ASYMPTOTIC EFFICIENCY OF JOINT ESTIMATION OF PARAMETERS OF A COMPOUND POISSON DISTRIBUTION

A. G. Belov and V. Ya. Galkin

1. PRELIMINARY REMARKS

This is a numerical study of the asymptotic efficiency of parameter estimators of the compound Poisson distribution [1, p. 53; 2, p. 115] with the generating function [3; 4; 5, p. 15]

$$P(z)=\exp\Big\{\sum_{\nu=1}^{k}\lambda_\nu((\bar\varepsilon+\varepsilon z)^\nu-1)\Big\}=\exp\Big\{\sum_{\nu=1}^{k}\theta_\nu(z^\nu-1)\Big\}, \tag{1}$$

$\lambda_1,\dots,\lambda_k>0$, $0<\varepsilon\le 1$ is given, $\bar\varepsilon=1-\varepsilon$,

$$\theta_\nu=\varepsilon^\nu\sum_{\mu=\nu}^{k}C_\mu^\nu\bar\varepsilon^{\mu-\nu}\lambda_\mu,\quad \Big(\lambda_\nu=\sum_{\mu=\nu}^{k}C_\mu^\nu\varepsilon^{-\mu}(-\bar\varepsilon)^{\mu-\nu}\theta_\mu\Big),\quad \nu=\overline{1,k}, \tag{2}$$

or in matrix form

$$\theta=C\lambda,\quad C=C^{k\times k}=\|c_{\nu\mu}\|=\|\varepsilon^\nu C_\mu^\nu\bar\varepsilon^{\mu-\nu}\|. \tag{2'}$$

The probability function of this distribution has the form

$$p_n=P\{\xi=n\}=p_0\sum\prod_{\nu=1}^{k}\theta_\nu^{i_\nu}/i_\nu!,\quad n=0,1,\dots, \tag{3}$$

where $p_0=\exp\Big\{-\sum_{\nu=1}^{k}\theta_\nu\Big\}=\exp\Big\{\sum_{\nu=1}^{k}\lambda_\nu(\bar\varepsilon^\nu-1)\Big\}$, and the summation is over integer nonnegative solutions of $\sum_{\nu=1}^{k}\nu i_\nu=n$.

We use the following results for the distribution (3) [3-5]: the derivatives with respect to the parameters θ_ν, λ_ν, $\nu=1,\dots,k$, are representable in the form

$$\frac{\partial p_n}{\partial\theta_\nu}=\sum_{j=0}^{\min\{1,[n/\nu]\}}(-1)^{1-j}p_{n-\nu j},\quad \frac{\partial p_n}{\partial\lambda_\nu}=\sum_{j=0}^{\min\{\nu,n\}}C_\nu^j\varepsilon^j\bar\varepsilon^{\nu-j}p_{n-j}-p_n; \tag{4}$$

the probabilities p_n, $n=1,2,\dots$, are expressible by the recurrences

$$np_n=\sum_{\nu=1}^{\min\{k,n\}}\nu\theta_\nu p_{n-\nu}=\sum_{\mu=1}^{k}\lambda_\mu\sum_{\nu=1}^{\min\{\mu,n\}}\nu C_\mu^\nu\bar\varepsilon^{\mu-\nu}p_{n-\nu}; \tag{5}$$

the cumulants $\varkappa_r$, $r=1,2,\dots$, are expressible by

$$\varkappa_r=\sum_{\nu=1}^{k}\nu^r\theta_\nu=\sum_{\mu=1}^{k}\lambda_\mu\sum_{\nu=1}^{\mu}\nu^rC_\mu^\nu\varepsilon^\nu\bar\varepsilon^{\mu-\nu}=\sum_{l=1}^{r}\sigma_r^{(l)}\varepsilon^l\sum_{\mu=l}^{k}(\mu)_l\lambda_\mu, \tag{6}$$

where $\sigma_r^{(l)}$ are Stirling numbers of second kind, $(\mu)_l=\mu(\mu-1)\dots(\mu-l+1)$.

Translated from Chislennye Metody Resheniya Obratnykh Zadach Matematicheskoi Fiziki, pp. 46-57, 1988.

1046-283X/91/0202-0121$12.50

2. DETERMINANT OF INFORMATION MATRIX

The regularity conditions for the first and second derivatives with respect to λ_ν and θ_ν from (4) are easily verified. First consider the structure and the properties of the determinant of the information matrix $B = B_\lambda^{k\times k} = \|b_{\nu\mu}\|$, $b_{\nu\mu} = E_\lambda \partial^2 \ln L/\partial\lambda_\nu\partial\lambda_\mu$ ($L = \prod_{i=1}^{N} p_{n_i}$ is the likelihood matrix). The elements of this information matrix are given by

$$b_{\nu\mu} = -E_\lambda\, \partial^2 \ln L/\partial\lambda_\nu\partial\lambda_\mu = N\sum_{n=0}^{\infty} p_n^{-1}\partial p_n/\partial\lambda_\nu\, \partial p_n/\partial\lambda_\mu = Ni_{\nu\mu},$$

$$i_{\nu\mu} = \bar{\varepsilon}^{\nu+\mu}\sum_{i=0}^{\nu}\sum_{j=0}^{\mu} C_\nu^i C_\mu^j (\varepsilon/\bar{\varepsilon})^{i+j}\sum_{n=\max(i,j)}^{\infty} p_{n-i}p_{n-j}/p_n - 1 = \tag{7}$$

$$= \bar{\varepsilon}^{\nu+\mu}\sum_{i=1}^{\nu}\sum_{j=1}^{\mu} C_\nu^i C_\mu^j (\varepsilon\bar{\varepsilon}^{-1})^{i+j}\,\bar{i}_{ij},\quad \nu,\mu=\overline{1,k},$$

where $N\bar{i}_{ij} = N\left(\sum_{n=\max(i,j)}^{\infty} p_{n-i}p_{n-j}/p_n - 1\right)$ are the elements of the information matrix $B_\theta^{k\times k} = N\bar{I}_\theta^{k\times k}$ in the parameter system θ_ν, $B_\lambda^{k\times k} = NI_\lambda^{k\times k}$.

Proposition. For the determinant of the matrix $B_\lambda^{k\times k}$ with $k \geq 2$ we have the recurrences

$$|B| = \det B_\lambda^{k\times k} = \frac{N^k}{k^2\lambda_k^2}\begin{vmatrix} I_\lambda^{k-1\times k-1} & \begin{matrix}\varepsilon\\ 2\varepsilon\\ \vdots\\ (k-1)\varepsilon\end{matrix}\\ \varepsilon 2\varepsilon\ \dots\ (k-1)\varepsilon & \mu_2\end{vmatrix} = \tag{8}$$

$$= \varepsilon^{k(k+1)}\det B_\theta^{k\times k} = \frac{N^k\varepsilon^{k(k-1)}}{k^2\lambda_k^2}\begin{vmatrix} \bar{I}_\theta^{k-1\times k-1} & \begin{matrix}1\\ 2\\ \vdots\\ k-1\end{matrix}\\ 1\ 2\ \dots\ k-1 & \mu_2\end{vmatrix}.$$

To prove the first relationship, note that any element in kth row (column) of the matrix I_λ (7) is expressible in terms of the elements of the preceding rows (columns):

$$i_{k\nu} = i_{\nu k} = \frac{1}{k\lambda_k}\left(\varepsilon\nu - \sum_{\mu=1}^{k-1}\mu\lambda_\mu i_{\nu\mu} + \bar{\varepsilon}\sum_{\mu=2}^{k}\mu\lambda_\mu i_{\nu\mu-1}\right),\quad \nu=\overline{1,k}.$$

It suffices to substitute p_{n-k} from (4) in (7) and to carry out simple manipulations. For the last element

$$i_{kk} = \frac{1}{k^2\lambda_k^2}\left(2\varepsilon k^2\lambda_k - \mu_2 + \sum_{\mu=1}^{k-1}\sum_{l=1}^{k-1}\mu\lambda_\mu l\lambda_l i_{\mu l} - 2\bar{\varepsilon}\sum_{\mu=1}^{k-1}\sum_{l=2}^{k}\mu\lambda_\mu l\lambda_l i_{\mu l-1} + \bar{\varepsilon}^2\sum_{\mu=2}^{k}\sum_{l=2}^{k}\mu\lambda_\mu l\lambda_l i_{\mu-1\, l-1}\right).$$

Successively multiplying the rows and then the columns in I_λ by $\mu\lambda_\mu$, $\mu = 1, \dots, k-1$, and adding them to row k and then to column k, we obtain the determinant of a matrix with the last row (column)

$$i'_{k\nu} = i'_{\nu k} = \varepsilon\nu + \bar{\varepsilon}\sum_{\mu=2}^{k}\mu\lambda_\mu i_{\nu\mu-1},\quad \nu=\overline{1,k-1},$$

$$i'_{kk} = \varepsilon\sum_{\mu=1}^{k}\mu^2\lambda_\mu + \varepsilon\bar{\varepsilon}\sum_{\mu=2}^{k}\mu(\mu-1)\lambda_\mu + \bar{\varepsilon}^2\sum_{\mu=2}^{k}\sum_{l=2}^{k}\mu\lambda_\mu l\lambda_l i_{\mu-1\, l-1}.$$

Successively multiplying the first $k-1$ rows by $\bar{\varepsilon}(r+1)\lambda_{r+1}$, $r = 1, \dots, k-1$, subtracting them from the last row, and repeating the same procedure for the columns, we obtain the first equality in (8).

To verify the second equality in (8), we first factor out ε^2 from each row in $|B|$. In the first step, successively multiplying the first row by $\nu\bar{\varepsilon}^{\nu-1}$, $\nu = 2, \ldots, k$, we subtract it from the second, third, ..., kth row and factor out ε from each row; precisely the same procedure is applied to the first column. Then successively multiplying the second row by $C_\nu^{\,2}\bar{\varepsilon}^{\nu-2}$, $\nu = 3, \ldots, k$, we subtract it from the third, fourth, ..., kth row and factor out ε from each row; and similarly for the second column. Continuing in the same way, in the $(k-1)$-th step we multiply the $(k-1)$-th row by $C_k^{\,k-1}\bar{\varepsilon}$, subtract it from the kth row and factor out ε; the same combination of columns $(k-1)$ and k gives the multiplier ε in the determinant, as required.

Finally, by (5), write the last row and column in $|B_\theta^{k\times k}| = N^k |\bar{\mathrm{I}}_\theta^{k\times k}|$ in the form

$$\bar{i}_{k\nu} = \bar{i}_{\nu k} = k^{-1}\theta_k^{-1}\left(\nu - \sum_{r=1}^{k-1} r\theta_r \bar{i}_{\nu r}\right), \quad \nu = \overline{1, k-1},$$

$$\bar{i}_{kk} = k^{-1}\theta_k^{-1}\left(k - k^{-1}\theta_k^{-1}\left(\sum_{r=1}^{k-1} r\theta_r\left(r - \sum_{l=1}^{k-1} l\theta_l \bar{i}_{lr}\right)\right)\right).$$

Multiply the first $k-1$ rows and columns by $r\theta_r$, $r = 1, \ldots, k-1$, and add to the kth row and column respectively. Seeing that from (2) $\theta_k = \varepsilon^k\lambda_k$, we obtain the third equality in (8).

Note that for $k = 1$ the distribution (1), (3) coincides with the Poisson distribution $\mathrm{Po}(\varepsilon\lambda_1)$ and clearly

$$\det B_{\lambda_1}^{1\times 1} = N i_{11} = N\varepsilon^2 \bar{i}_{11} = N\varepsilon^2\left(\sum_{n=1}^{\infty} p_{n-1}^2/p_n - 1\right) = N\varepsilon/\lambda_1.$$

Let us now investigate the asymptotic efficiency of various consistent estimators $\tilde{\lambda}$, which is defined [6, p. 389] as

$$e_0(\tilde{\lambda}) = e_0(\tilde{\lambda}\,|\,\lambda) = (\lim_{N\to\infty} |B|\,|V(\tilde{\lambda})|)^{-1},$$

where $V(\boldsymbol{\lambda}) = \|\mathrm{cov}(\lambda_\nu, \lambda_\mu)\|$ is the estimate covariance matrix. As alternatives to the maximum likelihood method [9], which produces asymptotically efficient estimators, we consider five simpler methods for estimation of the parameters $\boldsymbol{\lambda}$ of the distribution (3).

3. METHOD OF MOMENTS ESTIMATORS

Using (6), we see that the method of moments estimators λ^{m} for our distribution can be obtained from the system of linear algebraic equations

$$WC\boldsymbol{\lambda} = \mathbf{k},$$

where $W = \|W^{k\times k}\| = \|\mu^\nu\|$, $\nu, \mu = 1, \ldots, k$, is the Vandermonde matrix, $\mathbf{k}$ is the column vector of the sample cumulants. Then

$$V(\boldsymbol{\lambda}^{\mathrm{M}}) = V(C^{-1}W^{-1}\mathbf{k}) = C^{-1}W^{-1}V(\mathbf{k})(W^{-1}C^{-1})',$$
$$e_0(\boldsymbol{\lambda}^{\mathrm{M}}) = |W|^2|C|^2/|B|\,|V(\mathbf{k})|.$$

In the important experimental case $k = 2$ [7, 8], taking $1 < m_2/a_1 < 1 + \varepsilon$, we obtain

$$\lambda_1^{\mathrm{m}} = \varepsilon^{-2}((1+\varepsilon)a_1 - m_2), \quad \lambda_2^{\mathrm{m}} = 1/2\varepsilon^{-2}(m_2 - a_1),$$

where a_1 is the sample mean, m_2 is the sample second central moment. The estimate bias is written in the form

$$E(\lambda_1^{\mathrm{m}} - \lambda_1) = \mu_2(\varepsilon^2 N)^{-1}, \quad E(\lambda_2^{\mathrm{m}} - \lambda_2) = -\mu_2(2\varepsilon^2 N)^{-1},$$

and the elements of the estimate covariance matrix are accurate to $O(N^{-2})$:

$$\varepsilon^4 N D\lambda_1^{\mathrm{m}} = \varepsilon^2\mu_2 + 4(1-2\varepsilon)\theta_2 + 2\mu_2^2 + O(N^{-1}),$$
$$\varepsilon^4 N \,\mathrm{cov}(\lambda_1^{\mathrm{m}}, \lambda_2^{\mathrm{m}}) = -2\bar{\varepsilon}\theta_2 - \mu_2^2 + O(N^{-1}),$$
$$2\varepsilon^4 N D\lambda_2^{\mathrm{m}} = 2\theta_2 + \mu_2^2 + O(N^{-1}),$$

TABLE 1. $\varepsilon = 0.1$

λ_2 \ λ_1	0,1	0,3	0,5	0,8	1,0	1,5	2	3	4	5
0,1	92	94	96	97	98	99	100	100	100	100
0,3	87	89	91	94	95	97	98	99	100	100
0,5	85	87	89	91	93	96	97	98	99	100
0,8	85	87	89	91	92	94	96	98	99	99
1,0	85	87	88	90	92	94	95	97	98	99
1,5	86	87	89	90	91	93	95	96	98	98
2,0	87	88	90	91	92	93	94	96	97	98
3,0	90	91	91	92	93	94	95	96	97	98
4,0	91	92	92	93	93	94	95	96	97	97
5,0	93	93	93	94	94	95	95	96	97	97
10,0	96	96	96	96	96	97	97	97	97	98

TABLE 2. $\varepsilon = 1$

λ_2 \ λ_1	0,1	0,3	0,50	0,8	1,0	1,5	2,0	3,0	4,0	5,0
0,1	18	38	57	78	85	94	97	99	99	100
0,3	3	11	22	43	57	78	87	94	97	98
0,5	1	5	11	26	39	66	80	91	95	97
0,8	1	2	5	15	25	57	74	87	92	94
1,0		2	4	11	19	47	68	84	90	93
1,5		1	2	6	11	35	61	82	88	91
2,0			1	3	7	27	56	81	87	90
3,0				1	4	17	47	81	88	90
4,0				1	2	12	39	82	89	91
5,0				1	1	8	33	83	90	92
10,0						2	14	80	93	94

where $\mu_2 = \varkappa_2 = \theta_1 + 4\theta_2 = \varepsilon\lambda_1 + 2\varepsilon(1 + \varepsilon)\lambda_2$, $\theta_1 = \varepsilon\lambda_1 + 2\varepsilon\bar{\varepsilon}\lambda_2$, $\theta_2 = \varepsilon^2\lambda_2$. Denoting $\Delta = |\bar{I}^{2\times 2}| = \mu_2 i_{11} - 1$ and seeing that from (8) $|B_\lambda^{2\times 2}| = N^2\varepsilon^2\Delta/4\lambda_2^2$, we obtain an explicit expression for the asymptotic efficiency:

$$e_0(\lambda_1^{m}, \lambda_2^{m}) = 8\theta_2^2/\Delta(2\theta_1\theta_2 + \mu_2^3). \quad (9)$$

The computation of asymptotic efficiency thus reduces to summation of one numerical series

$$\bar{i}_{11} = \sum_{n=1}^{\infty} p_{n-1}^2/p_n - 1.$$

The values of e_0 as a function of the parameters λ_1, λ_2 for various ε are listed in Tables 1 and 2 (in percent) and are shown in the form of e_0 level lines in Fig. 1. We see that for small ε the estimators λ_1^m, λ_2^m are almost asymptotically efficient for virtually all λ_1, λ_2. As ε increases, a low-efficiency region ($e_0 < 0.9$) clearly emerges for small λ_1, λ_2. For $\varepsilon = 1$ (see Table 2, the blank entries are zeros), the low-efficiency region is essentially $\lambda_1 < 4$.

4. ESTIMATORS USING EVEN FREQUENCIES AND THE SAMPLE MEAN

For any discrete distribution, from the definition of generating function we obtain $P(1) + P(-1) = 2\sum_{n=0}^{\infty} p_{2n}$, i.e., if we write $P_e = \sum_{n=0}^{\infty} p_{2n}$, then $P(-1) = 2P_e - 1$. In our case ($k = 2$ by (1)), denoting the sum of even frequencies by $H_e = \sum_n h_{2n}$ this gives $\varepsilon\lambda_1 + 2\varepsilon\bar{\varepsilon}\lambda_2 = \frac{1}{2}\ln(2H_e - 1)$. Combined with the equation $\varkappa_1 = a_1$, this leads to the "even-frequency" estimators

$$\lambda_1^{ef} = -(\ln(2H_e - 1) + 2\bar{\varepsilon}a_1)/2\varepsilon^2, \quad \lambda_2^{ef} = (\ln(2H_e - 1) + 2\alpha_1)/4\varepsilon^2,$$

which are admissible for $\frac{1}{2}(1 + e^{-2a_1}) < H_e < \frac{1}{2}(1 + e^{-2\bar{\varepsilon}a_1})$.

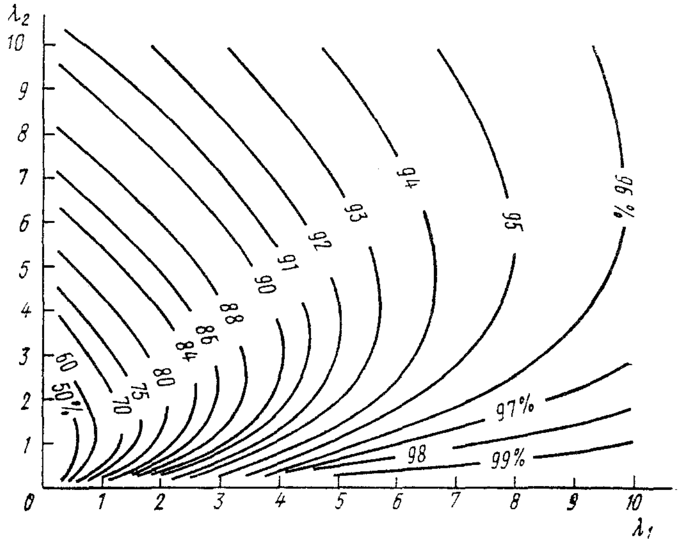

Fig. 1. $\varepsilon = 0.5$.

It is easy to verify that up to $o(N^{-1})$

$$E(\lambda_1^{ef} - \lambda_1) = (e^{4\theta_1} - 1)/4\varepsilon^2 N, \quad E(\lambda_2^{ef} - \lambda) = (1 - e^{4\theta_1})/8\varepsilon^2 N.$$

To find the estimate variances (functions of a_1, H_e), we need DH_e and $\mathrm{cov}(a_1, H_e)$ in addition to Da_1 [12]. Clearly,

$$DH_e = \sum_n Dh_{2n} + 2\sum_{i<j} \mathrm{cov}(h_{2i}, h_{2j}) = N^{-1}P_e(1 - P_e),$$

$$\mathrm{cov}(a_1, H_e) = E\Big(\sum_n nh_n - \varkappa_1\Big)(H_e - P_e) = \sum_n n\,\mathrm{cov}(h_n, H_e).$$

But it is easy to show that

$$\mathrm{cov}(h_n, H_e) = \begin{cases} N^{-1}p_{2i}(1 - P_e), & n = 2i, \\ N^{-1}p_{2i+1}P_e, & n = 2i + 1, \; i = 0, 1, \ldots . \end{cases}$$

Hence $\mathrm{cov}(a_1, H_e) = N^{-1}\Big(\sum_{i=0}^{\infty} 2ip_{2i} - \varkappa_1 P_e\Big)$. Now note that for any discrete distribution $2\sum_{i=0}^{\infty} 2ip_{2i} = P'(1) - P'(-1)$.

Finally, for our case

$$\mathrm{cov}(a_1, H_e) = \varepsilon(\lambda_1 + 2\bar{\varepsilon}\lambda_2)(1 - 2P_e)N^{-1} = -\theta_1 e^{-2\theta_1}N^{-1}.$$

We can now write out the elements of the matrix $\|\mathrm{cov}(\lambda_\nu^{ef}, \lambda_\mu^{ef})\|$:

$$\varepsilon^4 ND\lambda_1^{ef} = 1/4(e^{4\theta_1} - 1) - (1 - \varepsilon^2)\theta_1 + 4\bar{\varepsilon}^2\theta_2^2 + O(N^{-1/2}),$$

$$2\varepsilon^4 N\,\mathrm{cov}(\lambda_1^{ef}, \lambda_2^{ef}) = \theta_1 - 4\bar{\varepsilon}\theta_2 - 1/4(e^{4\theta_1} - 1) + O(N^{-1/2}),$$

$$4\varepsilon^4 ND\lambda_2^{ef} = 1/4(e^{4\theta_1} - 1) - \theta_1 + 4\theta_2 + O(N^{-1/2}).$$

Note that $\mathrm{cov}(\lambda_1^{ef}, \lambda_2^{ef})$ is best computed from $\lambda_1^{ef} + 2\lambda_2^{ef} = a_1\varepsilon^{-1}$.

Thus, for the asymptotic efficiency of the even-frequency estimators we have

$$e_0(\lambda_1^{ef}, \lambda_2^{ef}) = 64\theta_2^2/\Delta(\mu_2(e^{4\theta_1} - 1) - 4\theta_1^2). \tag{10}$$

The computed values of e_0 (10) are shown in Fig. 2 (the level lines of the function $e_0(\lambda_1^{ef}, \lambda_2^{ef} \mid \lambda_1, \lambda_2))$ and in Table 3 (the blank entries are zeros). We see that the high-efficiency region ($e_0 > 0.9$) is observed only for small λ_1, λ_2, and ε. Thus, for $\varepsilon = 0.1$, we have $e_0 > 0.9$ in the triangle $\lambda_1 + \lambda_2 < 1.5$. As ε increases, the asymptotic efficiency decreases, and the high-efficiency regions shifts to small λ_1, λ_2. For example, for $\varepsilon = 0.5$, we have $e_0 > 0.9$ in the triangle $2\lambda_1 + \lambda_2 < 1$.

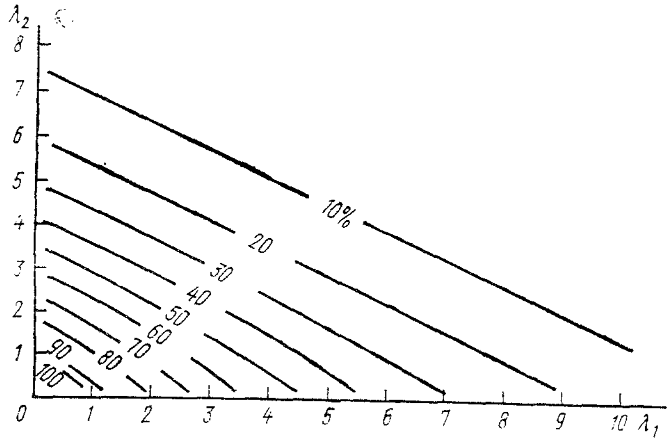

Fig. 2. $\varepsilon = 0.1$.

TABLE 3. $\varepsilon = 0.5$.

λ_2 \ λ_1	0,1	0,3	0,5	0,8	1,0	1,5	2,0	3,0	4,0	5,0
0,1	100	96	87	68	56	31	17	5	1	0
0,3	99	95	87	70	58	33	17	4	1	
0,5	97	92	83	66	55	31	16	4	1	
0,8	92	84	74	58	47	26	13	3	1	
1,0	87	78	68	51	41	22	11	2		
1,5	70	60	49	35	27	14	6	1		
2,0	51	41	33	22	17	8	4	1		
3,0	20	15	11	7	5	2	1			
4,0	6	4	3	2	1	1				
5,0	1	1	1	0						

5. ESTIMATORS USING THE ZEROTH FREQUENCY AND THE SAMPLE MEAN

Here $p_0 = h_0$ and $\varkappa_1 = a_1$. We write these estimators in explicit form with the corresponding admissibility condition:

$$\left.\begin{aligned}\lambda_1^0 &= \varepsilon^{-2}\left((\varepsilon-2)a_1 - 2\ln h_0\right),\\ \lambda_2^0 &= \varepsilon^{-2}(a_1 + \ln h_0)\end{aligned}\right\},\quad 1-\varepsilon/2 < -\ln h_0/a_1 < 1.$$

The expressions for biases, variances, and covariances are written in the form

$$E(\lambda_1^0-\lambda_1) = (1-p_0)/\varepsilon^2 p_0 N + o(N^{-1}),\quad E(\lambda_2^0-\lambda_2) = (p_0-1)/2\varepsilon^2 p_0 N + o(N^{-1}),$$

$$\varepsilon^4 N D\lambda_1^0 = 4(p_0^{-1}-1) + \varepsilon(\varepsilon-2)\left((\varepsilon+2)\lambda_1 + 2(2-\varepsilon\bar{\varepsilon})\lambda_2\right) + O(N^{-1/2}),$$

$$\varepsilon^4 N \operatorname{cov}(\lambda_1^0, \lambda_2^0) = 1 - p_0^{-1} + \varepsilon\left(\lambda_1 + (2-2\varepsilon+\varepsilon^2)\lambda_2\right) + O(N^{-1/2}),$$

$$\varepsilon^4 N D\lambda_2^0 = p_0^{-1} - 1 - \varepsilon(\lambda_1 + 2\bar{\varepsilon}\lambda_2) + O(N^{-1/2}).$$

The asymptotic efficiency of the estimators is thus given by

$$e_0(\lambda_1^0,\ \lambda_2^0) = 4p_0\theta_2^2/\Delta(\mu_2 - p_0\alpha_2), \tag{11}$$

where $\alpha_2 = \theta_1 + 4\theta_2 + (\theta_1 + 2\theta_2)^2$ is the second initial moment.

Analysis of the computation results for $e_0(\lambda_1^0, \lambda_2^0)$ as a function of λ_1, λ_2, and ε shows that for large ε the high-efficiency region is localized. For instance, $0.5 < \lambda_1 + \lambda_2 < 3$ for $\varepsilon = 0.5$ and $\lambda_1 < 1.5$, $0.4 < \lambda_1 - \lambda_2 < 0.8$ for $\varepsilon = 1$. As ε decreases, this region becomes substantially larger: for $\varepsilon = 0.1$, we have $\lambda_1 + \lambda_2 < 4$ (see level lines in Fig. 3).

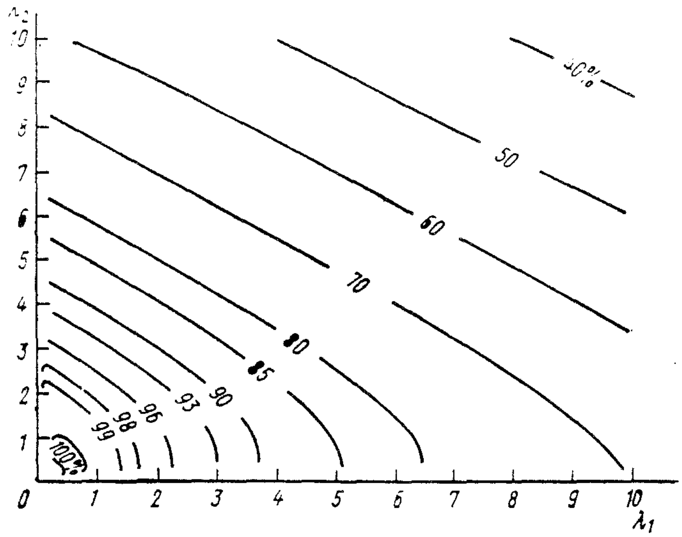

Fig. 3. $\varepsilon = 0.1$.

6. ESTIMATORS USING THE SAMPLE MEAN AND THE RATIO OF FIRST TWO FREQUENCIES

In this case ($\varkappa_1 = a_1, p_1/p_0 = h_1/h_0$), the estimators are written in the form

$$\lambda_1^{rf} = \varepsilon^{-2}(h_1/h_0 - \bar{\varepsilon}a_1),\ \lambda_2^{rf} = \frac{1}{2}\varepsilon^{-2}(a_1 - h_1/h_0),\ (a_1\bar{\varepsilon} < h_1/h_0 < a_1).$$

Here

$$E(\lambda_1^{rf} - \lambda_1) = p_1/\varepsilon^2 p_0^2 N + o(N^{-1}),\ E(\lambda_2^{rf} - \lambda_2) = -p_1/2\varepsilon^2 p_0^2 N + o(N^{-1}),$$

$$p_0^3\varepsilon^4 ND\lambda_1^{rf} = p_1^2 + p_0 p_1 + p_0^3\bar{\varepsilon}^2\mu_2 + O(N^{-1/2}),$$

$$4p_0^3\varepsilon^4 N\operatorname{cov}(\lambda_1^{rf}, \lambda_2^{rf}) = -p_1^2 - p_0 p_1 - p_0^3\bar{\varepsilon}\mu_2 + O(N^{-1/2}),$$

$$4p_0^3\varepsilon^4 ND\lambda_2^{rf} = p_1^2 + p_0 p_1 + p_0^3\mu_2 + O(N^{-1/2}),$$

where $p_1 = \theta_1 p_0 = \varepsilon(\lambda_1 + 2\bar{\varepsilon}\lambda_2)p_0$, and therefore

$$e_0(\lambda_1^{rf}, \lambda_2^{rf}) = 16p_0^3\theta_2^2/\Delta\mu_2 p_1(p_0 + p_1). \tag{12}$$

Table 4 lists the values of e_0 (12) (in percent) for $\varepsilon = 0.1$. We see that $e_0 < 0.6$ for all λ_1, λ_2. As ε increases, e_0 increases. Thus, for $\varepsilon = 0.5$, we have $0.8 < e_0 < 0.9$ in the parallelogram $1 < \lambda_1 + \lambda_2 < 2, \lambda_1 < 1$ (see level lines in Fig. 4).

7. ESTIMATORS USING THE FIRST TWO FREQUENCIES

In this case ($p_0 = h_0, p_1 = h_1$), the estimators have the form

$$\lambda_1^{f} = \varepsilon^{-2}(2\bar{\varepsilon}\ln h_0 + (1+\bar{\varepsilon})h_1/h_0),\quad \lambda_2^{f} = -\varepsilon^{-2}(\ln h_0 + h_1/h_0)$$

and are admissible for $1 < -h_0 \ln h_0/h_1 < \frac{1}{2}(1 + \bar{\varepsilon}^{-1})$.

Their biases are expressed up to $o(N^{-1})$ by

$$E(\lambda_1^{f} - \lambda_1) = ((1+\bar{\varepsilon})p_1 - \bar{\varepsilon}p_0(1-p_0))/\varepsilon^2 p_0^2 N,\ E(\lambda_2^{f} - \lambda_2) = (p_0(1-p_0) - 2p_1)/2\varepsilon^2 p_0^2 N,$$

and the elements of the covariance matrix are

$$p_0^3\varepsilon^4 ND\lambda_1^{f} = 4\bar{\varepsilon}^2 p_0^2(1-p_0) + (1+\bar{\varepsilon})(1-3\bar{\varepsilon})p_0 p_1 + (1+\bar{\varepsilon})^2 p_1^2 + O(N^{-1/2}),$$

$$p_0^3\varepsilon^4 N\operatorname{cov}(\lambda_1^{f}, \lambda_2^{f}) = 2\bar{\varepsilon}p_0 p_1 - 2\bar{\varepsilon}p_0^2(1-p_0) + (\varepsilon - 2)p_1^2 + O(N^{-1/2}),$$

$$p_0^3\varepsilon^4 ND\lambda_2^{f} = p_1(p_1 - p_0) + p_0^2(1-p_0) + O(N^{-1/2}).$$

TABLE 4. $\varepsilon = 0.1$

λ_2 \ λ_1	0,1	0,3	0,5	0,8	1,0	1,5	2	3	4
0,1	17	16	17	20	22	27	31	37	40
0,3	26	26	27	[illegible]	30	34	37	41	43
0,5	32	32	33	[illegible]	36	39	41	44	45
0,8	39	40	40	42	43	44	46	47	47
1,0	43	43	44	45	46	47	48	48	48
1,5	50	51	51	51	52	52	52	51	49
2	55	55	55	55	55	55	54	52	49
3	58	58	58	57	57	55	54	51	47
4	57	57	56	55	54	53	51	47	44
5	54	53	52	51	50	48	46	43	39

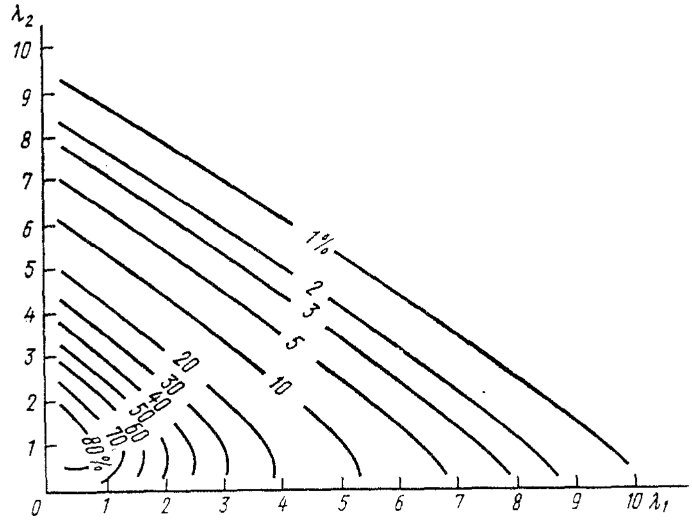

Fig. 4. $\varepsilon = 0.5$.

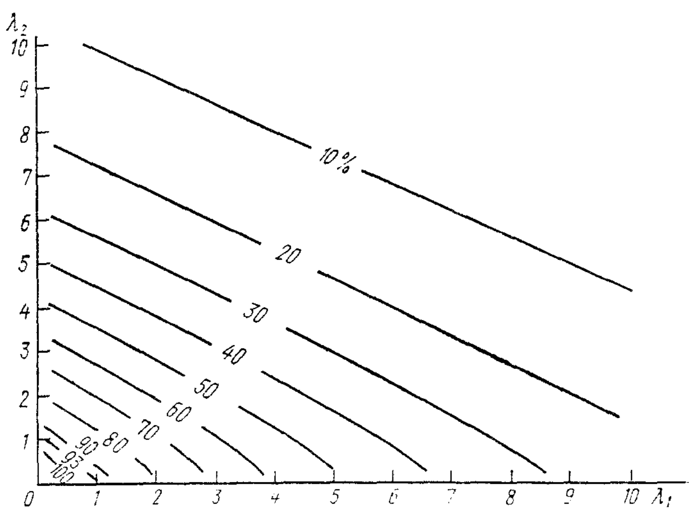

Fig. 5. $\varepsilon = 0.1$

The asymptotic efficiency in this case is given by

$$e_0(\lambda_1^{f}, \lambda_2^{f}) = 4p_0^3\theta_2^2/\Delta p_1(1-p_0-p_1). \tag{13}$$

Numerical calculations show that high-efficiency regions are observed only for small λ_1, λ_2, e.g., for $\varepsilon = 0.1$ we have $e_0 > 0.9$ in the region $\lambda_1 + \lambda_2 < 1.5$ (Fig. 5). With the increase of ε the efficiency declines and the region where $e_0 > 0.9$ becomes smaller. Thus, for $\varepsilon = 0.5$, we have $\lambda_1 + \lambda_2 < 0.5$.

8. COMPARATIVE ANALYSIS

Our numerical analysis of asymptotic efficiencies leads to the following conclusions. For small ε, the most efficient estimates are λ_1^m, λ_2^m. Other estimators can be used only for small parameter values. For instance, for $\varepsilon = 0.1$, we have $\lambda_1^{ef} + \lambda_2^{ef} < 1.5$, $\lambda_1^0 + \lambda_2^0 < 4$, $\lambda_1^e + \lambda_2^e < 1.5$. As ε increases, the low-efficiency region for λ_1^m, λ_2^m becomes larger, but it is completely covered by the high-efficiency regions of the estimators λ_1^0, λ_2^0 and λ_1^{rf}, λ_2^{rf} (see Figs. 1 and 4). Thus, for small λ_1, λ_2, the estimators λ_1^0, λ_2^0 and λ_1^{rf}, λ_2^{rf} are efficient for any ε in addition to λ_1^m, λ^{2m}.

LITERATURE CITED

1. Yu. V. Prokhorov and Yu. A. Rozanov, Probability Theory [in Russian], Nauka, Moscow (1973).
2. V. S. Korolyuk, N. I. Portenko, A. V. Skorokhod, and A. F. Turbin, Handbook of Probability Theory and Mathematical Statistics [in Russian], Nauka, Moscow (1985).
3. V. Ya. Galkin and M. V. Ufimtsev, "Analysis of direct stochastic problems for recording the yield of multiple nuclear processes," in: Some Topics of Computer-Aided Analysis and Interpretation of Physical Experiments [in Russian], No. 2, Izd. Moscow State Univ. (1973), pp. 81-116.
4. V. Ya. Galkin, "Direct problems for separation of multiple processes," Dokl. Akad. Nauk SSSR, **216**, No. 5, 1014-1017 (1974).
5. A. G. Belov, V. Ya. Galkin, and M. V. Ufimtsev, Probabilistic-Statistical Problems in Experimental Separation of Multiple Processes [in Russian], Moscow State Univ. (1985).
6. S. Wilks, Mathematical Statistics [Russian translation], Nauka, Moscow (1967).
7. V. Ya. Galkin, B. I. Goryachev, V. N. Orlin, and M. V. Ufimtsev, "On optimal analysis and organization of experiments with statistical separation of the yields of nuclear reactions of various multiplicities," in: Computational Methods and Programming [in Russian], No. 18, Moscow State Univ. (1972), pp. 161-172.
8. V. Ya. Galkin, Direct and Inverse Problems for Separation of Multiple Nuclear Processes [in Russian], Preprint D10-77-07 OIYaI, Izd. OIYaI, Dubna (1974), pp. 93-99.
9. V. Ya. Galkin and M. V. Ufimtsev, "Inverse problems in recording the yield of multiple nuclear reactions," in: Some Topics of Computer-Aided Analysis and Interpretation of Physical Experiments [in Russian], No. 3, Moscow State Univ. (1975), pp. 3-26.
10. A. W. Kemp and C. D. Kemp, "Some properties of Hermite distribution," Biometrika, **52**, 381-394 (1965).
11. Y. C. Patel, "Even point estimation and moment estimation in Hermite distribution," Biometrics, **32**, 865-873 (1976).
12. H. Cramer, Mathematical Methods of Statistics [Russian translation], Mir, Moscow (1975).