr. Phil. Soc., **24**, 180 (1928).
3. M. Fréchet, Ann. Soc. Polonaise Math., **6**, 92 (1927).
4. E. J. Gumbel, Ann. Inst. Henri Poincaré, **4**, 115 (1935).
5. N. Smirnov, Metron, **12**, No. 2, 59 (1935).
6. E. Gumbel, *Statistical Theory of Extreme Values*, Moscow, 1965.
7. K. Pearson, Biometrika, **13**, 113 (1920).
8. F. C. Mosteller, Ann. Math. Stat., **17**, 377 (1946).
9. N. V. Smirnov, Proc. Steklov Math. Inst. Acad. Sci. USSR, **15** (1949).
10. D. M. Chibisov, Theory of Probability and Its Applications, **9**, issue 1, 159 (1964).

Note: Figure translations are in progress. See original paper for figures.