Functional central limit theorems for infinite urn models

Mikhail Chebunin
Novosibirsk State University, Russia.
e-mail: chebuninmikhail@gmail.com

We consider the classical multinomial occupancy scheme in which balls are thrown independently at a fixed infinite series of urns, with probability p_i of hitting the ith urn. The frequencies $(p_i,\ i = 1, 2, \dots)$ are assumed nonincreasing, strictly positive and satisfying $\sum_i p_i = 1$. As n balls are thrown, their allocation is captured by the array $X_n = (X_{n,j},\ j = 1, 2, \dots)$, where $X_{n,j}$ is the number of balls out of the first n that fall in urn j.

A functional of X_n which appears in many contexts is the number of nonempty urns

$$R_n = \#\{j : X_{n,j} > 0\}.$$

R_n is sometimes regarded as a measure of diversity of the sample. More detailed information is carried by the counts of urns occupied by exactly k balls

$$R_{n,k} = \#\{j : X_{n,j} = k\} \quad (k = 1, 2, \dots),$$

so that $R_n = \sum_k R_{n,k}$ and $\sum_k kR_{n,k} = n$.

The first systematic study of the asymptotics R_n and $R_{n,k}$ appeared in a remarkable paper by Karlin (1967), in which he proved a central limit theorem under a condition of regular variation on the frequencies $\{p_i\}$.

Dutko (1989) showed that R_n is asymptotically normal, assuming only the necessary condition that its variance tends to infinity with n.

Gnedin, Hansen and Pitman (2007) focused on study of conditions for convergence $\mathbf{Var}R_n \to \infty$.

Hwang and Janson (2008) proved local limit theorems for finite and infinite number of cells.

Barbour and Gnedin (2009) showed that $R_{n,k}$ is asymptotically normal if variances go to infinity.

We will discuss the validity of the functional central limit theorem (Chebunin and Kovalevskii (2016)) provided under a condition of regular variation on the frequencies $\{p_i\}$.

References

Barbour, A. D., 2009. Univariate approximations in the infinite occupancy scheme. Alea 6, 415–433.

Barbour, A. D., Gnedin, A. V., 2009. Small counts in the infinite occupancy scheme. Electronic Journal of Probability, Vol. 14, Paper no. 13, 365–384.

Chebunin M., Kovalevskii A., 2016. Functional central limit theorems for certain statistics in an infinite urn scheme. Statistics and probability letters, Vol. 119, 344–348.

Dutko, M., 1989. Central limit theorems for infinite urn models, Ann. Probab. 17, 1255–1263.

Gnedin, A., Hansen, B., Pitman, J., 2007. Notes on the occupancy problem with infinitely many boxes: general asymptotics and power laws. Probability Surveys, Vol. 4, 146–171.

Hwang, H.-K., Janson, S., 2008. Local Limit Theorems for Finite and Infinite Urn Models. The Annals of Probability, Vol. 36, No. 3, 992–1022.

Karlin, S., 1967. Central Limit Theorems for Certain Infinite Urn Schemes. Jounal of Mathematics and Mechanics, Vol. 17, No. 4, 373–401.