A BERRY-ESSEEN BOUND FOR EMPTY BOXES STATISTIC ON THE SCHEME AN ALLOCATIONS OF SEVERAL TYPE BALLS*

S.A. Mirakhmedov & O. Saidova

Abstract

A Berry-Esseen bound for the number of empty cells in the scheme of independent and random allocation of balls of $s$ type into different cells is obtained.

Key words and phrases: Central limit theorem, empty cells, random allocations.

Introduction

Let $n_1$ balls be of first type and $n_2$ be balls of a second type, etc., and $n_s$ balls of $s$th type be distributed independently and randomly into $N$ different cells, in such a way that each ball of $i$th type has probability $p_{ik}$ of landing into $k$th cell, $p_{i1} + \cdots + p_{iN} = 1$, $i = 1, \ldots, s$. Let $\mu_0(s) = \mu_0(s, N, n_1, \ldots, n_s)$ be a number of empty cells after all $n_1, \ldots, n_s$ losses. If $s = 1$ we deal with multinomial scheme of an allocation and $\mu_0(1)$ is a well-known empty box test statistic (see, for example, Koichin, Sevastjanov, Chistjacov (1976)). For example, the random variable (r.v.) $\mu_0(s)$ used as test statistic for verification of homogeneity hypothesis.

Here we get a bound for remainder in the central limit theorem for $\mu_0(s)$. Our theorem generalizes the result of Quine and Robinson (1982).

Result. We consider the case that $s$ is fixed and $N = N(n_1, \ldots, n_s)$ is growing as one of $n_1, \ldots, n_s$ increases. Suppose that for all $j = 1, \ldots, s$ and $k = 1, \ldots, N$

$$Np_{jk} \leq C_0 \quad \text{and} \quad n_i \leq \exp(C_1N). \quad (1)$$

Here and in what follows, $C_j, C_j(\cdot)$ are positive constants not dependent on $N, n_1, \ldots, n_s$.

Denote

$$\lambda_{jm} = n_j p_{jm}, \quad \lambda_m^{(\epsilon)} = \lambda_{1m} + \cdots + \lambda_{sm}, \quad \alpha_i = n_i/N.$$
Theorem. Under condition (1) for arbitrary $\nu \geq 0$ there exist $C(s, C_0, \nu)$ such that

$$
\omega_N(x) \leq C(s, C_0, \nu) \left[ \frac{1}{\sigma_N(s)} + \sum_{j=1}^{s} \frac{1}{\sqrt{n_j}} \right].
$$

Corollary 1. Under condition (1) there exist $C(s) > 0$ such that

$$
\sup_x \omega_N(x) \leq C(s) \left[ \frac{1}{\sigma_N(s)} + \sum_{j=1}^{s} \frac{1}{\sqrt{n_j}} \right].
$$

Corollary 2. Suppose that $N, n_1, \ldots, n_s \to \infty$ in such a way that $\sigma_N(s) \to \infty$ and (1) is hold true. Then $\mu_0(s)$ is asymptotically normal.

The result of Quine and Robinson (1982) is $\omega_N^{(1)}(x) \leq C\sigma_N^{-1}(1)$, which also follows from our theorem since $\sigma_N^2(1) \leq n_1$: but in the general case we have $\sigma_N^2(s) \leq n_1 + \cdots + n_s$.

Proof. It is clear that

$$
\mu_0(s) = \sum_{m=1}^{N} f(\eta_{1m}, \ldots, \eta_{sm}),
$$

where $\eta_{ik}$ is a number of balls of $i$th type in $m$th cell after $n_1, \ldots, n_s$ tosses, and $f(0, \ldots, 0) = 1$ and $f(y_1, \ldots, y_s) = 0$ if $y_i > 0$ for some $i = 1, \ldots, s$.

Let $\zeta_{jm}$ be a Poisson with parameter $\lambda_{jm}$, $\zeta_m^{(s)} = (\zeta_{1m}, \ldots, \zeta_{sm})$, then

$$
g_m(\zeta_m^{(s)}) = f(\zeta_m^{(s)}) - \exp\{-\lambda_m\} + \sum_{i=1}^{s} a_i(\zeta_{im} - \lambda_{im}).
$$

From Corollary 2 of Mirakhmedov (1987) we get

$$
\omega_N^{(s)}(x) \leq \frac{C(s, k)}{1 + |x|^k} \left[ \beta_{1N} + \beta_{k+2, N} + \sum_{i=1}^{s} \frac{1}{\sqrt{n_i}} \right],
$$

(2)
for any integer \( k > 0 \), where

\[
\beta_{kN} = \frac{1}{\sigma^2_N(s)} \sum_{m=1}^{N} E |g_m(c_m^{(s)})|^k.
\]

We remark that

\[
\sigma^2_N(s) = \sum_{m=1}^{N} D g_m(c_m^{(s)}).
\]

We rewrite \( \sigma^2_N(s) \) and \( g_m(c_m^{(s)}) \) as follows:

\[
\sigma^2_N(s) = \sum_{m=1}^{N} [1 - (1 + \lambda_m) \exp\{-\lambda_m\}] \exp\{-\lambda_m\} + \sum_{m=1}^{N} \sum_{j=1}^{s} \lambda_{jm}(\exp\{-\lambda_m\} - a_j)^2,
\]

\[
g_m(c_m^{(s)}) = f(c_m^{(s)}) + \exp\{-\lambda_m\} \sum_{i=1}^{s} \zeta_{im} - (1 + \lambda_m) \exp\{-\lambda_m\} + \sum_{i=1}^{s} (a_i - \exp\{-\lambda_m\})(\zeta_{im} - \lambda_{im}).
\]

Then for arbitrary \( b > 1 \) we get

\[
E |g_m(c_m^{(s)})|^b \leq 2^{b-1} E \left| f(c_m^{(s)}) + \exp\{-\lambda_m\} \sum_{i=1}^{s} \zeta_{im} - (1 + \lambda_m) \exp\{-\lambda_m\} \right|^b + (2s)^{b-1} \sum_{i=1}^{s} |a_i - \exp\{-\lambda_m\}|^b E |\zeta_{im} - \lambda_{im}|^b \equiv 2^{b-1} \Delta_{1m} + (2s)^{b-1} \Delta_{2m}.
\]

The r.v. \( f(c_m^{(s)}) \) has the same distribution as r.v. \( \varphi(\zeta_{1m} + \cdots + \zeta_{sm}) \) where \( \varphi(0) = 1 \) and \( \varphi(x) = 0 \) if \( x > 0 \). Thus

\[
\Delta_{1m} = E \left| \varphi(\zeta_{1m} + \cdots + \zeta_{sm}) + \exp\{-\lambda_m\} \sum_{i=1}^{s} -(1 + \lambda_m) \exp\{-\lambda_m\} \right|^b \leq \exp\{-\lambda_m\}(1 - \exp\{-\lambda_m\}(1 + \lambda_m))^{b} + \lambda_m^{b+1} \exp\{- (b - 1) \lambda_m \}
\]

\[
+ \sum_{j=2}^{\infty} (j - 1\lambda_m)^b \exp\{- (b + 1) \lambda_m \} \frac{\lambda_j}{j!} \equiv \Delta'_{1m} + \Delta''_{1m} + \Delta'''_{1m}
\]

because \( \zeta_{1m} + \cdots + \zeta_{sm} \) is Poisson with parameter \( \lambda_m \). Since \( (1 + u)e^{-u} < 1 \) for \( u > 0 \) and (3), we have

\[
\sum_{m=1}^{N} \Delta'_{1m} \leq \sum_{m=1}^{N} \exp\{-\lambda_m\}(1 - \exp\{-\lambda_m\}(1 + \lambda_m)) \leq \sigma^2_N(s).
\]
Therefore we get

\[ \sum_{m=1}^{N} \Delta_{1m}'' \leq \sum_{m=1}^{N} \lambda_m^2 \exp\{-2\lambda_m\} \leq 2 \sum_{m=1}^{N} (1 - (1 + \lambda_m)) \exp\{-\lambda_m\} \leq 2\sigma_N^2(s). \quad (7) \]

Let \( b \) be odd, \( \sum_{\lambda_m \leq 1} \) and \( \sum_{\lambda_m \geq 1} \) be a sum on \( m \) such that \( \lambda_m \leq 1 \) and \( \lambda_m \geq 1 \), correspondingly. We have

\[
\sum_{\lambda_m \leq 1} \Delta_{1m}'' = \sum_{\lambda_m \leq 1} \exp\{-b\lambda_m\}[E(\zeta_{1m} + \cdots + \zeta_{sm} - 1 - \lambda_m)^b + (-1)^b (1 + \lambda_m)^b \exp\{-\lambda_m\}] + (-1)^{b+1} \exp\{-\lambda_m\},
\]

if \( \lambda_m \leq 1 \) then

\[ E(\zeta_{1m} + \cdots + \zeta_{sm} - 1 - \lambda_m)^b = \sum_{i=1}^{b} C_i^b (-1)^i E(\zeta_{1m} + \cdots + \zeta_{sm})^{b-1} \leq C(b) \lambda_m^2 - 1 - (b-1)\lambda_m. \]

Therefore we get

\[
\sum_{\lambda_m \leq 1} \Delta_{1m}'' \leq \sum_{\lambda_m \leq 1} \exp\{-b\lambda_m\}(C(b) \lambda_m^2 - (1 + (b-1) \lambda_m))(1 - (1 + \lambda_m) \exp\{-\lambda_m\})
\]

\[ \leq C(b) \sum_{m=1}^{N} \lambda_m^2 \exp\{-2\lambda_m\} \leq C(b)\sigma_N^2(s). \quad (8) \]

Since \( (1 + \lambda_m)^b \leq 1 + (b-1)\lambda_m + C(b) \lambda_m^2 \), if \( \lambda_m \leq 1 \). Using well known inequality between moments of r.v., we obtain:

\[
\sum_{\lambda_m \geq 1} \Delta_{1m}'' \leq \sum_{\lambda_m \geq 1} \exp\{-b\lambda_m\} E|\zeta_{1m} + \cdots + \zeta_{sm}|^b
\]

\[ \leq \sum_{\lambda_m \geq 1} (E(\zeta_{1m} + \cdots + \zeta_{sm} - 1 - \lambda_m)^{b+1})^{b+1} \exp\{-b\lambda_m\}
\]

\[ \leq C(b) \sum_{\lambda_m \geq 1} \lambda_m^2 \exp\{-b\lambda_m\} \leq C(b) \sum_{\lambda_m \geq 1} \lambda_m^2 \exp\{-2\lambda_m\} \leq C(b)\sigma_N^2(s). \quad (9) \]

From (5), (6), (7), (8) and (9) follows

\[ \sum_{m=1}^{N} \Delta_{1m} \leq C(b)\sigma_N^2(s). \quad (10) \]

Let us estimate \( \sum_{m=1}^{N} \Delta_{2m} \). We have

\[
\sum_{m=1}^{N} |a_k - \exp\{-\lambda_m\}|^b E|\zeta_{km} - \lambda_{km}| \leq \sum_{\lambda_{km} \leq 1} (a_k - \exp\{-\lambda_m\})^b E|\zeta_{km} - \lambda_m|^b
\]

50
Here, we used that \( \lambda_{km} \leq C_0 \alpha_k, \ldots, a_k \leq 1 \) and \( \lambda_{km} \leq C_0 \) if \( \alpha_k \leq 1 \), \( E|\zeta_{km} - \lambda_{km}|^i \leq C(i)\lambda_{km} \).

Let \( \lambda_{km} > 1, \alpha_k > 1 \). Since \( a_k \leq \alpha_k^{-1} \) we have

\[
|a_k - \exp(-\lambda_m)|^{1/2} \leq a_k\lambda_m^{1/2} + 1 \leq \sqrt{C_0/\alpha_k} + 1 \leq \sqrt{C_0} + 1.
\]

Therefore

\[
\sum_{\lambda_{km} > 1, \alpha_k > 1} (a_k - \exp(-\lambda_m))\lambda_{km}^{1/2} \leq (\sqrt{C_0} + 1)^{b-2} \Sigma(a_k - \exp(-\lambda_m))^2 \lambda_{km}
\]

\[
\leq (\sqrt{C_0} + 1)^{b-2} \sigma_N^2(s).
\]

From this and (11)

\[
\sum_{m=1}^N \Delta_{2m} \leq C(b)\sigma_N^2(s). \tag{12}
\]

Thus if \( b \) is odd, then by (4), (10), (11) it follows that

\[
\sum_{m=1}^N E|g_m(\zeta_m^{(a)})|^b \leq C(b)\sigma_N^2(s). \tag{13}
\]

If \( b \) is odd then the theorem follows from (2) and (13). If \( b \) is even then the theorem follows from the well-known inequality between Ljapunov’s ratio and (2), (13).

Proof of theorem is complete. \( \square \)
References


Serzod A. MIRAKHMEDOV & Odila SAIDOVA
Department of Mathematics
Tashkent State University
700095, Tashkent - UZBEKISTAN

Received 13.05.1996