DEMONSTRATIO MATHEMATICA
Vol. XXX No 3 1997

Dobiesław Bobrowski, Barbara Popowska

THE ESTIMATE OF THE ERROR FOR APPROXIMATION OF SOME DISCRETE DISTRIBUTION BY THE GEOMETRIC DISTRIBUTION

We consider discrete distribution from the class (D)NBUE. When we know the mean ($1 < \mu < \infty$) and the variance ($\sigma^2 < \infty$), then we can estimate the error of the approximate of such distribution by a geometric distribution with some parameter p. The estimate is sharp. Moreover we show that a sequence of distributions from (D)NBUE tends to the geometric one if and only if the sequence of parameters $a_m \to 0$ when $m \to \infty$.

Let $(p_n) = (p_n, n \in N)$ be given discrete probability distribution such that there exists its finite expected value $\mu > 1$. For this distribution we introduce the following denotations, for all $n \geq 0$,

$$R_n = \sum_{k=n+1}^{\infty} p_k, \tag{1}$$

$$\overline{G}_n = \frac{1}{\mu} \sum_{k=n}^{\infty} R_k, \tag{2}$$

and

$$\Delta_n = R_n - \overline{G}_n. \tag{3}$$

Hence we have

$$R_0 = 1,\ 0 \leq R_n \leq 1,\ \sum_{k=0}^{\infty} R_k = \mu, \tag{4}$$

AMS 1991 subject classification: Primary 62E17 Secondary 60E99, 62N05, 90B25.
Keywords: Discrete distribution, geometric distribution, error of approximation.

(5) $$\overline{G}_0 = 1,$$

(6) $$\Delta_0 = 0 \text{ and } \lim_{n\to\infty} \Delta_n = 0.$$

Moreover denote

(7) $$p = \frac{1}{\mu} \text{ and } q = 1 - p.$$

The assumption about μ implies that $p, q \in (0,1)$.

Ma Becheng [1] defined among others two classes of discrete distributions:

— (D)HNBUE (discrete harmonic new better than used in expectation class) if

(8) $$\overline{G}_n \leq q^n,\ n \in N_0,$$

— (D)NBUE (discrete new better than used in expectation class) if

(9) $$\overline{G}_n \leq R_n,\ n \in N_0.$$

In [2] the last class was denoted by $(D)S_3$. It is easy to prove that

(10) $$(D)NBUE \subset (D)HNBUE.$$

First of all we will show that if $(p_n) \in (D)HNBUE$, then it has finite each moment. By inclusion (10) the same is true for distribution $(p_n) \in (D)NBUE$.

LEMMA 1. *If the distribution $(p_n) \in (D)NHBUE$, then it has finite all moments and they are bounded by geometric distribution moments with parameter p, namely*

(11) $$m_v \leq \mu^2 \gamma_v,\ v = 1, 2, \ldots,$$

where γ_v is the v-th moment of the geometric distribution with parameter $p = \frac{1}{\mu}$.

Proof. By inequality (8) we have $(v \in N)$

$$m_v = \sum_{k=1}^{\infty} k^v p_k \leq \sum_{k=1}^{\infty} k^v R_{k-1} \leq \mu \sum_{k=1}^{\infty} k^v \overline{G}_{k-1} \leq$$

$$\leq \mu \sum_{k=1}^{\infty} k^v q^{k-1} = \mu^2 \sum_{k=1}^{\infty} k^v p(1-p)^{k-1} = \mu^2 \gamma_v.$$

Because every geometric distribution has finite moments of any order, the same holds for the distribution $(p_n) \in (D)HNBUE$.

For $v = 1$ the estimation (11) is trivial ($m_1 = \mu \le \mu^3$).

Remark 1. Let us note that, if $(p_n) \in (D)NBUE$, then $\Delta_n \ge 0, n \in N_0$. In [2] it was proved, that

$$\alpha := \sum_{k=0}^{\infty} \Delta_n = \frac{1}{2\mu}(\mu^2 - \mu - \sigma^2), \tag{12}$$

where σ is the standard deviation of distribution (p_n). From (12) it follows $\sigma^2 \le \mu^2 - \mu$.

Hence the second moment

$$m_2 \le 2\mu^2 - \mu. \tag{13}$$

By Lemma 1 we obtain only that

$$m_2 \le \mu^3(2\mu - 1). \tag{14}$$

Now we prove the following lemma.

LEMMA 2. *If* $(p_n) \in (D)NBUE$, *then*

$$R_n - q^n = \Delta_n - \frac{p}{q}\sum_{k=0}^{n-1} \Delta_k q^{n-k}. \tag{16}$$

Proof. By (2) and (3) we get

$$\Delta_n = R_n - \overline{G}_n = \frac{1}{p}(\overline{G}_n - \overline{G}_{n+1}) - \overline{G}_n = \frac{q}{p}\overline{G}_n - \frac{1}{p}\overline{G}_{n+1}.$$

Hence

$$\overline{G}_{n+1} = q\overline{G}_n - p\Delta_n$$

then we have

$$\overline{G}_n = q^n - p\sum_{k=0}^{n-1} \Delta_k q^{n-k-1},$$

and

$$\overline{G}_{n+1} = q^{n+1} - p\sum_{k=0}^{n} \Delta_k q^{n-k}.$$

Therefore,

$$R_n = \frac{1}{p}(\overline{G}_n - \overline{G}_{n+1}) = q^n - \frac{p}{q}\sum_{k=0}^{n-1} \Delta_k q^{n-k} + \Delta_n,$$

which implies (16).

LEMMA 3. *If* $(p_n) = (D)NBUE$, *then*

$$R_n - q^n = \sum_{k=1}^{n} q^{n-k}(\Delta_k - \Delta_{k-1}).$$

Proof. According to the well known Abel's transformation (see [3])

$$\sum_{k=1}^{n-1} \Delta_k q^{-k} = \sum_{k=0}^{n-1} \Delta_k q^{-k} = \Delta_n \sum_{k=0}^{n-1} q^{-k} - \sum_{k=0}^{n-1} \Big[(\Delta_{k+1} - \Delta_k)\sum_{r=0}^{k} q^{-r}\Big]$$

$$= \Delta_n \frac{1-q^{-n}}{1-q^{-1}} - \sum_{k=1}^{n} \Big[(\Delta_k - \Delta_{k-1})\sum_{r=0}^{k-1} q^{-r}\Big]$$

$$= -\frac{q}{p}\Delta_n(1-q^{-n}) - \frac{q}{p}\sum_{k=1}^{n}(\Delta_k - \Delta_{k-1})(q^{-k} - 1).$$

From this

$$-\frac{p}{q}\sum_{k=1}^{n-1} \Delta_k q^{n-k} = \Delta_n(q^n - 1) + \sum_{k=1}^{n}(q^{n-k} - q^n)(\Delta_k - \Delta_{k-1})$$

$$= -\Delta_n + \sum_{k=1}^{n} q^{n-k}(\Delta_k - \Delta_{k-1}).$$

Therefore by Lemma 2 we get (17).

THEOREM 1. *If* $(p_n) \in (D)NBUE$, *then*

$$|R_n - q^n| \le \alpha. \tag{18}$$

Proof. By Lemma 3 we get

$$R_n - q^n \le \sum_{k=1}^{n} q^{n-k}\Delta_k \le \sum_{k=1}^{n} \Delta_k \le \sum_{k=0}^{\infty} \Delta_k = \alpha.$$

By the other hand

$$-R_n + q^n \geq -\sum_{k=1}^{n} q^{n-k}\Delta_{k-1} \geq -\sum_{k=0}^{n}\Delta_{k-1} = -\sum_{k=0}^{n-1}\Delta_k \geq -\sum_{k=0}^{\infty}\Delta_k = -\alpha.$$

The theorem is proved.

Remark 2. Theorem 1 gives better estimation of the error of the approximation $|R_n - q^n|$ for unknown distribution $(p_n) \in (D)NBUE$ than the results got in [2], where this error was estimated by $\sqrt{2\alpha}$ and 2α, respectively.

Remark 3. The estimation given in Theorem 1 is sharp, in the sense that it is the best for whole class. The following example shows this fact.

EXAMPLE 1. Consider two points distribution (p_n): $p_1 = r$, $(0 < r < 1)$, $p_2 = 1 - r$, and $p_k = 0$, if $k > 2$. Then, $R_0 = 1$, $R_1 = 1 - r$, $R_2 = R_3 = \ldots = 0$, $\mu = 2 - r > 1$, $\overline{G}_0 = 1$, $\overline{G}_1 = \frac{1-r}{2-r}$, $\overline{G}_2 = \overline{G}_3 = \ldots = 0$. Hence, the distribution belongs to the class (D)NBUE, and

$$\Delta_0 = 0,\ \Delta_1 = \frac{(1-r)^2}{2-r},\ \Delta_2 = \Delta_3 = \ldots = 0.$$

Moreover $m_2 = 4 - 3r$, $\alpha = \frac{(1-r)^2}{2-r}$,

$$R_0 - q^0 = 0,$$

Since

$$R_1 - q = 1 - r - \frac{1-r}{2-r} = \alpha,$$

and

$$|R_n - q^n| = \left(\frac{1-r}{2-r}\right)^n \text{ if } n > 1.$$

We obtain

$$\sup_{n\geq 0}|R_n - q^n| = \frac{(1-r)^2}{2-r} = \alpha.$$

Hence the estimation is sharp.

THEOREM 2. *Let $((p_n)_m; n, m \in N)$ be a sequence of discrete distribution from the class $(D)NBUE$. The limit of the sequence $((p_n)_m)$ is a geometric* *distribution if and only if*

$$\lim_{m\to\infty} \alpha_m = 0, \tag{19}$$

where

$$\alpha_m = \mu - \sum_{n=0}^{\infty} (\overline{G}_n)_m. \tag{20}$$

Proof. *

(i) According to Theorem 1 if $\alpha_m \to 0$ for every $n \geq 0$, then

$$\lim_{m\to\infty} |(R_n)_m - q_m^n| = 0,$$

i.e.

$$\lim_{m\to\infty} (R_n)_m = \lim_{m\to\infty} q_m^n =: q^n.$$

It means that

$$\lim_{m\to\infty} \sum_{k=n+1}^{\infty} (p_n)_m = q^n.$$

Because the series are convergent (by assumption)

$$\sum_{k=n+1}^{\infty} \lim_{m\to\infty} (p_k)_m = q^n.$$

Now denote

$$\widehat{p}_k = \lim_{m\to\infty} (p_k)_m, k = 1, 2, \ldots$$

Since

$$\sum_{k=n+1}^{\infty} \widehat{p}_k = q^n,$$

we have

$$\widehat{p}_n = \sum_{k=n}^{\infty} \widehat{p}_k - \sum_{k=n+1}^{\infty} \widehat{p}_k = q^{n-1} - q^n = (1-q)q^{n-1}.$$

So the distribution is geometric.

---

*The problems of convergence of distribution are considered e.g. in [4].

(ii) Let the sequence $((p_n)_m)$ of distributions tends to geometric distribution with parameter $p = \frac{1}{\mu}$. Then for arbitrary $n \in N_0$

$$\lim_{m\to\infty}(R_n)_m = q^n, q = 1 - p.$$

As $((p_n)_m) \in (D)NBUE$

$$\lim_{m\to\infty}\mu_m = \lim_{m\to\infty}\sum_{n=0}^{\infty}(R_n)_m = \sum_{n=0}^{\infty}\lim_{m\to\infty}(R_n)_m = \sum_{n=0}^{\infty}q^n = \mu.$$

Because of the property of the class (D)NBUE

$$(\overline{G}_n)_m \leq (R_n)_m, n \geq 0.$$

The serie $\sum_{n=0}^{\infty}(\overline{G}_n)_m$ has a convergent majorant

$$\sum_{n=0}^{\infty}(R_n)_m = \sum_{n=0}^{\infty}q_m^n = \frac{1}{1-q_m}\mu_m.$$

Hence

$$\lim_{m\to\infty}\sum_{n=0}^{\infty}(\overline{G}_n)_m = \sum_{n=0}^{\infty}\lim_{m\to\infty}(\overline{G}_n)_m$$

$$= \sum_{n=0}^{\infty}(\lim_{m\to\infty}p_m)\sum_{k=n}^{\infty}\lim_{m\to\infty}(R_n)_m = p\sum_{n=0}^{\infty}\frac{q^n}{1-q} = \mu.$$

Therefore

$$\lim_{m\to\infty}\alpha_m = \lim_{m\to\infty}\mu_m - \lim_{m\to\infty}\sum_{n=0}^{\infty}(\overline{G}_n)_m = 0.$$

The proof is completed.

EXAMPLE 2. Consider the sequence of two-points distribution for $m = 1, 2, \ldots$

$$(p_1)_m = \frac{m}{m+1}, \ (p_2)_m = \frac{1}{m+1}, \ (p_k)_m = 0, k > 2$$

belonging to the class (D)NBUE. We have

$$\mu_m = \frac{m+2}{m+1} > 1, \ (m_2)_m = \frac{m+4}{m+1}.$$

Therefore

$$\alpha_m = \frac{1}{m^2 + 3m + 1}.$$

If $m \to \infty$ then $\alpha_m \to 0$.

On the other hand in this case $(p_1)_m \to 1$, and $(p_2)_m \to 0$. So we obtain (degenerate) geometric distribution with the parameter $p = 1$.

References

[1] M. Becheng, Ch. Kan, *Characterizations of geometric distributions based on moments*, Acta Math. Appl. Sinica, 8 (1992), 168–181.

[2] D. Bobrowski, B. Lutomska, *On approximation of the discrete life time distribution by the geometric distribution*, Discuss. Math., Algebra and Stochastic Methods, 16 (1996), 197–203.

[3] V. L. Kocic and G. Ladas, *Global behavior of nonlinear difference of higher order with applications*, Kluwer Academic Publishers, Dordrecht, 1993.

[4] D. Pollard, *Convergence of Stochastic Processes*, Springer Vlg., New York, 1984.

Dobiesław Bobrowski
DEPARTMENT OF MATHEMATICS
AND COMPUTER SCIENCE
A. MICKIEWICZ UNIVERSITY
J. Matejki 48/49
60-769 POZNAŃ, POLAND

Barbara Popowska
INSTITUTE OF MATHEMATICS
TECHNICAL UNIVERSITY OF POZNAŃ
Piotrowo 3a
60-965 POZNAŃ, POLAND

Received August 8, 1996.