ISSN 2347-1921
Volume 14 Number1
Journal of Advances in Mathematics

Functional moments estimators analysis by the Monte-Carlo method for model of mixture with varying concentrations

Oksana Kubaychuk
Department of Applied Mathematics, National Aviation University, Kyiv, Ukraine
kubaychuk@gmail.com

ABSTRACT

The functional moments estimation by the sample from the mixture with varying concentrations is studied. The problem of efficiency the simple linear estimator with fixed weight against the adaptive or improved estimators with random weight is considered. By the Monte-Carlo method it is shown that simple linear estimator is better for small sample sizes, but for large samples the adaptive and improved estimators are more efficient.

Indexing terms/Keywords

Monte-Carlo method; mixture with varying concentrations; linear estimator; adaptive estimator; improved estimator

Academic Discipline And Sub-Disciplines

Mathematics, Probability and Statistics

SUBJECT CLASSIFICATION

62G05, 62G20, 60G15

INTRODUCTION

Often, to describe the statistical data the model of a mixture with varying concentrations is used. A sample $\Xi_N = (\xi_{j:N}, j = 1,\dots,N)$ in this model [1] consists of jointly independent random variables with distributions

$$\Pr\{\xi_{j:N} < x\} = \sum_{m=1}^{M} w_{j:N}^{m} H_m(x),$$

where M is the total number of components in the mixture, H_m is the distribution function of the m-th component, and $w_{j:N}^{m}$ is the probability to observe an object from the m-th component in the j-th observation. Probability $w_{j:N}^{m}$ is called concentration or mixing probability. We assume that concentrations of components are known, and the distributions H_m are unknown.

1. SETTING OF THE PROBLEM

Let consider some functional moment for the k-th component of the mixture, namely

$$\bar{g}^k = \int g(x) H_k(dx)$$

where the real valued function g is fixed.

1.1 Linear and adaptive estimators

The problem of estimation of functional moment studied in [2]. One can consider the linear estimator for $\bar{g}^k$ becomes of the form

$$\hat{g}_N^k = \hat{g}_N(\vec{a}^k) = \int g(x) \hat{F}_N(dx, \vec{a}^k) = \frac{1}{N}\sum_{j=1}^{N} a_j^k g(\xi_j) \tag{1}$$

where

$$\hat{F}_N(x, \vec{a}) = \frac{1}{N}\sum_{j=1}^{N} a_j \mathbf{1}\{\xi_j < x\} \tag{2}$$

is the weighted empirical distribution function with some nonrandom weight $\vec{a}$.

The estimator, defined (1), is unbiased, consistent, asymptotically normal however, generally speaking, not efficient. It is proposed in [2] to use the adaptive method of estimation. The adaptive estimator $\tilde{g}_N^k = \hat{g}_N(\vec{a}^k(\hat{\lambda}_N))$ is the estimator (1) with adaptive random weights $\vec{a}^k(\hat{\lambda}_N)$.

1.2 Improved estimators

It can be that some coefficients a_j in (2) are negative, then the weighted distribution function is not nondecreasing, and therefore it isn't a probability distribution function. Put

$$F_N^+(x,\vec{a}) \triangleq \sup_{y<x} \hat{F}_N(y,\vec{a}). \quad (3)$$

The function $F_N^+(x,\vec{a})$ is nondecreasing and assumes only positive values, but it can assumes values greater then 1. Thus we consider the function

$$\Phi_N^+(x,\vec{a}) = \min(1, F_N^+(x,\vec{a})) . \quad (4)$$

Accordingly, the estimators for functional moments becomes of the form

$$\tilde{g}_N^{+k} = \int g(x)\Phi_N^+(dx,\vec{a}^k) = \frac{1}{N}\sum_{j=1}^{N} b_j^{+k} g(\xi_j), \quad (5)$$

where b_j^{+k} are some coefficients that depend on the sample Ξ_N. To obtain the improved distribution function $\Phi_N^+(x,\vec{a})$ (namely, coefficients b_j^{+k}) we can use the algorithm from [3].

Similarly, we can construct an other estimators for functional moments.For example, Put

$$F_N^-(x,\vec{a}) \triangleq \inf_{y>x} \hat{F}_N(y,\vec{a}), \quad (6)$$

$$\Phi_N^-(x,\vec{a}) = \max(0, F_N^-(x,\vec{a})) , \quad (7)$$

and the combination of (4), (7), for example

$$\Phi_N^{\pm}(x,\vec{a}) = \frac{1}{2}[\Phi_N^+(x,\vec{a}) + \Phi_N^-(x,\vec{a})]. \quad (8)$$

And, accordingly

$$\tilde{g}_N^{-k} = \int g(x)\Phi_N^-(dx,\vec{a}^k) = \frac{1}{N}\sum_{j=1}^{N} b_j^{-k} g(\xi_j), \quad (9)$$

$$\tilde{g}_N^{\pm k} = \int g(x)\Phi_N^{\pm}(dx,\vec{a}^k) = \frac{1}{N}\sum_{j=1}^{N} b_j^{\pm k} g(\xi_j). \quad (10)$$

where b_j^{-k}, $b_j^{\pm k}$ are some coefficients that depend on the sample Ξ_N. To obtain the coefficients b_j^{-k}, $b_j^{\pm k}$ we can use the algorithms from [4].

2. SIMULATION RESULTS

2.1 Two component mixture model

The aim of present paper is studying the efficiency of the simple linear $\hat{g}_N^k$ against $\tilde{g}_N^k$, $\tilde{g}_N^{+k}$, $\tilde{g}_N^{-k}$, $\tilde{g}_N^{\pm k}$ adaptive and improved estimators by Monte-Carlo method. We used the two-component mixture with simple linear concentrations for estimators analysis. Namely, $M = 2$ and

$$w_{j:N}^1 = \frac{j}{N}, w_{j:N}^2 = 1 - \frac{j}{N},\ j = 1,\dots,N.$$

The distributions of the components were normal, uniform and Pareto's. Let, $g(x) = x$. The mean square errors $\mathrm{MSE}(\hat{g}_N^k)$ and $\mathrm{MSE}(\tilde{g}_N^k)$ of the estimators were estimated. The relative efficiency

$$\mathrm{RE_k} = \frac{\mathrm{MSE}(\hat{g}_N^k)}{\mathrm{MSE}(\tilde{g}_N^k)},\ k = 1, 2 \qquad (11)$$

were calculated for adaptive estimators. The formulas to obtain the relative efficiency RE_k^+, RE_k^-, $\mathrm{RE}_k^\pm$ for improved estimators are similar.

2.2 Additional transformations

All computations were made in Mathcad v.7. But it can't to generate a vector of random numbers having the Pareto's distributions. To solve this problem, we use the built-in function *rbeta* to generate a vector of random numbers having the beta distribution and transformation of the random variable.

Really, if the random variable ξ has the beta distribution with parameters $\gamma > 0$ and 1, then the random variable $\eta = \frac{1}{\xi}$ has Pareto's distribution with parameter $\gamma > 0$.

$$F_\eta(t) = \Pr\{\eta < t\} = \Pr\left\{\frac{1}{\xi} < t\right\} = \Pr\left\{\xi > \frac{1}{t}\right\} = 1 - F_\xi\left(\frac{1}{t}\right).$$

Hence,

$$p_\eta(t) = p_\xi\left(\frac{1}{t}\right)\cdot\frac{1}{t^2} = \frac{1}{\mathrm{B}(m,n)}\cdot\frac{1}{t^2}\cdot\frac{1}{t^{m-1}}\cdot\left(1-\frac{1}{t}\right)^{n-1} =$$

$$= \frac{1}{\mathrm{B}(m,n)}\cdot\frac{1}{t^2}\cdot\frac{(t-1)^{n-1}}{t^{m-1+n-1}} = \frac{1}{\mathrm{B}(m,n)}\cdot\frac{(t-1)^{n-1}}{t^{m+n}} = \frac{1}{\mathrm{B}(\gamma,1)}\cdot\frac{1}{t^{\gamma+1}} =$$

$$= \frac{\Gamma(\gamma+1)}{\Gamma(\gamma)\cdot\Gamma(1)}\cdot\frac{1}{t^{\gamma+1}} = \gamma\cdot\frac{1}{t^{\gamma+1}}, 1 < t < \infty\left(0 < \frac{1}{t} < 1\right)$$

where $\mathrm{B}(m,n)$, $\Gamma(n)$ are the beta and gamma-functions.

The Pareto's density function is

$$p(x) = \begin{cases} \gamma\cdot\frac{1}{x^{\gamma+1}}, 1 < x < \infty, \\ 0, x \le 1, \end{cases},\ \gamma > 0.$$

Also, remind that $\mathbf{E}\eta < \infty$ if $\gamma > 1$, and $\mathbf{E}\eta^2 < \infty$ if $\gamma > 2$.

2.3 Adaptive estimators simulation

The relative efficiency (11) for two-component mixture is displayed for sample size N for L simulated samples at the Fig.1-3.

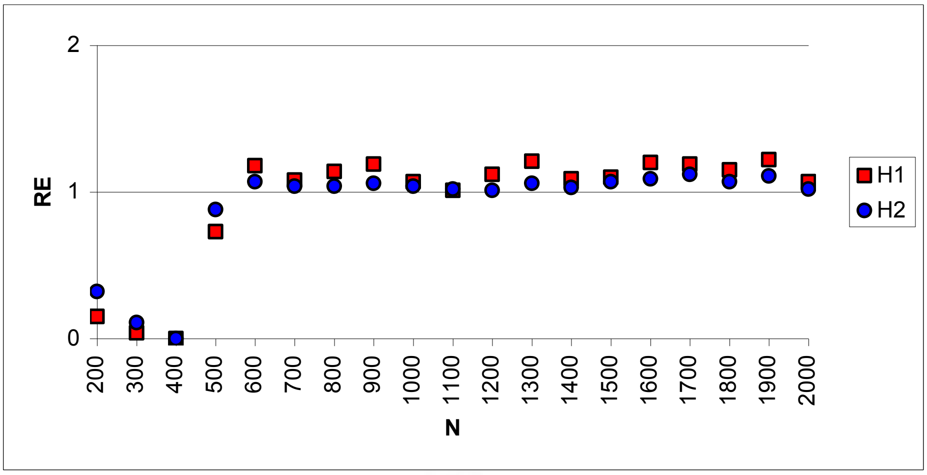

Fig 1: $H_1 \sim N(0,1)$, $H_2 \sim N(1,4)$, $L = 200$

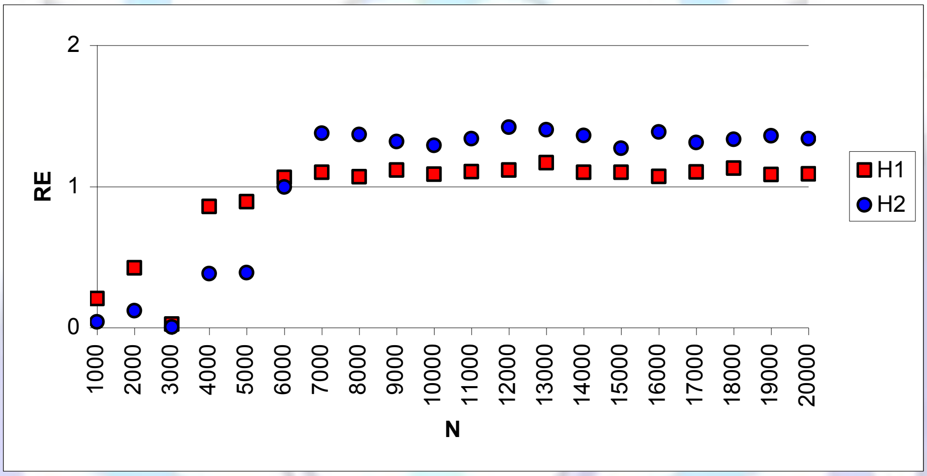

Fig 2: $H_1 \sim N(0,1)$, $H_2 \sim U_{2,4}$, $L = 1000$

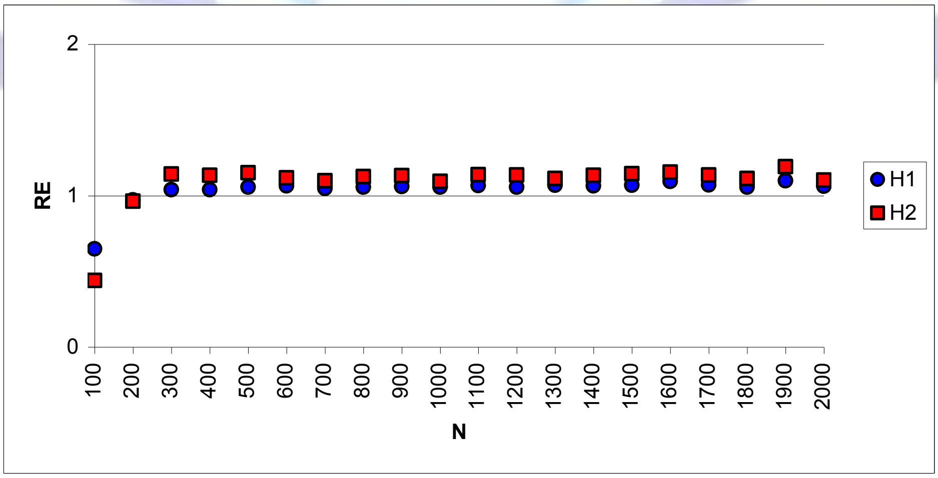

Fig 3: $H_1 \sim U_{1,3}$, $H_2 \sim U_{2,4}$, $L = 1000$

In case one or both components have Pareto's distributions we can't take an results as above. For example, it is no sense to construct the adaptive estimator if $\gamma \le 2$. The simulation results are established in Table 1 when the first component has the normal distribution and the second component has Pareto's distribution with parameter $\gamma = 3$ for $L = 100$ simulated samples.

Table 1. The component distributions are normal and Pareto.

N	MSE		RE_1	MSE		RE_2
	$\hat{\mathbf{g}}_N^1$	$\tilde{\mathbf{g}}_N^1$		$\hat{\mathbf{g}}_N^2$	$\tilde{\mathbf{g}}_N^2$	
1E+3	5.133E-3	0.032	0.159	4.64E-3	0.023	0.2
2E+3	2.104E-3	2.003E-3	1.051	2.054E-3	1.858E-3	1.106
3E+3	1.493E-3	0.018	0.083	1.196E-3	0.014	0.085
4E+3	1.063E-3	7.726E-3	0.138	1.262E-3	8.5E-3	0.149
5E+3	8.772E-4	9.203E-4	0.953	7.699E-4	7.157E-4	1.076
6E+3	9.668E-4	44.554	2.17E-5	8.264E-4	29.86	2.768E-5
7E+3	7.107E-4	6.699E-4	1.061	5.769E-4	5.027E-4	1.148
8E+3	5.784E-4	5.384E-4	1.074	5.748E-4	5.009E-4	1.148
9E+3	7.288E-4	0.021	0.035	5.461E-4	0.014	0.038
1E+4	4.271E-4	4.158E-4	1.027	5.289E-4	4.949E-4	1.069
1.1E+4	3.595E-4	1.515E-3	0.237	3.399E-4	1.006E-3	0.338
1.2E+4	4.26E-4	4.194E-4	1.016	3.528E-4	3.587E-4	0.984
1.3E+4	3.416E-4	0.059	5.829E-3	4.276E-4	0.048	8.89E-3
1.4E+4	4.012E-4	3.946E-4	1.017	2.87E-4	2.81E-4	1.021
1.5E+4	4.485E-4	0.042	0.011	3.071E-4	0.029	0.011
1.6E+4	2.861E-4	2.869E-4	0.997	2.676E-4	2.594E-4	1.032
1.7E+4	2.992E-4	0.034	8.732E-3	2.783E-4	0.024	0.012
1.8E+4	3.27E-4	3.26E-4	1.003	3.123E-4	3.049E-4	1.024
1.9E+4	2.518E-4	2.458E-4	1.024	2.373E-4	2.314E-4	1.026
2E+4	1.925E-4	1.893E-4	1.017	1.905E-4	1.848E-4	1.031

As shown from Table. 1, the estimators $\hat{g}_N^k$ and $\tilde{g}_N^k$, $k = 1,2$ have a good statistical property, but we can't fix the sample size N when the estimator $\tilde{g}_N^k$ is better then $\hat{g}_N^k$. Perhaps, the sample size may be too large.

2.4 Simulation of imroved estimators

The relative efficiency RE_k, RE_k^+, RE_k^-, $\mathrm{RE}_k^\pm$ for two-component mixture is displayed for sample size N for $L = 100$ simulated samples at the Fig.4-5. Namely, the relative efficiency of the first component is shown at the Fig. 4, second component is shown at the Fig. 5.

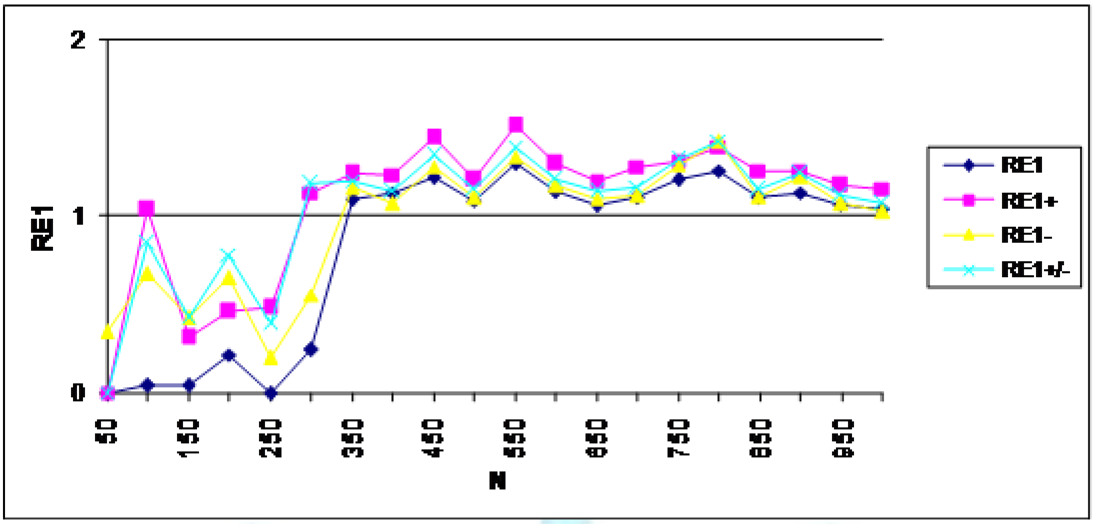

Fig 4: $H_1 \sim N(0,1)$, $H_2 \sim N(1,4)$, $L = 100$, $k = 1$

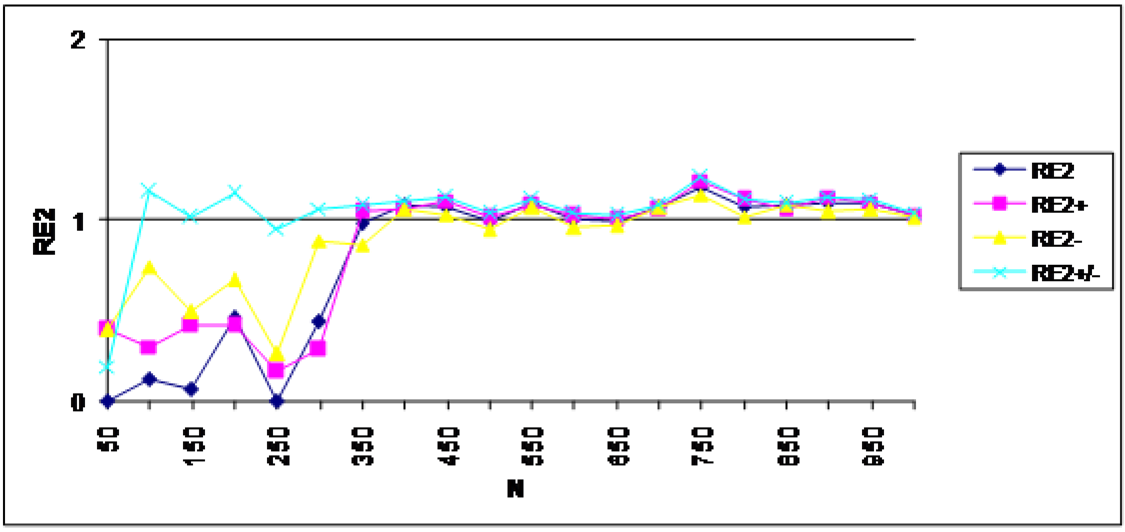

Fig 5: $H_1 \sim N(0,1)$, $H_2 \sim N(1,4)$, $L = 100$, $k = 2$

CONCLUSIONS

The simulation results demonstrates that the mean square error of the simple linear estimator is less for small sample sizes, however for large samples the adaptive and improved estimators are more efficient. It is no sense to use the adaptive estimator, when one of distributions are Pareto's.

REFERENCES

1. R. E. Maıboroda (1996) Estimation of the distributions of the components of mixtures having varying concentrations, Ukrain. Matem. Zh. 48, no. 4, 562–566; English transl. in Ukrainian Math. J. 48 (1996), no. 4, 618–622.
2. Kubaychuk O. O. (2002) Estimation of moments by observations from mixtures with varying concentrations // Theory of Stochastic Processes. Vol. 8(24), no.3–4.– P. 226 – 232.
3. R. Maıboroda and O. Kubaıchuk (2003) Asymptotic normality of improved weighted empirical distribution functions, Teor. Imovirnost. ta Matem. Statist. 69, 89–95; English transl. In Theor. Probability and Math. Statist. 69 (2004), 95–102.
4. Kubaychuk, O.O. (2003): Estimation of moments from mixtures using the corrected weighted empirical distribution functions. *Visnyk KNU, Ser. Matematika. Mekhanika*, 9, 48–52. (Ukrainian).