ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
21,4 (1993), pp. 503–510

S. LOUKAS (Ioannina)

SOME METHODS OF ESTIMATION FOR A TRIVARIATE POISSON DISTRIBUTION

Abstract. A four-parameter trivariate Poisson distribution was considered by Loukas and Papageorgiou [6] where various properties of the distribution were examined and the maximum likelihood estimation was obtained by iteration techniques.

In the present paper we study several alternative estimation procedures which lead to explicit expressions of the parameter estimates in terms of simple sample statistics. Asymptotic efficiencies are examined and recommendations made. Applications of the estimation methods to simulated data sets are also given.

1. Introduction. Multivariate generalizations of the Poisson distribution, with $2^k - 1$ non-negative parameters for the k-variate Poisson distribution, as limiting forms of multivariate binomial distributions were considered by Krishnamoorthy [4], Teicher [15] and Kawamura [3] and the corresponding seven-parameter trivariate Poisson distribution by Mahamunulu [7], Kawamura [2] and Loukas and Kemp [5]. A simpler version of the k-variate Poisson distribution with $k+1$ parameters was considered by Johnson and Kotz [1, §11.4] and Šidák [10]–[12]. In this paper, a parameter estimation for the corresponding four-parameter trivariate Poisson distribution is studied.

The random vector (X, Y, Z) has the trivariate Poisson distribution with parameters a, b, c, d if

$$X = X' + T\,, \quad Y = Y' + T\,, \quad Z = Z' + T\,, \tag{1}$$

where X', Y', Z' and T are independent Poisson variables with parameters $a - d$, $b - d$, $c - d$ and d, respectively. The probability generating function

1991 *Mathematics Subject Classification*: 62E15, 62F10.

Key words and phrases: trivariate Poisson distribution; estimation; method of moments; method of zero frequency; method of even points; asymptotic efficiency; simulation.

(p.g.f.) of the trivariate Poisson with the structure (1) is

$$(2)\quad g_{x,y,z}(u,v,w) = \exp\{(a-d)(u-1)+(b-d)(v-1)+(c-d)(w-1)+d(uvw-1)\}\,,$$

where $a, b, c > 0$ and $0 < d < \min(a, b, c)$. The marginal distributions are Poisson with $E(X) = a$, $E(Y) = b$, $E(Z) = c$ and $\mathrm{Cov}(X,Y) = \mathrm{Cov}(X,Z) = \mathrm{Cov}(Y,Z) = d$.

In [6] various properties of the trivariate Poisson distribution were examined, maximum likelihood estimators of the parameters were obtained by iteration techniques, and the information matrix was derived.

In this paper, we examine some alternative estimation procedures, namely the methods of moments, zero frequency and even points, which lead to explicit expressions of the parameters in terms of simple sample statistics. Computations of asymptotic efficiencies are also included. These simple ad hoc estimation techniques often provide reasonable alternatives in terms of the trade-off between modest computational requirements and an adequate level of asymptotic efficiency over certain regions of the parameter space. Finally, applications of the estimation methods to trivariate Poisson simulated data are also given.

2. The method of moments.

Since

$$\mu_{111} = E(X-E(X))(Y-E(Y))(Z-E(Z)) = d$$

for the trivariate Poisson distribution with p.g.f. given by (2), moment estimators can be obtained by using the equations

$$\widehat{a} = \overline{x}\,,\quad \widehat{b} = \overline{y}\,,\quad \widehat{c} = \overline{z}\,,$$
$$\widehat{d} = m_{111} = \frac{1}{n}\sum_{x,y,z}(x-\overline{x})(y-\overline{y})(z-\overline{z})\,,$$

where n represents the sample size and m_{111} must satisfy $0 < m_{111} < \min(\overline{x}, \overline{y}, \overline{z})$.

Asymptotic efficiency is defined as $E = |I^{-1}|/|V|$ where $|I^{-1}|$, the determinant of the inverse of the information matrix for the maximum likelihood estimators, is given by (cf. [6])

$$|I^{-1}| = \frac{n^{-4}(a-d)^2(b-d)^2(c-d)^2}{(abc-ad^2-bd^2-cd^2+2d^3)(Q-1)-(ab+ac+bc-2ad-2bd-2cd+3d^2))},$$

where $P_{x,y,z} = P(X=x, Y=y, Z=z)$ and $Q = \sum_{x,y,z} P^2_{x-1,y-1,z-1}/P_{x,y,z}$.

The determinant of the asymptotic covariance matrix for the moment estimates is given by

$$(3)\qquad |V| = |M|/|G|^2\,,$$

where M is the asymptotic covariance matrix of the sample moments and G is the Jacobian matrix of the transformation from the moment estimators to the sample moments.

It can be easily shown that $|G| = 1$.

Formulae for calculating most of the elements of M are given by Stuart and Ord [13] and Subrahmaniam and Subrahmaniam [14]. In addition,

$$\begin{aligned}\operatorname{Cov}(m_{111}, \bar{x}) = {} & \operatorname{Cov}\left(\frac{1}{n}\sum_{x,y,z} xyz, \bar{x}\right) - \operatorname{Cov}\left(\frac{1}{n}\bar{x}\sum_{y,z} yz, \bar{x}\right) \\ & - \operatorname{Cov}\left(\frac{1}{n}\bar{y}\sum_{x,z} xz, \bar{x}\right) - \operatorname{Cov}\left(\frac{1}{n}\bar{z}\sum_{x,y} xy, \bar{x}\right) \\ & + \operatorname{Cov}(2\bar{x}\,\bar{y}\,\bar{z}, \bar{x}).\end{aligned}$$

After some tedious algebra we obtain

$$\begin{aligned}\operatorname{Cov}(m_{111}, \bar{x}) &= \frac{1}{n}(\mu'_{211} - \mu'_{210}\mu'_{001} - \mu'_{201}\mu'_{010} - \mu'_{200}\mu'_{011} \\ &\quad - 2\mu'_{111}\mu'_{100} - 2\mu'_{110}\mu'_{101} + 2\mu'_{011}\mu'^{2}_{100} + 4\mu'_{110}\mu'_{100}\mu'_{001} \\ &\quad + 4\mu'_{101}\mu'_{100}\mu'_{010} + 2\mu'_{200}\mu'_{010}\mu'_{001} - 6\mu'^{2}_{100}\mu'_{010}\mu'_{001}) \\ &= k_{211}/n = d/n,\end{aligned}$$

where $\mu'_{\dots}$ and $k_{\dots}$ are the trivariate moments about the origin and cumulants, respectively. Similarly, $\operatorname{Cov}(m_{111}, \bar{y}) = k_{121}/n = d/n$ and $\operatorname{Cov}(m_{111}, \bar{z}) = k_{112}/n = d/n$. Also, following Stuart and Ord [13, Ch. 13] we obtain

$$\begin{aligned}\operatorname{Var}(m_{111}) &= \frac{1}{n}(k_{222} + k_{200}k_{022} + k_{020}k_{202} + k_{002}k_{220} + 2k_{110}k_{112} \\ &\quad + 2k_{101}k_{121} + 2k_{011}k_{211} + 3k_{111}^2 + 2k_{210}k_{012} + 2k_{120}k_{102} \\ &\quad + 2k_{201}k_{021} + k_{200}k_{020}k_{002} + 2k_{110}k_{101}k_{011} + k_{200}k_{011}^2 \\ &\quad + k_{020}k_{101}^2 + k_{002}k_{110}^2) \\ &= \frac{1}{n}(abc + ad^2 + bd^2 + cd^2 + 2d^3 + ad + bd + cd + 15d^2 + d) \\ &\equiv A.\end{aligned}$$

Finally, the determinant of M is found to be

$$\begin{aligned}|M| = \frac{1}{n^4}\{A(a-d)(b-d)(c-d) - d(d-A)[(a-d)(b-d) \\ +(a-d)(c-d) + (b-d)(c-d)]\}.\end{aligned}$$

The asymptotic efficiency of the method was examined for $c = b = \lambda a$, $d = \nu \min(a, b, c)$, where $\lambda = 1/4, 1/2, 1, 2, 4$, $\nu = 0.01, 0.02, \dots, 0.99$ and $a = 0.005, 0.01, \dots, 2$. A selection of the results for the symmetric case

$a = b = c$ is given in Figure 1. The method is highly efficient for small values of a, b, c and small values of d relative to $\min(a, b, c)$.

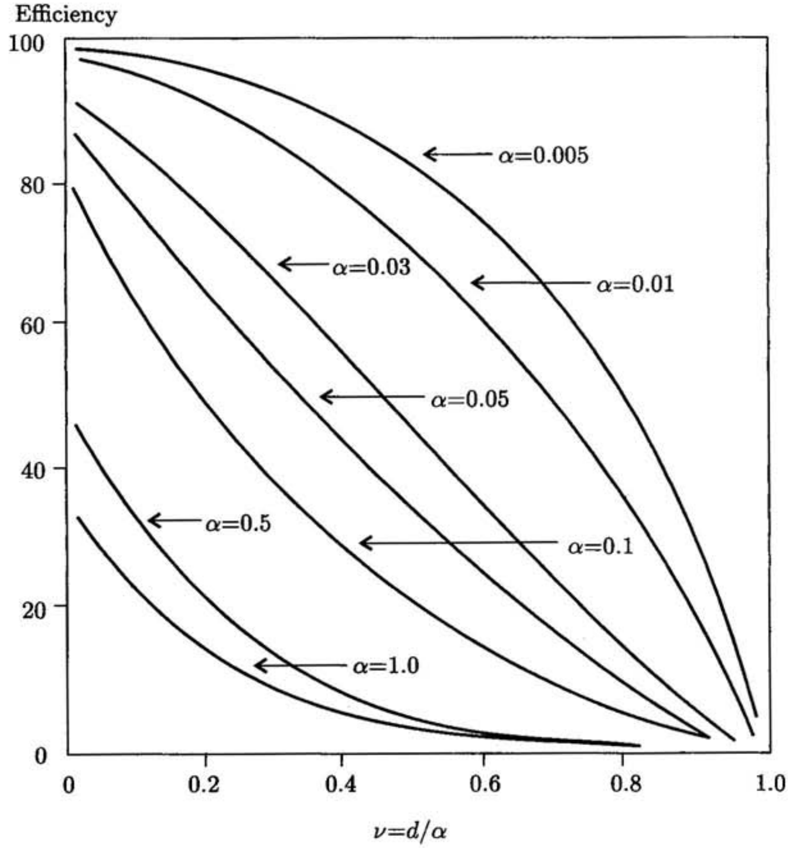

Fig. 1. Efficiency of the method of moments

3. The method of zero frequency. If we denote by f_{000} the observed proportion of observations in the $(0, 0, 0)$ cell, zero frequency estimators can be obtained by using the equations

$$\widehat{a} = \overline{x}\,, \quad \widehat{b} = \overline{y}\,, \quad \widehat{c} = \overline{z}\,,$$
$$\widehat{d} = \tfrac{1}{2}(\overline{x} + \overline{y} + \overline{z} + \log f_{000})\,,$$

where $f_{000} > 0$.

The determinant of the asymptotic covariance matrix for the zero frequency estimates can be obtained by using expression (3). From the well-known expressions for variances and covariances of proportions we have

$$\mathrm{Var}(f_{000}) = \frac{1}{n} e^{-B}(1 - e^{-B})\,, \quad \mathrm{Cov}(\overline{x}, f_{000}) = -\frac{1}{n} a e^{-B}\,,$$

and finally,

$$|M| = \frac{1}{n^4}[e^{-2B}(a^2d^2 + b^2d^2 + c^2d^2 + 6abcd \\ - a^2bc - ab^2c - abc^2 - 2abd^2 - 2acd^2 - 2bcd^2) \\ + e^{-B}(1 - e^{-B})(abc - ad^2 - bd^2 - cd^2 + 2d^3)]\,,$$

where $B = a + b + c - 2d$. Also, it is easy to show that the determinant of the Jacobian of the transformation from the zero frequency estimators to $(\overline{x}, \overline{y}, \overline{z}, f_{000})$ is given by $|G| = 2e^{-(a+b+c-2d)}$.

Calculations of the asymptotic efficiency have been carried out for the parameter values given in Section 2. A selection of the results for the symmetric case $a = b = c$ is given in Figure 2. The method is highly efficient for small values of a, b, c and moderate values of d relative to $\min(a, b, c)$. As can also be seen from Figures 1 and 2, the method is more efficient than the method of moments in a considerable region of the parameter space.

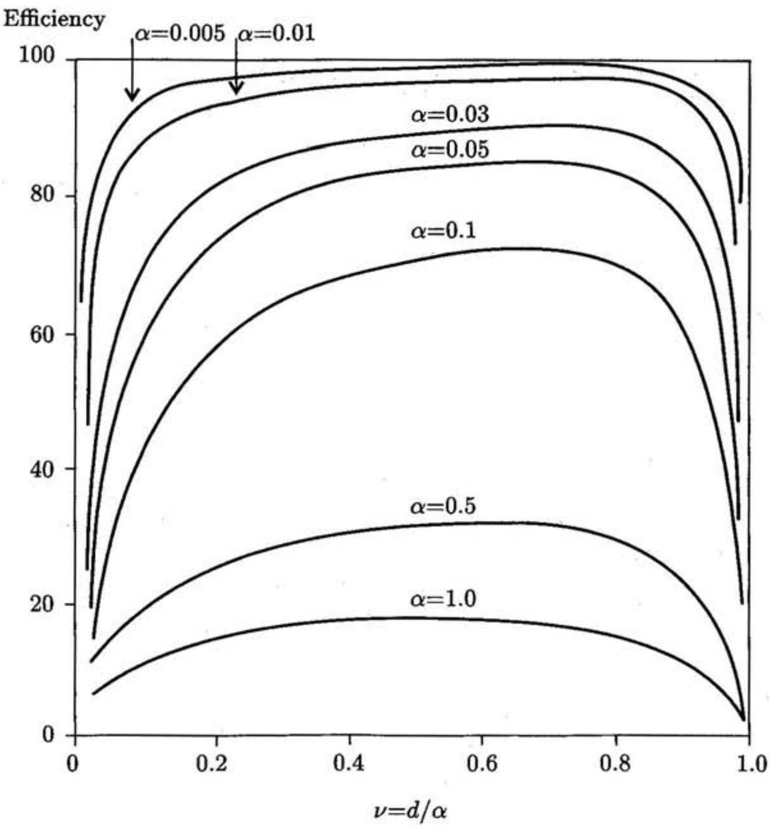

Fig. 2. Efficiency of the zero frequency method

4. The method of even points. This method was introduced by Patel [9] and extended to the bivariate case by Papageorgiou, Kemp and Loukas [8]. In the particular case of the trivariate Poisson distribution, evaluation of the p.g.f. (2) at $(u,v,w) = (1,1,1)$ and $(u,v,w) = (-1,-1,-1)$ yields the relation

$$\begin{aligned} g(1,1,1) + g(-1,-1,-1) &= 1 + \exp\{-2(a+b+c-2d)\} \\ &= 2\Big\{\sum P_{2x,2y,2z} + \sum P_{2x,2y+1,2z+1} \\ &\qquad + \sum P_{2x+1,2y,2z+1} + \sum P_{2x+1,2y+1,2z}\Big\}. \end{aligned}$$

Denote by S the observed proportion at points $(2x,2y,2z)$, $(2x,2y+1,2z+1)$, $(2x+1,2y,2z+1)$ and $(2x+1,2y+1,2z)$. Then, even point estimators can be obtained by using the equations

$$\begin{gathered} \widehat{a} = \overline{x}, \quad \widehat{b} = \overline{y}, \quad \widehat{c} = \overline{z}, \\ \widehat{d} = \tfrac{1}{2}(\overline{x} + \overline{y} + \overline{z}) + \tfrac{1}{4}\log(2S-1), \end{gathered}$$

where $S > 1/2$.

The determinant of the asymptotic covariance matrix for the even points estimates can be obtained by using expression (3). The determinant of the Jacobian of the transformation from the even points estimators to $(\overline{x}, \overline{y}, \overline{z}, S)$ is given by $|G| = 2e^{-2(a+b+c-2d)}$.

The elements of M can be derived by following the approach described in [8]. For n independent (x,y,z) observations the joint moment generating function of nS, $\sum x$, $\sum y$, $\sum z$ is given by

$$h(t_1,t_2,t_3,t_4) = \{\tfrac{1}{2}[(e^{t_1}+1)g(e^{t_2},e^{t_3},e^{t_4}) + (e^{t_1}-1)g(-e^{t_2},-e^{t_3},-e^{t_4})]\}^n,$$

where $g(t_2,t_3,t_4)$ is the p.g.f. (2) of the trivariate Poisson distribution. From this, the covariance matrix of $S, \overline{x}, \overline{y}, \overline{z}$ is easily obtained. For example,

$$\operatorname{Var}(S) = \frac{1}{n}\frac{1-e^{-4B}}{4}, \qquad \operatorname{Cov}(\overline{x}, S) = -\frac{1}{n}ae^{-2B},$$

where $B = a+b+c-2d$. The determinant of M is found to be

$$\begin{aligned} |M| = \frac{1}{n^4}\bigg[& e^{-4B}(a^2d^2 + b^2d^2 + c^2d^2 + 6abcd \\ & - a^2bc - ab^2c - abc^2 - 2abd^2 - 2acd^2 - 2bcd^2) \\ & + \frac{1-e^{-2B}}{4}(abc - ad^2 - bd^2 - cd^2 + 2d^3)\bigg]. \end{aligned}$$

Calculations of the asymptotic efficiency have been carried out for the parameter values given in Section 2. Regions of the parameter space with high efficiency are for small values of the parameters, but the method appears to be less efficient than either the method of moments or the zero frequency method.

5. Examples. Loukas and Papageorgiou [6] used simulated data from two trivariate Poisson distributions in order to apply the method of maximum likelihood. For comparison purposes we are using the same sets of data. For the first set of data, not reproduced here, of size 200 from a trivariate Poisson distribution with parameters $a = b = c = 0.5$ and $d = 0.1$, the maximum likelihood estimates were $\widehat{a} = 0.465$, $\widehat{b} = 0.470$, $\widehat{c} = 0.500$ and $\widehat{d} = 0.0575$ while the moment, zero frequency and even points estimates of d were $\widehat{d} = 0.0807$, $\widehat{d} = 0.1425$ and $\widehat{d} = 0.1657$, respectively. Note that the method of zero frequency is less efficient than the method of moments for this set of parameter values. For the second set of data of size 1000 from a trivariate Poisson distribution with parameters $a = b = c = 0.1$ and $d = 0.05$, the maximum likelihood estimates were $\widehat{a} = 0.087$, $\widehat{b} = 0.089$, $\widehat{c} = 0.089$ and $\widehat{d} = 0.0409$ while the moment, zero frequency and even points estimates of d were $\widehat{d} = 0.0352$, $\widehat{d} = 0.0405$ and $\widehat{d} = 0.04118$, respectively. Note that the method of moments is less efficient than the method of zero frequency for this set of parameter values.

References

[1] N. L. Johnson and S. Kotz, *Distributions in Statistics: Discrete Distributions*, Houghton Mifflin, Boston 1969.

[2] K. Kawamura, *The structure of trivariate Poisson distribution*, Kōdai Math. Sem. Rep. 28 (1976), 1–8.

[3] —, *The structure of multivariate Poisson distribution*, Kōdai Math. J. 2 (1979), 337–345.

[4] A. S. Krishnamoorthy, *Multivariate binomial and Poisson distributions*, Sankhyā 11 (1951), 117–124.

[5] S. Loukas and C. D. Kemp, *On computer sampling from trivariate and multivariate discrete distributions*, J. Statist. Comput. Simul. 17 (1983), 113–123.

[6] S. Loukas and H. Papageorgiou, *On a trivariate Poisson distribution*, Appl. Math. 36 (1991), 432–439.

[7] D. M. Mahamunulu, *A note on regression in the multivariate Poisson distribution*, J. Amer. Statist. Assoc. 62 (1967), 251–258.

[8] H. Papageorgiou, C. D. Kemp and S. Loukas, *Some methods of estimation for the bivariate Hermite distribution*, Biometrika 70 (1983), 479–484.

[9] Y. C. Patel, *Even point estimation and moment estimation in Hermite distribution*, Biometrics 32 (1976), 865–873.

[10] Z. Šidák, *A chain of inequalities for some types of multivariate distributions*, in: Colloq. Math. Soc. János Bolyai, North-Holland, 1972, 693–699.

[11] —, *A chain of inequalities for some types of multivariate distributions, with nine special cases*, Apl. Mat. 18 (1973), 110–118.

[12] Z. Šidák, *On probabilities in certain multivariate distributions: Their dependence on correlations*, ibid. 18 (1973), 128–135.
[13] A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics: Vol. 1, Distribution Theory*, fifth edition, Charles Griffin, London 1987.
[14] K. Subrahmaniam and K. Subrahmaniam, *On the estimation of the parameters in the bivariate negative binomial distribution*, J. Roy. Statist. Soc. Ser. B 35 (1973), 131–146.
[15] H. Teicher, *On the multivariate Poisson distribution*, Scand. Actuar. J. 37 (1954), 1–9.

SOTIRIOS LOUKAS
UNIVERSITY OF IOANNINA
DEPARTMENT OF MATHEMATICS
PROBABILITY, STATISTICS & OPERATIONAL RESEARCH UNIT
45110 IOANNINA, GREECE

Received on 6.6.1991