FASCICULI MATHEMATICI

Nr 28 1998

DEDICATED TO PROFESSOR DOBIESŁAW BOBROWSKI
ON THE OCCASION OF HIS 70[th] BIRTHDAY.

BARBARA POPOWSKA

GEOMETRIC APPROXIMATION OF MIXTURE OF DISCRETE LIFE TIME DISTRIBUTIONS

ABSTRACT. We consider a mixture of a distribution from the class (D)NBUE and a distribution, which describes the life time of small part of lower reliability elements. Given the first two moments of this mixture, the distribution can be approximated by a geometric distribution with some parametr p. An error of such approximation is also estimated.

KEY WORDS. mixture of discrete distributions, geometric distribution, error of approximation.

In two previous papers ([1], [2]) the possibility of approximation of unknow life time discrete distribution by the geometric distribution was studied. It was supposed that the unknown distribution belongs to the class (D)NBUE (discrete new better than used in expectation). In this paper we consider the same problem in a case when the unknown distribution is a mixture of a distribution from a given class (D)NBUE and other distribution. The second distribution describes the life time of rather small part of lower reliability elements e.g. elements stoped working on the beginning of exploatation time.

Let $(p_n) = (p_n, n \in N)$ be a mixture of independent discrete probability distributions (p_n') and (p_n'') i.e.

(1) $$p_n = (1-\delta)\, p_n' + \delta\, p_n'', \qquad n \in N,$$

where $\delta \in (0,1)$. We will assume that the mixture (p_n) has a finite first and second moment $\mu > 1$, $\mu_2 > 1$, respectively.

In this paper (p_n') and (p_n'') stands for the life time distribution of high reliabity elements and foulty elements fraction, respectively. For the mixture we introduce the folowing notation for all $n \geq 0$:

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

and

$$\bar{G}_n = \frac{1}{\mu} \sum_{k=n}^{\infty} R_k .$$

R_n is call residual probability.

Clearly

$$R_n = (1-\delta)R_n' + \delta R_n'',$$

where

$$R_n' = \sum_{k=n+1}^{\infty} p_k' \qquad \text{and} \qquad R_n'' = \sum_{k=n+1}^{\infty} p_k''$$

is a residual probability for high reliability elements fraction and foulty elements fraction, respectively.

Similarly denoting by μ' the expected value of life time of good elements fraction and μ'' the expected value of life time of foulty elements fraction we have

$$\mu = (1-\delta)\mu' + \delta\mu''.$$

For considered mixture we introduce following asumption for all $n \geq 0$

(2) $$R_n \leq R_n'.$$

Under our assumption that fraction of high reliability elements belongs to (D)NBUE class so from the definition of such a class the inequality

(3) $$\bar{G}_n' = \frac{1}{\mu'}\sum_{k=n}^{\infty} R_k' \leq R_n'$$

holds.

From the definition and asumptions (2) and (3) for our mixture we get following condition

$$\frac{1}{\mu}\sum_{k=n}^{\infty} R_k \leq \frac{1}{(1-\delta)^2} R_n.$$

Let $$\varepsilon = \varepsilon(\delta) = \frac{\delta(2-\delta)}{(1-\delta)^2} > 0 \quad \text{for} \quad \delta \in (0,1),$$

and let

$$\Delta_n = (1+\varepsilon)R_n - \bar{G}_n.$$

Then clearly

$$\bar{G}_n \leq (1+\varepsilon)R_n,$$

and

(i) $\Delta_n \geq 0,$

(ii) $\Delta_0 = \varepsilon > 0,$

(iii) $\lim_{n\to\infty} \Delta_n = 0.$

We begin with the following lemma

Lemma 1. Let the distribution (p_n) *be a mixture of distributions defined by* (1). *In this case*

(4) $$\sum_{n=0}^{\infty}\Delta_n=\frac{\mu}{(1-\delta)^2}-\frac{\mu_2}{2\mu}-\frac{1}{2}=:\alpha_\delta$$

Proof. We have

$$\sum_{n=0}^{\infty}\Delta_n=(1+\varepsilon)\sum_{n=0}^{\infty}R_n-\frac{1}{\mu}\sum_{n=0}^{\infty}\sum_{k=n}^{\infty}R_k=(1+\varepsilon)\sum_{n=0}^{\infty}R_n-\frac{1}{\mu}\sum_{k=0}^{\infty}(k+1)R_k=$$

$$=(1+\varepsilon)\mu-\frac{1}{\mu}\sum_{k=0}^{\infty}kR_k-1=(1+\varepsilon)\mu-1-\frac{1}{2\mu}\sum_{n=0}^{\infty}(n^2-n)p_n=$$

$$=(1+\varepsilon)\mu-1-\frac{1}{2\mu}(\mu_2-\mu)=\frac{\mu}{(1-\delta)^2}-\frac{\mu_2}{2\mu}-\frac{1}{2}$$

and the proof is completed.

Our next result is the following technical lemma about the residual probability R_n.

Lemma 2. If (p_n) *is a mixture of distributions defined by* (1) *then*

(5) $$R_n=\frac{1}{1+\varepsilon}q^n\left[1-\frac{p}{q}\sum_{k=0}^{n-1}q^{-k}\Delta_k\right]+\frac{1}{1+\varepsilon}\Delta_n,\quad n\geq 0,$$

where

(6) $$p=\frac{1}{(1+\varepsilon)\mu},\qquad \text{and}\qquad q=1-p.$$

Proof. By the definitions

$$\Delta_n=[(1+\varepsilon)\mu-1]\bar{G}_n-(1+\varepsilon)\mu\bar{G}_{n+1}.$$

Therefore

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

From this recurence relation we obtain

$$\bar{G}_n=q^n\left[1-\frac{p}{q}\sum_{k=0}^{n-1}q^{-k}\Delta_k\right].$$

Consequently

$$\bar{G}_n-\bar{G}_{n+1}=q^np-q^{n-1}p^2\sum_{k=0}^{n-1}q^{-k}\Delta_k+p\Delta_n.$$

Hence

$$\frac{1}{\mu}R_n = q^n p\left[1-\frac{p}{q}\sum_{k=0}^{n-1} q^{-k}\Delta_k\right] + p\Delta_n$$

which implies (5).

The proof of the next theorem relies on the following formula for $R_n - q^n$.

Lemma 3. If (p_n) *is a mixture of distributions defined by* (1) *then*

$$(1+\varepsilon)(R_n - q^n) = \sum_{k=1}^{n} q^{n-k}(\Delta_k - \Delta_{k-1}). \tag{7}$$

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

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

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

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

Hence

$$\frac{p}{q}\sum_{k=0}^{n-1}\Delta_k q^{-k} = -(1-q^{-n})\Delta_n + (\Delta_n - \Delta_0) - \sum_{k=1}^{n} q^{-k}(\Delta_k - \Delta_{k-1}) =$$

$$= q^{-n}\Delta_n - \Delta_0 + \sum_{k=1}^{n} q^{-k}(\Delta_k - \Delta_{k-1}).$$

Therefore by Lemma 2 we get (7).

Now we are ready to prove the main result of this paper.

Theorem. Let the discrete distribution (p_n) *be a mixture of distributions defined by* (1). *Then there exists* $q = 1 - \frac{1}{(1+\varepsilon)\mu} \in (0,1)$ *such that we get the following estimation*

$$|R_n - q^n| \le \beta_\delta, \tag{8}$$

where

$$\beta_\delta = (1-\delta)^2\,\alpha_\delta, \tag{9}$$

and α_δ *defined by Lemma* 1.

Proof. From Lemma 3 we get

$$R_n - q^n \le \frac{1}{1+\varepsilon}\sum_{k=1}^{n} q^{n-k}\Delta_k \le \frac{1}{1+\varepsilon}\sum_{k=0}^{\infty}\Delta_k = \frac{\alpha_\delta}{1+\varepsilon} = \beta_\delta.$$

On the other hand

$$R_n - q^n \ge \frac{1}{1+\varepsilon}\sum_{k=1}^{n} q^{n-k}\Delta_{k-1} \ge -\frac{1}{1+\varepsilon}\sum_{k=1}^{n}\Delta_{k-1} \ge$$

$$\ge -\frac{1}{1+\varepsilon}\sum_{k=0}^{\infty}\Delta_k = -\frac{\alpha_\delta}{1+\varepsilon} = -\beta_\delta.$$

The theorem is proved.

Remark. If $\varepsilon = 0$ then we get

$$|R_n - q^n| \le \alpha_0,$$

i.e. the result for clean population without foulty elements (see [1]).

Example. Consider the inflateted geometric distribution

$$p_n = \begin{cases} (1-\delta)a + \delta, & n = 1, \\ (1-\delta)a(1-a)^{n-1}, & n \ge 2. \end{cases}$$

i.e. the distribution (p_n) given by (1), where

$$p'_n = a(1-a)^{n-1}, \qquad p''_n = \begin{cases} 1, & n = 1, \\ 0, & n = 2,3..., \end{cases}$$

and $\delta \in (0,1)$.

Then

$$R'_n = (1-a)^n, \qquad R''_n = \begin{cases} 1, & n = 0, \\ 0, & n = 1,2..., \end{cases}$$

$$\bar{G}'_n = (1-a)^{n+1}, \qquad \bar{G}''_n = \begin{cases} 1, & n = 0, \\ 0, & n = 1,2..., \end{cases}$$

and

$$\mu' = \frac{1}{a}, \qquad \mu'' = 1,$$

$$\sigma^{2'} = \frac{1-a}{a^2}, \qquad \sigma^{2''} = 0.$$

Hence

$$R_n = \begin{cases} 1, & n = 1, \\ (1-\delta)(1-a)^n, & n = 1,2,..., \end{cases}$$

$$\bar{G}_n' = \begin{cases} (1-\delta)(1-a)+\delta, & n=0, \\ (1-\delta)(1-a)^{n+1}, & n=1,2,..., \end{cases}$$

and

$$\mu = \frac{1-\delta+\delta a}{a}, \qquad \sigma^2 = \frac{1-a-\delta+\delta a}{a^2}.$$

In this case we have by (4) and (9) that

$$\beta_\delta := \frac{(1+2\varepsilon)(1-\delta+\delta a)^2 - \delta a^2 + \delta - 1}{a(1+\varepsilon)(2-2\delta+2\delta a)}.$$

But by $\varepsilon = \frac{\delta(2-\delta)}{(1-\delta)^2}$, we have

$$\beta_\delta = \frac{(3+2a-a^2)\delta + (-7+2a+3a^2)\delta^2 + (5-6a+a^2)\delta^3 - (1-2a+a^2)\delta^4}{2a(1-\delta)^2 - (1-\delta+\delta a)},$$

omiting the terms with higher powers of δ we obtain the estimate

$$\beta_\delta \approx \frac{(3+2a-a^2)\delta}{2a(1-3\delta+\delta a)}.$$

For instance it $a = 0.9$ and $\delta = 0.01$ then the error of approximation $\beta_\delta = 0.0226$.

ACKNOWLEGEMENST: *I would like to thank to Professor Dobiesław Bobrowski for very many helpful comments and inspirational discussions.*

REFERENCES

[1] D. Bobrowski, B. Lutomska, On Approximation of the Discrete Life Time Distribution by the Geometric Distribution, *Discussiones Mathematicae Algebra and Stochastic Methods*, 16(1996), 1-7.

[2] D. Bobrowski, B. Popowska, The Estimate of the Error for Approximation of Some Discrete Distribution by the Geometric Distribution, *Demonstratio Mathematica*, 30(3)(1997), (in press).

[3] S.N. Singh, A Note on Inflantde Poisson Distribution, *J. Indian Statist. Assoc.*, 1(3)(1963), 140-144.

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

(Institute of Mathematics, Poznan University of Thechnology, 60-965 Poznań, Poland)

Received on 13.02.1997 and, in revised form, on 10.12.1997.