LIMIT DISTRIBUTION OF THE NUMBER OF PAIRS IN ONE URN SCHEME

E.V. Cherepanova

Karelian Research Center of RAS

Let mN different balls be located in an urn, where N is the number of balls' colours and m is the number of balls of identical colour. We are realizing equiprobable choice of n balls from the urn without return, $0 < n < mN$. We denote by τ the number of different pairs of balls of the same colour. We are intrested in the limit distribution of τ. In [1] this distribution was obtained for a fixed m and n, $N \longrightarrow \infty$, $0 < \alpha_0 \le n/(mN) \le \alpha_1 < 1$. The following theorem is the generalization of this result.

Theorem. *Let $m \ge 2$, n, $N \longrightarrow \infty$ in such a way that $mN - n \longrightarrow \infty$ and one of the following conditions is valid:*

1. $n/N \longrightarrow 0$, $n^2/N \longrightarrow \infty$;
2. $0 < c_1 \le n/N \le c_2 < \infty$, $n\left(1 - n/(mN)\right)^3 \longrightarrow \infty$;
3. $n/N \longrightarrow \infty$, $N^{4/3}\left(1 - n/(mN)\right)/n \longrightarrow \infty$, $n^3\left(1 - n/(mN)\right)/N^3 = O(1)$.

Then

$$\mathbf{P}\left\{\tau = k\right\} = \frac{1}{\sigma\sqrt{2\pi}} exp\left\{-\frac{(k-a)^2}{2\sigma^2}\right\}(1 + o(1))$$

uniformly in $(k-a)/\sigma$ lying in any fixed finite interval, where $a = (m-1)n^2/(2mN)$, $\sigma^2 = (m-1)(mN-n)^2 n^2/(2m^3N^3)$.

It was shown in [2] that

$$\mathbf{P}\left\{\eta_1 = k_1, \ldots, \eta_N = k_N\right\} = \\ = \mathbf{P}\left\{\xi_1 = k_1, \ldots, \xi_N = k_N \mid \xi_1 + \ldots + \xi_N = n\right\},$$

where η_i is the number of balls of the i-th colour, $1 \le i \le N$, $\xi_1, \ldots, \xi_N$ are independent identically distributed random variables with the binomial distribution. It folows from this that

$$\mathbf{P}\{\tau = k\} = \mathbf{P}\{\zeta_N = k \mid \mu_N = n\} = \frac{\mathbf{P}\{\zeta_N = k,\ \mu_N = n\}}{\mathbf{P}\{\mu_N = n\}},$$

where $\zeta_N = \xi_1(\xi_1 - 1)/2 + \ldots + \xi_N(\xi_N - 1)/2,\ \mu_N = \xi_1 + \ldots + \xi_N$. Proving local limit convergence for array schemes of the sums μ_N and (ζ_N, μ_N) is the main difficulty for obtaining the theorem.

Research is supported by the Russian Foundation for Basic Research, grant 00-01-00233.

References

1. S.V. Agievich. About one combinatorial allocation problem. Computer date analysis and modelling. Proc. of V Inter. conf., v.3, Minsk, 1998, 28-31 (in Russian).

2. V.F. Kolchin. Random Mappings. Springer, New York, 1986.