DEMONSTRATIO MATHEMATICA
Vol. IV No 3 1972

Włodzimierz Krysicki

ESTIMATION OF THE PARAMETERS OF THE MIXTURE OF AN ARBITRARY NUMBER OF EXPONENTIAL DISTRIBUTIONS

Introduction. K.Pearson in [1] first indicated the importance of estimating the parameters of a mixture of distributions. He gave a method of estimating five parameters of the mixture of two normal distributions $p\,N(m_1, \sigma_1) + (1-p)\,N(m_2, \sigma_2)$. Other papers on this topic have been subsequently published by K.Pearson [2], Muench [3], [4], Gumbel [5], Mendenhall and Hader [6], Rider [7,8], Blischke [9], Krysicki [10, 11], Kącki [12], Wasilewski [13,14], Cohen [15], Kącki and Krysicki [16], Behbodian [17], Falls [18]. The mentioned papers deal with mixtures of distributions of continuous as well as discrete type, but always of two components, depending on one or two parameters. The first paper concernig the estimation of the mixture of more than two Bernoulli distributions was written by Blischke [19]. He applied there factorial moments. Subsequently Hasselblad gave a method of the estimation of parameters of the mixture of $k \geqslant 3$ normal distributions $N(m_i, \sigma_i)$ derived from the maximum likelihood method. The most recent papers in this field are papers of Kabir [21] and Gridgeman [22]. The latter deals with the estimation of parameters of the mixture of normal distributions $N(0, \sigma_i)$. In the interesting paper [21] the author gives a method of the estimation of parameters of the mixture of $k \geqslant 3$ distributions belonging to the class of so-called one-parameter exponential distributions introduced by Koopman and Pitman in 1936. To use this method it is necessary to restrict the range of the random variable to a finite interval. Kabir used this method to the estimation of parameters of the mixtures of two exponential distributions and it allows to estimate jointly three parameters. In the present work we shall give a method of estimating parameters of the mixture of an arbitrary finite number of exponential distributions without the necessity of restricting the range to a finite interval.

1. FORMULATION OF THE PROBLEM

Consider the mixture of k $(k = 2,3,\ldots)$ exponential distributions whose density is expressed by the formula

$$f(x) = \begin{cases} \sum_{1}^{k} p_i \, m_i^{-1} \exp\left(-\frac{x}{m_i}\right) & \text{if } x \in (0, +\infty) \\ 0 & \text{otherwise.} \end{cases} \tag{1.1}$$

We may assume the following conditions

$$p_i > 0, \quad \sum_{1}^{k} p_i = 1 \tag{1.2}$$

$$0 < m_1 < m_2 \ldots < m_k \tag{1.3}$$

without the loss of generality.

On the basis of a sample taken from the general population, in which the examined trait X is subject to the distribution (1.1), assuming that the results of the sample are

$$x_1, x_2, \ldots, x_n, \tag{1.4}$$

we are to estimate the parameters

$$m_1, m_2, \ldots, m_k$$

$$p_1, p_2, \ldots p_{k-1},$$

that is, the $2k - 1$ parameters satisfying conditions (1.2) and (1.3).

2. THE PROBLEM OF UNIQUENESS

Already in 1948 Robbins 23 paid attention to the problem of identifiability of mixtures. This problem consists in showing that two different sets of values of the parameters of a mixture cannot determine the same distribution of the mixture. Namely, it is known that in general the class of continuous distributions, e.g. normal or two-parameters gamma, as well as the class of discrete distributions, e.g. of Bernoulli type, are not identifiable [24, 25].On the other hand, Teicher showed [26] that the class of all mixtures of a finite number of two-parameters gamma distributions is identifiable. Since the class of exponential distributions is contained in that class, it follows that the class of mixtures of a finite number of exponential distributions is identifiable. This means that the estimation of parameters of the mixture (1.1) on the basis of the sample (1.4) is uniquely feasible and hence thee posed problem has a unique solution.

3. THE SOLUTION OF THE PROBLEM

The ordinary moments of order r, $r = 1,2,\ldots,$ of an exponentially distributed random variable X with the parameter $m = E(X)$ are expressed by the formula:

$$\alpha_r = r!\, m^r, \quad r = 1,2,\ldots$$

Consequently, the ordinary moments of the mixture (1.1) are as follows

$$\bar{\alpha}_r = r! \sum_{1}^{k} p_i\, m_i^r \,. \tag{3.1}$$

Let us form 2k equations

$$
\begin{aligned}
&p_1m_1 + p_2m_2 + \dots + p_km_k = \bar{\alpha}_1 \\
&p_1m_1^2 + p_2m_m^2 + \dots + p_km_k^2 = \bar{\alpha}_2/2! \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&p_1m_1^{2k} + p_2m_2^{2k} + \dots + p_km_k^{2k} = \bar{\alpha}_{2k}/(2k!).
\end{aligned}
\tag{3.2}
$$

To simplify the notation put

$$\frac{\bar{\alpha}_r}{r!} = M_r, \quad r = 1,2,\dots, \quad 2k. \tag{3.3}$$

Let m_i, $i = 1,2,\dots k$ satisfy conditions (1.3). Assume that they are also the roots of the following equation with real coefficients

$$m^k + a_1m^{k-1} + \dots + a_{k-1}m + a_k = 0. \tag{3.4}$$

Multiply the first equation in (3.2) by a_k, the second by a_{k-1}, and so on, the k-th equation by a_1, and sum them side by side. Using the notation introduced in (3.3) we can write the result as follows

$$M_1a_k + M_2a_{k-1} + \dots + M_ka_1 = p_1(a_km_1 + a_{k-1}m_1^2 + \dots + a_1m_1^k) + \dots$$

$$\dots + p_k(a_km_k + a_{k-1}m_k^2 + \dots + a_1m_k^k).$$

Simplifying the right-hand side and using the fact that the m_i are the roots of equation (3.4), we obtain

$$M_1a_k + M_2a_{k-1} + \dots + M_ka_1 + M_{k+1} = 0.$$

Similarly we obtain the next $k-1$ equations. Hence we have a system of k equations

$$
\begin{aligned}
&M_1a_k + M_2a_{k-1} + \dots + M_ka_1 + M_{k+1} = 0\\
&M_2a_k + M_3a_{k+1} + \dots + M_ka_1 + M_{k+2} = 0\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \qquad (3.5)\\
&M_ka_k + M_{k+1}a_{k-1} + \dots + M_{2k-1}a_1 + M_{2k} = 0.
\end{aligned}
$$

Let M denote matrix of this system. M can be written as follows

$$
\begin{bmatrix} m_1 & m_2 & \dots & m_k \\ m_1^2 & m_2^2 & \dots & m_k^2 \\ \dots\dots \\ m_1^k & m_2^k & \dots & m_k^k \end{bmatrix}
\begin{bmatrix} p_1 & 0 & \dots & 0 \\ 0 & p_2 & \dots & 0 \\ \dots\dots \\ 0 & 0 & \dots & p_k \end{bmatrix}
\begin{bmatrix} 1 & m_1 & \dots & m_1^{k-1} \\ 1 & m_2 & \dots & m_2^{k-1} \\ \dots\dots \\ 1 & m_k & \dots & m_k^{k-1} \end{bmatrix}. \qquad (3.6)
$$

Since the determinants corresponding to the first and the third matrix are Vandermonde's determinants and consequantly in view of (1.3) are different from zero, we infer that the matrix M is non-singular (to obtain this conclusion we have applied also conditions (1.2)). This implies that the system (3.5) has exactly one solution expressed by Cramer's formulas. Hence the coefficients $a_1, a_2, \dots, a_k$ of equation (3.4) can be expressed rationally by means of the ordinary moments $\bar{\alpha}_r$ of the mixture, and consequently they can be expressed rationally by the ordinary moments m_r. As estimators $\hat{\bar{\alpha}}_r$ for the ordinary moments $\bar{\alpha}_r$ $(r = 1,2,\dots,2k)$ of the mixture (1.1) we take the ordinary moments of the sample (1.4), and as estimators for M_r defined by formulas (3.3) we take

$$
\hat{M}_r = \frac{1}{r!n} \sum_{i=1}^{n} x_i^r, \quad r = 1,2,\dots,2k. \qquad (3.7)
$$

It is known that the ordinary moments of a sample of an arbitrary order (provided they exist) are un-biased and consistent estimators of the ordinary moments of the general population. Substituting the estimators $\hat{M}_i$ defined by (3.7) for Mi (i=1,2,...2k) in formula (3.5) and solving the obtained system we obtain k estimators for the coefficients of equation (3.4):

$$\hat{a}_1,\ \hat{a}_2,\ \dots,\ \hat{a}_k. \tag{3.8}$$

Since the estimators (3.8) can be expressed rationally by the moments of a sample, Słucki's theorem applies and we can infer that they are stochastically convergent to the "true values" of the coefficients $a_i (i=1,\dots,k)$ of equation (3.4). In other words, the estimators (3.8) are consistent estimators of the coefficients a_i.

It may happen that the determinant obtained from the determinant of equation (3.5) after replacing M_i by the estimators $\hat{M}_i$ from the sample (1.4) becomes equal to 0, but this is practically impossible for a large sample in view of the fact the sample (1.4) has been taken from the mixture (1.1).

As a next step we substitute the values (3.8) for a_i in equation (3.8) and obtain the following equation with number coefficients:

$$m^k + \hat{a}_1 m^{k-1} + \dots + \hat{a}_{k-1} m + \hat{a}_k = 0. \tag{3.9}$$

Solving this equation we obtain the roots (with the required exactness), which are estimators of the parameters (1.3):

$$\hat{m}_1,\ \hat{m}_2,\ \dots,\ \hat{m}_k\ . \tag{3.10}$$

We cannot claim that the roots of equation (3.9) are always positive and all different, but in view of the assumption we have made above the event opposite to

$$0 < \hat{m}_1 < \hat{m}_2 \dots < \hat{m}_k \tag{3.11}$$

becomes practically impossible when $n \to \infty$.

The estimators (3.7) and (3.10) can now be substituted into the first k-1 equations of system (3.2). In view of (1.2) the resulting system can be written as follows:

$$(\hat{m}_1 - \hat{m}_k)p_1 + (\hat{m}_2 - \hat{m}_k)p_2 + \dots + (\hat{m}_{k-1} - \hat{m}_k)p_{k-1} = \hat{M}_1 - \hat{m}_k$$

$$(\hat{m}_1^2 - \hat{m}_k^2)p_1 + (\hat{m}_2^2 - \hat{m}_k^2)p_2 + \dots + (\hat{m}_{k-1}^2 - \hat{m}_k^2)p_{k-1} = \hat{M}_2 - \hat{m}_k^2$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \qquad (3.12)$$

$$(\hat{m}_1^{k-1} - \hat{m}_k^{k-1})p_1 + (\hat{m}_2^{k-1} + \hat{m}_k^{k-1})p_2 + \dots + (\hat{m}_{k-1}^{k-1} - \hat{m}_k^{k-1})p_{k-1} = \hat{M}_{k-1} - \hat{m}_k^{k-1}.$$

By (3,11) the determinant of this system is a Vandermonde's determinant and consequently it is different from zero.

Solving (3.12) we obtain the estimators

$$\hat{p}_1, \hat{p}_2, \dots, \hat{p}_{k-1}, \qquad (3.13)$$

and from formula (1.2) we can calculate $\hat{p}_k$. Again, it may happen that some $\hat{p}_i$ are negative, equal to 0, or greater than 1, but since the sample (1.4) has been taken from the mixture (1.1), the event opposite to $\left[(0 < p_i < 1) \cap (\sum_{1}^{k-1} p_i < 1)\right]$ is for a large sample practically impossible.

It is worth noticing that in this method we obtain estimators first for the whole group of parameters m_i $(i=1,2,\dots k)$, and then for the group of parameters $p_1, p_2, \dots, p_{k-1}$.

MATHEMATICAL INSTITUTE, TECHNICAL UNIVERSITY, ŁÓDŹ

REFERENCES

[1] K. Pearson: Contributions to the mathematical theory of evolution. Philos. Trans. Roy. Soc. London A 185(1894) 71-110.

[2] K. P e a r s o n: On certain types of compound frequency distributions in which the components can by individually described by binomial series. Biometrika 11(1915) 139-144.

[3] H. M u e n c h: Probability distributions of protection test results. J.A.S.A. 31(1936) 677-689.

[4] H. M u e n c h: Discrete frequency distributions arising from mixtures of several singly probability values. J.A.S.A. 33(1938) 390-398.

[5] E. G u m b e l: La dissection d'une répartition. Ann. Univ. Lyon 2 (1939) 39-51.

[6] W. M e n d e n h a l l, R. H a d e r: Estimation of parameters of mixed exponentially distributed failure time distributions from consored life test data. Biometrika 45(1938) 504-520.

[7] P. R i d e r: The method of moments applied to a mixture of two exponential distributions. Ann. Mathem. Statist. 32(1961) 143-147.

[8] P. R i d e r: Estimating the parameters of mixed Poisson, binomial and Weibull distributions by the method of moments.Bull. of the Intern. Statist. Instit. 39(1962) 225-232.

[9] W. B l i s c h k e: Moment estimators for the parameters of a mixture of two binomial distributions. Ann. Mathem. Statist. 33(19) ~~444-454~~.

[10] W. K r y s i c k i: Application de la méthode des moments a l'estimation des parametres d'un mélange de deux distributions de Rayleigh. Revue de Stat. Appl. 4(1963) 25-34.

[11] W. K r y s i c k i: Zastosowanie metody momentów do estymacji parametrów mieszaniny dwóch rozkładów Laplace'a. Zeszyty Nauk. Politech. Łódzkiej, Włókiennictwo 14(1966) 3-14.

[12] E. K ą c k i: Pewne przypadki estymacji parametrów mieszaniny dwóch rozkładów Laplace'a. Zeszyty Nauk. Politech. Łódzkiej, Elektryka, 20(1965).

[13] M. W a s i l e w s k i: Uber gewisse Merkmale der Mischung von zwei Maxwellschen - Verteilungen. Zeszyty Nauk.Politech.Łódzkiej, Włókiennictwo, 16(1966) 5-21.

[14] M. W a s i l e w s k i: Sur certaines propriétés de la distribution gamma generalisée. Revue de Stat. Appl. 1(1967) 95-105.

[15] A. C o h e n: Estimation in mixtures of two normal distributions. Technometrics 1(1967) 15-28.

[16] E. K ą c k i; W. K r y s i c k i: Die Parameterschätzung einer Mischung von zwei Laplaceschen Verteilungen in allgemeinen Fall. Commentationes Mathematicae XI, (1967) 23-31.

[17] J. B e h b o o d i a n: On the modes of a mixture of two normal distributions. Technometrics 12,1, rok 131-140.

[18] L. F a l l s: Estimation of parameters in compound Weibull distributions. Technometrics 2(1970) 399-407.

[19] W. B l i s c h k e: Estimating the parameters of mixtures of binomial distributions. J.A.S.A. 59(1964) 510-528.

[20] V. H a s s e l b l a d: Estimations of parameters for a mixture of normal distributions. Technometrics 3(1966) 431-444.

[21] A. K a b i r: Estimation of parameters of a fimite mixture of distributions. Journ. Royal Statist. Soc. 3(1968) 472-482.

[22] N. G r i d g e m a n: A comparison of two methods of a analysis of mixtures of normal distributions. Technometrics 4(1970) 823-834.

[23] H. R o b b i n s: Mixtures of distributions.Ann.Mathem.Statist. 19(1948) 360-369.

[24] H. T e i c h e r: On the mixture of distributions. Ann. Mathem. Statist. 31(1960) 35-73.

[25] H. T e i c h e r: Identifiability of mixtures. Ann.Mathem. Statist. 32(1961) 244-248.

[26] H. T e i c h e r: Identifiability of finite mixtures. Ann. Mathem. Statist. 34(1963) 1265-1269.

[27] V. H a s s e l b l a d: Estimation of finite mixtures of distributions from the exponential family. J.A.S.A. 64(1969) 1459-1471.

Received May 3rd, 1972.

Address of the Author: prof. dr Włodzimierz Krysicki, Łódź, ul. Próchnika 25 m 7.