Ulyanov V.V. (Moscow, Russia), **Christoph G.** (Magdeburg, Germany), **Monakhov M.M.** (Moscow, Russia) — **Asymptotic expansions for the distributions of statistics with random sample size.**

In practice, we often encounter situations where a sample size is not defined in advance and can be random itself. In [1] it is demonstrated that the asymptotic properties of the statistics can be radically changed when the non-random sample size is replaced by a random value. In the talk, the second order Chebyshev–Edgeworth and Cornish–Fisher type expansions (see [2]) based on Student's t- and Laplace distributions and their quantiles are derived for sample mean and sample median with random sample size of a special kind. We use general transfer theorem (see [3]) which enables us to construct the asymptotic expansions for the distributions of the randomly normalized statistics applying the asymptotic expansions for the distributions of the considered non-randomly normalized statistics and for the distributions of the random size of the underlying sample.

Let the random variables (r.v.) $X, X_1, X_2, \ldots \in \mathbf{R}$ and $N_1, N_2, \ldots \in \mathbf{N}$ be defined on a probability space $(\Omega, \mathbf{A}, \mathbf{P})$. In mathematical statistic, $X_1, X_2, \ldots$ are the observations and N_n is a random size of the sample, depending on the integer parameter $n \in \mathbf{N}$. Suppose, that for any $n \in \mathbf{N}$, r.v. $N_n \in \mathbf{N}$ is independent of $X_1, X_2, \ldots$. Let $T_m := T_m(X_1, \ldots, X_m)$ be a statistic based on the sample with fixed size $m \in \mathbf{N}$. Put $T_{N_n}(\omega) := T_{N_n(\omega)}\left(X_1(\omega), \ldots, X_{N_n(\omega)}\right)$, $\omega \in \Omega$, that is T_{N_n} is the statistic based on the statistic T_m with random sample size N_n. For example, take the sample median as T_m. Let the distribution function of N_n for integer k be $(k/(1+k))^n$, that corresponds to the fact: N_n is the maximum over n independent r.v. with some discrete Pareto distribution. Then one can get non-asymptotic second order approximations for the distribution of N_n (see [4]), and for T_m (see [5]) under some regularity conditions on the density function of X_1. Moreover, by the refined version of the transfer theorem (see [5]), we get inequality $\sup_x |F(T_{N_n}, x) - L(n, x)| \le c/n^{-3/2}$, where $F(T_{N_n}, x)$ is the distribution function of the normalized random sample median T_{N_n} and $L(n, x)$ is the second order approximation with respect to powers $n^{-1/2}$ with Laplace distribution as the limit function with the density function as $\exp\{-\sqrt{2}\,|x|\}/\sqrt{2}$. Other options for T_m with non-asymptotic results for it see in [7], Ch.13–16, including the case of the high-dimensional observations.

The formulations and proofs of the results mentioned in the talk see in [4–6].

REFERENCES

1. *Gnedenko B. V.* An estimate of the distribution of the unknown parameters with a random number of independent observations. [in Russian], Proceedings of Tbilisi Math. Inst., AN GSSR, 1989, v. 92, 146–150.
2. *Ulyanov V.V., Aoshima M., Fujikoshi Y.*, Non-asymptotic results for Cornish–Fisher expansions, J. Math. Sci., 2016, vol. 218, No. 3, 363–368.
3. *Bening V. E., Galieva N. K., Korolev V. Yu.* Asymptotic expansions for the distribution functions of statistics constructed from samples with random sizes [in Russian], Informatics and its Applications, IPI RAN, 2013, v. 7, no 2, 75–83.
4. *Christoph G., Monakhov M.M. and Ulyanov V.V.*, Second order Chebyshev–Edgeworth and Cornish–Fisher expansions for distributions of statistics constructed from samples with random sizes, (in Russian), Zapiski Nauchnykh Seminarov POMI, 2017, v. 466, 167–207. English version see: http://www.math.sci.hiroshima-u.ac.jp/stat/TR/TR17/TR17-15.pdf
5. *Christoph G., Ulyanov V.V., Bening V.E.*, Second Order Expansions for Sample Median with Random Sample Size, ArXiv 1905.07765, 2019, 1–17.
6. *Markov A. S., Monakhov M. M. and Ulyanov V. V.*, Generalized Cornish–Fisher expansions for distributions of statistics based on samples of random size. [in Russian], Inform. Primen., 2008, v. 10, no 2, 84–91.
7. *Fujikoshi Y., Ulyanov V.V., Shimizu R.* Multivariate Statistics: High-Dimensional and Large-Sample Approximations. Wiley Series in Probability and Statistics, Wiley, Hoboken, N.J., 2010.

This work was supported by the RSF grant No 18-11-00132 and by the Russian Academic Excellence Project '5–100'.