UDC 519.226.2, 519.234.2

A.Iskakova, G.Zhaxybayeva

L.N. Gumilyov Eurasian National University, Astana, Kazakhstan
(E-mail: ayman.astana@gmail.com)

Maximum likelihood estimates of some probability model of discrete distributions

In this work the new multivariate discrete probability model of distribution of random sums with unobserved components is proposed.The maximum likelihood estimates for this model are determined in the case that all the elements of the sample implementation, namely the observed sums of unobserved components have only singular partition. In the case, that some element of the sample implementation has more then one partition, it is not possible to establish the maximum likelihood estimates.

Keywords: probability, multivariate distributions, maximum likelihood estimates.

1 Introduction

Models of probability distributions are a powerful and effective tool for studying diverse objects, systems and processes in various areas of human activity. In recent years, a significant number of probabilistic models have been developed.

Nevertheless, many unresolved problems remain, when it is possible to observe only the sums of components that can not be detected as a result of observations. For example, due to of the increasing use of digital media, there are failures of noise immunity, explained by randomly overlapping in one frequency band [1-3]. The use of probabilistic and static methods, namely, probabilistic modeling, allows us to present an approach to reducing the failures of noise immunity in information of digital media.

Also, an exclusive relevant example of the use of such a model is the advertising industry, where it is necessary to link the distribution of consumer interests with relevant advertising in various sources. Similar problems are very common in meteorology and in other areas. The probabilistic models describing such situations was considered in [4-6], where unbiased estimates were presented using the Rao-Blackwell-Kolmogorov method. Unlike the works [4-7], for these distributions in this paper we consider the maximum likelihood estimates and the conditions for the existence of their analytic derivation.

As is known, the maximum likelihood method is one of the most effective methods in terms of ensuring a minimum variance of the estimated parameters of probability distributions [5; 229]. The method is rigorous in the mathematical (probability-theoretic) plan. And its application is especially justified when there are both uncorrelated and correlated measurements in the processed information [8].

The strong consistency, the asymptotic unbiasedness, the asymptotic normality, asymptotic and the efficiency of maximum likelihood estimates provides their advantages in applied problems. Therefore the maximum likelihood estimates determined by the needs of practice, especially when using a large sample size.

The efficiency of the second order distinguishes this method of estimation among other asymptotically effective ones [9]. Invariance of maximum likelihood estimation ensures successful application of this method when estimating functions of distribution's parameters [10].

2 The model of multivariate discrete probability distribution

Consider a following probabilistic model by the example of the urn scheme with balls. Suppose that the urn contains balls, and each ball in the urn is marked by some value of number from the set of the random numbers $L_1, L_2, \ldots, L_d$. Let's the elements of the vector $\mathbf{p} = (p_1, p_2, \ldots, p_d)$ are the probabilities of retrieval from the urn of a ball marked by corresponding numbers $L_1, L_2, \ldots, L_d$ and

$$\sum_{\alpha=1}^{d} p_\alpha = 1.$$

By successive extraction of n balls from the urn with return we have the following situation. After the successive removal of n balls from the urn with the return, it is not known exactly what the balls were taken out of the urn. Only value u is known, which is the sum of the values of the numbers on the extracted balls from the urn. To study this situation, it is necessary to construct a probability distribution u. It is obvious that u is the realization of some random variable U.

Let‘s say, that V_u represents the number of possible combinations $r_{1_{v_u}}L_1, r_{2_{v_u}}L_2, \ldots, r_{d_{v_u}}L_d$, which in the sum formed the number u, where $r_{1_{v_u}}, r_{2_{v_u}}, \ldots, r_{d_{v_u}}$ determine the possible number of removed balls that are marked with the corresponding the random numbers L_1, L_2, $\ldots$, L_d. In other words, V_u is the number of partitions of the u into parts L_1, L_2, $\ldots$, L_d [11; 1].

The following assertion follows from the results of [4-6]. The probability that the random variable U will take the value u, is

$$P(U = u) = \sum_{v_U=1}^{V_u} n! \prod_{\alpha=1}^{d} \frac{p_\alpha^{r_{\alpha_{v_u}}}}{r_{\alpha_{v_u}}!}. \tag{1}$$

3 Formulation of the problem

Obviously, in practice, the elements of the vector $\mathbf{p}$ are not known. Consequently, formula (1) does not find actual application. In this connection, it becomes necessary to determine the probability estimate (1).

It is also L_1, L_2, $\ldots$,, L_d, which give the sum u, are not known.

Let's $\mathbf{x} = (x_1, \ldots, x_k)$ can be interpreted as a realization of a sample $\mathbf{X} = \{X_1, ..., X_k\}$ with size k, whose elements have distribution (1). We denote vector $\mathbf{r}_{v_\beta} = (r_{1_{v_\beta}}, \ldots, r_{d_{v_\beta}})$, which defines v_β - th solution of equation

$$\begin{cases} \sum\limits_{\alpha=1}^{d} L_\alpha r_{\alpha_{v_\beta}} = \mathbf{x}_\beta; \\ \sum\limits_{\alpha=1}^{d} r_{\alpha_{v_\beta}} = n, \end{cases} \tag{2}$$

where $v_\beta = 1, \ldots, V_\beta, V_\beta$ is the number of partitions of the x_β on the L_1, L_2, $\ldots$,, L_d. Using the L_1, L_2, $\ldots$,, L_d, and the realization of simple $\mathbf{x}$ in the system of equations (2) we define for each $\beta = 1, \ldots, k$ the number of partitions V_β of the sum $\mathbf{x}_\beta$ on L_1, L_2, $\ldots$,, L_d, and vectors $\mathbf{r}_{1_\beta}$, $\ldots$, $\mathbf{r}_{V_\beta}$.

Suppose that for each $j = 1, \ldots, \mu$, where

$$\mu = \prod_{\beta=1}^{k} V_\beta,$$

there is a vector $\mathbf{z}_j = (z_{1_j}, \ldots, z_{d_j})$, defined as

$$\mathbf{z}_j = \sum_{\beta=1}^{k} \mathbf{r}_{v_\beta}, \tag{3}$$

and the indices on the right and left side are linked one-to-one correspondence, which is not unique. For example, this line can be described by the following form

$$j = v_1 + (v_2 - 1)V_1 + (v_3 - 1)V_1V_2 + \ldots + (v_k - 1)\prod_{\beta=1}^{k} V_\beta. \tag{4}$$

Also, it can be represented as

$$j = v_k + (v_{k-1} - 1)V_k + (v_{k-2} - 1)V_kV_{k-1} + \ldots + (v_1)\prod_{\beta=2}^{k} V_\beta.$$

That is, if used (4), then (3) can be represented as the following systems of equations

$$\mathbf{z}_1 = \mathbf{r}_{1_1} + \mathbf{r}_{1_2} + \mathbf{r}_{1_3} + \ldots + \mathbf{r}_{1_k};$$

$$\mathbf{z}_2 = \mathbf{r}_{2_1} + \mathbf{r}_{1_2} + \mathbf{r}_{1_3} + \ldots + \mathbf{r}_{1_k};$$

$$\ldots$$

$$\mathbf{z}_{V_1} = \mathbf{r}_{1_{V_1}} + \mathbf{r}_{1_2} + \mathbf{r}_{1_3} + \ldots + \mathbf{r}_{1_k};$$

$$\mathbf{z}_{V_1+1} = \mathbf{r}_{1_1} + \mathbf{r}_{2_2} + \mathbf{r}_{1_3} + \ldots + \mathbf{r}_{1_k};$$

$$\mathbf{z}_{V_1+2} = \mathbf{r}_{2_1} + \mathbf{r}_{2_2} + \mathbf{r}_{1_3} + \ldots + \mathbf{r}_{1_k};$$

$$\ldots$$

$$\mathbf{z}_{\mu} = \mathbf{r}_{V_1} + \mathbf{r}_{V_2} + \mathbf{r}_{V_3} + \ldots + \mathbf{r}_{V_k}.$$

The following Lemma allows to determine which vectors $\mathbf{r}_{v_1}, \mathbf{r}_{v_2}, \ldots, \mathbf{r}_{v_k}$ form the vector $\mathbf{z}_j$. Suppose that for some real value a the value $\langle a \rangle$ determines the integer part of a.

Lemma. If the indices in the right and left sides of the equation (3) are interconnected in form (4), then

$$v_k = \left\langle \frac{j-1}{\prod_{i=1}^{k-1} V_i} \right\rangle + 1;$$

$$v_{k-1} = \left\langle \frac{j - (v_k - 1)\prod_{i=1}^{k-1} V_i - 1}{\prod_{i=1}^{k-2} V_i} \right\rangle + 1;$$

$$v_{k-2} = \left\langle \frac{j - (v_{k-1} - 1)\prod_{i=1}^{k-2} V_i - (v_k - 1)\prod_{i=1}^{k-1} V_i - 1}{\prod_{i=1}^{k-3} V_i} \right\rangle + 1;$$

$$\ldots$$

$$v_2 = \left\langle \frac{j - (v_3 - 1)V_1V_2 - (v_4 - 1)V_1V_2V_3 - \ldots - (v_k - 1)\prod_{i=1}^{k-1} V_i - 1}{\prod_{i=1}^{k-3} V_i} \right\rangle + 1;$$

$$v_1 = j - (v_2 - 1)V_1 - (v_3 - 1)V_1V_2 - (v_4 - 1)V_1V_2V_3 - \ldots - (v_k - 1)\prod_{i=1}^{k-1} V_i.$$

Proof. From (4) it follows that

$$v_k = \frac{j-c}{\prod_{i=1}^{k-1} V_i} + 1, \tag{5}$$

where

$$c = v_1 + (v_2 - 1)V_1 + (v_3 - 1)V_1V_2 + \ldots + (v_{k-1} - 1)\prod_{i=1}^{k-2} V_i.$$

It is obvious that the latter can be represented as follows

$$c = v_{k-1}\prod_{i=1}^{k-2} V_i - \left[(V_{k-2} - v_{k-2})\prod_{i=1}^{k-3} V_i (V_{k-3} - v_{k-3})\prod_{i=1}^{k-4} V_i + \ldots + V_1 - v_1 \right].$$

Since

$$(V_{k-2} - v_{k-2})\prod_{i=1}^{k-3} V_i (V_{k-3} - v_{k-3})\prod_{i=1}^{k-4} V_i + \ldots + V_1 - v_1 \geq 0,$$

then

$$c \leq (v_{k-1} - 1) \prod_{i=1}^{k-2} V_i \leq \prod_{i=1}^{k-1} V_i$$

or

$$\frac{c-1}{\prod\limits_{i=1}^{k-1} V_i} < 1. \tag{6}$$

By the fact that (5), then we obtain receive

$$v_k = \frac{j-1}{\prod\limits_{i=1}^{k-1} V_i} - \frac{c-1}{\prod\limits_{i=1}^{k-1} V_i} + 1.$$

Since we have (6) and v_k is non-negative integer, then

$$v_k = \left\langle \frac{j-1}{\prod\limits_{i=1}^{k-1} V_i} \right\rangle + 1.$$

The same way as v_k has been determined in the last formula, v_{k-1}, v_{k-2}, $\ldots$, v_1 are determined. Lemma is proved.

4 Construction of maximum likelihood estimates for the distribution parameters of the model investigated

Find maximum likelihood estimates for the parameters $p_1, \ldots, p_d$ of distribution (1). The likelihood function of distribution (1) has form

$$L(\mathbf{x};\mathbf{p}) = \prod_{\beta=1}^{k} P(U = x_i) = \prod_{\beta=1}^{k} \sum_{v_\beta=1}^{V_\beta} n! \prod_{\alpha=1}^{d} \frac{p_\alpha^{r_{\alpha v_\beta}}}{r_{\alpha_{v_\beta}}!},$$

which can present in form

$$L(\mathbf{x};\mathbf{p}) = \prod_{\beta=1}^{k} P(U = x_i) = \prod_{\beta=1}^{k} \frac{n! \sum\limits_{v_\beta=1}^{V_\beta} \prod\limits_{\alpha=1}^{d} \frac{p_\alpha^{r_{\alpha v_\beta}}}{r_{\alpha_{v_\beta}}!}}{\left(\sum\limits_{\alpha=1}^{d} p_\alpha\right)^{nV_\beta}}.$$

Accordingly, the log-likelihood for the parameters $p_1, \ldots, p_d$ of distribution (1) is

$$\ln L(\mathbf{x};\mathbf{p}) = k \ln n! + \sum_{\beta=1}^{k} \ln \sum_{v_\beta=1}^{V_\beta} \prod_{\alpha=1}^{d} \frac{p_\alpha^{r_{\alpha v_\beta}}}{r_{\alpha_{v_\beta}}!} - n\eta \ln \left(\sum_{\alpha=1}^{d} p_\alpha\right)^{nV_\beta},$$

where $\eta = \sum\limits_{\beta=1}^{k} V_\beta$.

It follows that for any $\alpha^* = 1, \ldots, d$ we have

$$\frac{\partial \ln L(\mathbf{x};\mathbf{p})}{\partial p_{\alpha^*}} = \sum_{\beta=1}^{k} \frac{\sum\limits_{v_\beta=1}^{V_\beta} \frac{p_{\alpha^*}^{r_{\alpha^* v_\beta}-1}}{\left(r_{\alpha^* v_\beta}-1\right)!} \prod\limits_{\substack{\alpha=1 \\ \alpha \neq \alpha^*}}^{d} \frac{p_\alpha^{r_{\alpha v_\beta}}}{r_{\alpha_{v_\beta}}!}}{\sum\limits_{v_\beta=1}^{V_\beta} \prod\limits_{\alpha=1}^{d} \frac{p_\alpha^{r_{\alpha v_\beta}}}{r_{\alpha_{v_\beta}}!}} - n\eta;$$

or

$$\frac{\partial \ln L(\mathbf{x};\mathbf{p})}{\partial p_{\alpha^*}} = \sum_{\beta=1}^{k}\sum_{v_\beta=1}^{V_\beta} \frac{r_{\alpha^* v_\beta}}{p_{\alpha^*}\Lambda_{v_\beta}} - n\eta, \tag{7}$$

where for $\beta = 1, \ldots, k$, $v_\beta = 1, \ldots, V_\beta$

$$\Lambda_{v_\beta} = 1 + \sum_{\substack{w_\beta = 1 \\ w_\beta \neq v_\beta}}^{V_\beta} \prod_{\alpha=1}^{d} \frac{r_{\alpha_{v_\beta}}!}{r_{\alpha_{w_\beta}}!} p_\alpha^{r_{\alpha_{w_\beta}} - r_{\alpha_{v_\beta}}}. \tag{8}$$

As it is known, the maximum likelihood estimations $\hat{\mathbf{p}} = (\hat{p}_1, \ldots, \hat{p}_d)$ for vector of parameters $\mathbf{p} = (p_1, \ldots, p_d)$ satisfy the following condition for any $\alpha = 1, \ldots, d$

$$\left.\frac{\partial \ln L(\mathbf{x};\mathbf{p})}{\partial p_\alpha}\right|_{\mathbf{p}=\hat{\mathbf{p}}} = 0. \tag{9}$$

It follows that $\ln L(\mathbf{x}, \mathbf{p})$ reaches a local maximum at the point $\hat{\mathbf{p}} = (\hat{p}_1, \ldots, \hat{p}_d)$, for any $\alpha_1, \alpha_2 = 1, \ldots, s$, $s = 2, \ldots, d$ carried the following conditions

$$\begin{cases} \det \left\| \frac{\partial^2 \ln L(\mathbf{x},\mathbf{p})}{\partial p_{\alpha_1} \partial p_{\alpha_2}} \right\|_{s\times s} > 0, & \text{if } s \text{ is even;} \\ \det \left\| \frac{\partial^2 \ln L(\mathbf{x},\mathbf{p})}{\partial p_{\alpha_1} \partial p_{\alpha_2}} \right\|_{s\times s} < 0, & \text{otherwise.} \end{cases} \tag{10}$$

From (7) and(9) it follows that for $\alpha = 1, \ldots, d$

$$\hat{p}_\alpha = \frac{\sum\limits_{\beta=1}^{k}\sum\limits_{v_\beta=1}^{V_\beta} \frac{r_{\alpha_{v_\beta}}}{\Lambda_{v_\beta}}}{n\eta}. \tag{11}$$

Since

$$\sum_{\alpha=1}^{d} \hat{p}_\alpha = 1,$$

then in conformity with (11)

$$\sum_{\alpha=1}^{d} \hat{p}_\alpha = \frac{\sum\limits_{\alpha=1}^{d}\sum\limits_{\beta=1}^{k}\sum\limits_{v_\beta=1}^{V_\beta} \frac{r_{\alpha_{v_\beta}}}{\Lambda_{v_\beta}}}{n\eta} = 1.$$

Hence, we have

$$\sum_{\alpha=1}^{d}\sum_{\beta=1}^{k}\sum_{v_\beta=1}^{V_\beta} \frac{r_{\alpha_{v_\beta}}}{\Lambda_{v_\beta}} = n\eta. \tag{12}$$

From (8) it is evident that for any $\beta = 1, \ldots, k$, $v_\beta = 1, \ldots, V_\beta$ $\Lambda_{v_\beta} \geq 1$. And $\Lambda_{v_\beta} = 1$, if $V_\beta = 1$, otherwise $\Lambda_{v_\beta} > 1$. Thus, we have

$$\sum_{\alpha=1}^{d}\sum_{\beta=1}^{k}\sum_{v_\beta=1}^{V_\beta} \frac{r_{\alpha_{v_\beta}}}{\Lambda_{v_\beta}} \leq \sum_{\alpha=1}^{d}\sum_{\beta=1}^{k}\sum_{v_\beta=1}^{V_\beta} r_{\alpha_{v_\beta}} = n\eta.$$

That is

$$\sum_{\alpha=1}^{d}\sum_{\beta=1}^{k}\sum_{v_\beta=1}^{V_\beta} \frac{r_{\alpha_{v_\beta}}}{\Lambda_{v_\beta}} \neq n\eta,$$

if for some $\beta = 1, \ldots, k$ $\Lambda_{v_\beta} > 1$. So (12) is satisfied if for all $\beta = 1, \ldots, k$ $V_\beta = 1$. Consequently, the construction of maximum likelihood estimates for the distribution parameters $\mathbf{p} = (p_1, \ldots, p_d)$ of this model (1) is possible only when all elements of realization of sample have no more than one partition on the submitted $L_1, L_2, \ldots, L_d$.

In other words, if for all $\beta = 1, \ldots, k$ $V_\beta = 1$, then from (8) it implies $\Lambda_{v_\beta} = 1$, and by (11) for $\alpha = 1, \ldots, d$ we have

$$\hat{p}_\alpha = \frac{\sum\limits_{\beta=1}^{k} \sum\limits_{v_\beta=1}^{V_\beta} r_{\alpha_{v_\beta}}}{n\eta} = \frac{\sum\limits_{\beta=1}^{k} r_{\alpha_1}}{nk}.$$

That is

$$\hat{p}_\alpha = \frac{\sum\limits_{\beta=1}^{k} r_{\alpha_1}}{nk}. \qquad (13)$$

From (7) and (13) it's following that for any $\alpha, \alpha_1, \alpha_2 = 1, \ldots, d$ if $\alpha_1 \neq \alpha_2$

$$\left.\frac{\partial^2 \ln L(\mathbf{x}, \mathbf{p})}{\partial p_{\alpha_1} \partial p_{\alpha_1}}\right|_{\mathbf{p}=\hat{\mathbf{p}}} = 0$$

and

$$\left.\frac{\partial^2 \ln L(\mathbf{x}, \mathbf{p})}{\partial p_\alpha^2}\right|_{\mathbf{p}=\hat{\mathbf{p}}} = -\sum_{\beta=1}^{k} \frac{r_{\alpha_1}}{\hat{p}_\alpha^2} < 0.$$

So for any $s = 2, \ldots, d$ we have the following

$$\det \left\| \frac{\partial^2 \ln L(\mathbf{x}, \mathbf{p})}{\partial p_{\alpha_1} \partial p_{\alpha_2}} \right\|_{s \times s} = \prod_{\alpha=1}^{s} \left.\frac{\partial^2 \ln L(\mathbf{x}, \mathbf{p})}{\partial p_\alpha^2}\right|_{\mathbf{p}=\hat{\mathbf{p}}}$$

or

$$\det \left\| \frac{\partial^2 \ln L(\mathbf{x}, \mathbf{p})}{\partial p_{\alpha_1} \partial p_{\alpha_2}} \right\|_{s \times s} = \prod_{\alpha=1}^{s} \left(\frac{r_{\alpha_1}}{\hat{p}_\alpha^2} \right).$$

Consequently, if for all $\beta = 1, \ldots, k$ $V_\beta = 1$, then the elements of the vector

$$\hat{\mathbf{p}} = (\hat{p}_1, \ldots, \hat{p}_d),$$

defined in (12) satisfy (8)–(9) and they are the maximum likelihood estimates for the parameters $\mathbf{p} = (p_1, \ldots, p_d)$ of disrtibution (1). Thus, the following Theorem holds.

Theorem. If all elements of realization of sample $\mathbf{x} = (x_1, \ldots, x_k)$ of distibution (1) have no more than one partition on the on the submitted $L_1, L_2, \ldots, L_d$, then there are maximum likelihood estimates for the parameters of the distribution (1), which is defined in (12).

Consequence. If some element from realization of sample $\mathbf{x} = (x_1, \ldots, x_k)$ of distibution (1) have more than one partition on the on the submitted $L_1, L_2, \ldots, L_d$, then there not are maximum likelihood estimates for the parameters of the distribution (1).

Thus, it can not always possible to construct maximum likelihood estimators for the parameters of the distribution (1).

5 Conclusion

The analysis conducted in this paper studies allows us to formulate the following conclusion: Found that for this model maximum likelihood estimates exist if all the elements of observations have no more than one part by partition.

As is known, in practice, often some element of the implementation of the sample has more than one partition. That is, the method for determining the likelihood estimation is not actually applicable to this model. Of course, it is possible to use modified likelihood estimates by means of the apparatus of numerical methods, for example, to solve the system of maximum likelihood equations by the iterative method [12] or directly maximize the likelihood function of the type [13].

Obviously, the application of numerical methods generates numerous problems. Namely, the convergence of iterative methods requires justification [14; 202], the likelihood function can has a several local maxima [15], the choice of the moment of termination of calculations in connection with the achievement of the required accuracy requires justification [16], also the accuracy of the computation depends to the sample size [17].

Thus, if any element of the sampling of the given distribution model has more than one decomposition, then when finding the maximum likelihood estimates, we have a number of computational problems that call into question the practicality of using maximum likelihood estimates.

There is no need to absolutize the maximum likelihood estimates. In addition to these, there are other types of estimates that have good asymptotic properties. An example is the most suitable unbiased estimates presented in [4-6].

References

1 Mu T. Optimal design and performance metric of broadband full-Stokes polarimeters with immunity to Poisson and Gaussian noise / T.Mu, Z.Chen, C.Zhang, R.Liang // Optics express. — 2016. — 24(26). — P. 29691–29704.

2 An J. Automatic on-chip backup clock changer for protecting abnormal MCU operations in unsafe clock frequency / J.An, M.G. Seok, D.Park // IEICE Electronics Express. — 2016. — Vol. 13. — No. 24. — P. 20160808–20160808.

3 Savchenko V.V. Enhancement of the noise immunity of a voice-activated robotics control system based on phonetic word decoding method / V.V. Savchenko // Journal of Communications Technology and Electronics. — 2016. — Vol. 61. — No. 12. — P. 1374–1378.

4 Ayman I. Construction of the most suitable unbiased estimate distortions of radiation processes from remote sensing data / I.Ayman // Journal of Physics: Conference Series. — 2014. — Vol. 490. — No. 1. — P. 012113. IOP Publ.

5 Ayman I. Statistical Research for Probabilistic Model of Distortions of Remote Sensing / I.Ayman // Journal of Physics: Conference Series. — 2016, August. — Vol. 738. — No. 1. — P. 012004. IOP Publ.

6 Искакова А.С. Определение наиболее подходящей несмещенной оценки вероятности оправдываемости прогноза в метеорологии / А.С. Искакова // Сибирский журнал индустриальной математики. — 2002. — № 1. — С. 79–84.

7 Ibragimov I.A. Statistical estimation: asymptotic theory / I.A. Ibragimov, R.Z.Has' Minskii. — Berlin: Springer Science & Business Media, 2013. — Vol. 16.

8 Bai J. Maximum likelihood estimation and inference for approximate factor models of high dimension / J.Bai, K.Li // Review of Economics and Statistics. — 2016. — No. 98(2). — P. 298–309.

9 Jiao J. Maximum likelihood estimation of functionals of discrete distributions / J.Jiao, K.Venkat, Y.Han, T.Weissman // IEEE Transactions on Information Theory, 2017. — Vol. 63. — No. 10. — P. 6774–6798.

10 Schmid H. Probabilistic part-of-speech tagging using decision trees / H.Schmid // In New methods in language processing. — 2013, November. — P. 154.

11 Andrews G.E. The theory of partitions / G.E.Andrews. — Cambridge: Cambridge University Press, 1998. — No. 2.

12 Ziskind I. Maximum likelihood localization of multiple sources by alternating projection / I.Ziskind, M.Wax // IEEE Transactions on Acoustics, Speech, and Signal Processing. — 1988. — Vol. 36. — No. 10. — P. 1553–1560.

13 Neal R.M. A view of the EM algorithm that justifies incremental, sparse, and other variants / R.M. Neal, G.E. Hinton // Learning in graphical models. — Luxembourg: Springer Netherlands, 1998.—P. 355–368.

14 Aster R.C. Parameter estimation and inverse problems / R.C. Aster, B. Borchers, C.H. Thurber. — Cambridge: Academic Press, 2011. — Vol. 90.

15 Fan J. Strong oracle optimality of folded concave penalized estimation / J.Fan, L.Xue, H.Zou // Annals of statistics. — 2014. — Vol. 42. — No. 3. — P. 819.

16 Inglis M., Gilmore C. Indexing the approximate number system / M.Inglis, C.Gilmore // Acta Psychologica. — 2014. — Vol. 145. — P. 147–155.

17 Mengersen K.L. Bayesian computation via empirical likelihood / K.L. Mengersen, P.Pudlo, C.P. Robert // Proceedings of the National Academy of Sciences. — 2013. — 110(4). — P. 1321–1326.

А.Искакова, Г.Жаксыбаева

Дискретті үлестірімдердің бір ықтималдық моделінің шындыққа ұқсас максималды бағалары

Мақалада бақыланатын компоненттерімен берілген кездейсоқ қосындылар үлестірімінің жаңа көпөлшемді дискретті ықтималдық моделі ұсынылған. Осы модель үшін шындыққа ұқсас бағалары анықталған, оның ішінде егер таңдаманы жүзеге асырудағы барлық элементтері, дәл осы бақыланбайтын компоненттердің бақыланатын қосындылары тек жалғыз бөліктеуге ие болса. Сонымен қатар таңдаманы жүзеге асырудағы элемент жалғыз бөліктеуге ие болмаса, максималды шындыққа ұқсас бағаларды алу мүмкін емес.

Кілт сөздер: ықтималдық, көпөлшемді үлестірім, максималды шындыққа ұқсас бағалар.

А.Искакова, Г.Жаксыбаева

Оценки максимального правдоподобия одной вероятностной модели дискретных распределений

В статье представлена новая многомерная дискретная вероятностная модель распределения случайных сумм с ненаблюдаемыми компонентами. Определены оценки максимального правдоподобия для этой модели в том случае, если все элементы реализации выборки, а именно наблюдаемые суммы ненаблюдаемых компонентов, имеют только единственные разбиения. В случае если какой-нибудь элемент реализации выборки имеет не единственное разбиение, то оценки максимального правдоподобия невозможно установить.

Ключевые слова: вероятность, многомерные распределения, оценки максимального правдоподобия.

References

1 Mu, T., Chen, Z., Zhang, C., & Liang, R. (2016). Optimal design and performance metric of broadband full-Stokes polarimeters with immunity to Poisson and Gaussian noise. *Optics express, 24(26)*, 29691–29704.

2 An, J., Seok, M. G., & Park, D. (2016). Automatic on-chip backup clock changer for protecting abnormal MCU operations in unsafe clock frequency. *IEICE Electronics Express, Vol. 13*, 24, 20160808-20160808.

3 Savchenko, V.V. (2016). Enhancement of the noise immunity of a voice-activated robotics control system based on phonetic word decoding method. *Journal of Communications Technology and Electronics, Vol. 61*, 12, 1374–1378.

4 Ayman, I. (2014). Construction of the most suitable unbiased estimate distortions of radiation processes from remote sensing data. *Journal of Physics: Conference Series, Vol. 490*, 1, 012113. IOP Publishing.

5 Ayman, I. (2016). Statistical Research for Probabilistic Model of Distortions of Remote Sensing. *Journal of Physics: Conference Series, Vol. 738*, 1, 012004. (2016, August). IOP Publishing.

6 Iskakova, A.S. (2002). Opredelenie naibolee podkhodiashchei nesmeshchennoi otsenki veroiatnosti opravdyvaemosti prohnoza v meteorolohii [Determination of the most suitable unbiased estimate for a weather forecast being correc]. *Sibirskii zhurnal industrialnoi matematiki – Siberian Journal of Industrial Mathematics, 1*, 79–84 [in Russian].

7 Ibragimov, I.A., & Has' Minskii, R.Z. (2013). Statistical estimation: asymptotic theory (Vol. 16). Berlin: Springer Science & Business Media.

8 Bai, J., & Li, K. (2016). Maximum likelihood estimation and inference for approximate factor models of high dimension. *Review of Economics and Statistics, 98(2)*, 298–309.

9 Jiao, J., Venkat, K., Han, Y., & Weissman, T. (2017). Maximum likelihood estimation of functionals of discrete distributions. *IEEE Transactions on Information Theory, Vol. 63*, 10, 6774–6798.

10 Schmid, H. (2013, November). Probabilistic part-of-speech tagging using decision trees. *New methods in language processing*, 154.

11 Andrews, G.E. (1998). The theory of partitions, No. 2. Cambridge: Cambridge university press.

12 Ziskind, I., & Wax, M. (1988). Maximum likelihood localization of multiple sources by alternating projection. *IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 36*, 10, 1553–1560.

13 Neal, R.M., & Hinton, G.E. (1998). A view of the EM algorithm that justifies incremental, sparse, and other variants. *Learning in graphical models*, 355–368. Luxembourg: Springer Netherlands.

14 Aster, R.C., Borchers, B., & Thurber, C.H. (2011). Parameter estimation and inverse problems, Vol. 90. Cambridge: Academic Press.

15 Fan, J., Xue, L., & Zou, H. (2014). Strong oracle optimality of folded concave penalized estimation. *Annals of statistics, Vol. 42*, 3, 819.

16 Inglis, M., & Gilmore, C. (2014). Indexing the approximate number system. *Acta Psychologica, Vol. 145*, 147–155.

17 Mengersen, K.L., Pudlo, P., & Robert, C.P. (2013). Bayesian computation via empirical likelihood. *Proceedings of the National Academy of Sciences, 110(4)*, 1321–1326.