THE UNIVERSITY OF MANITOBA

BIVARIATE VERSIONS OF THE POISSON-BINOMIAL DISTRIBUTION

by

ASMA ALAVI

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

Department of Statistics
WINNIPEG, MANITOBA

July, 1996

ISBN 0-612-16077-7

ame ______________________

ssertation Abstracts International and Masters Abstracts International are arranged by broad, general subject categories. ease select the one subject which most nearly describes the content of your dissertation or thesis. Enter the corresponding ur-digit code in the spaces provided.

URE SCIENCES (STATISTICS) — SUBJECT TERM

0 4 6 3 — SUBJECT CODE — UMI

ıbject Categories

HE HUMANITIES AND SOCIAL SCIENCES

MMUNICATIONS AND THE ARTS
chitecture 0729
t History 0377
nema 0900
nce 0378
e Arts 0357
ormation Science 0723
urnalism 0391
rary Science 0399
ss Communications 0708
isic 0413
eech Communication 0459
eater 0465

UCATION
neral 0515
ministration 0514
ult and Continuing 0516
ricultural 0517
.... 0273
ngual and Multicultural 0282
siness 0688
mmunity College 0275
rriculum and Instruction 0727
rly Childhood 0518
mentary 0524
ance 0277
idance and Counseling 0519
alth 0680
gher 0745
tory of 0520
me Economics 0278
ustrial 0521
nguage and Literature 0279
thematics 0280
isic 0522
losophy of 0998
ysical 0523
Psychology 0525
Reading 0535
Religious 0527
Sciences 0714
Secondary 0533
Social Sciences 0534
Sociology of 0340
Special 0529
Teacher Training 0530
Technology 0710
Tests and Measurements 0288
Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS
Language
General 0679
Ancient 0289
Linguistics 0290
Modern 0291
Literature
General 0401
Classical 0294
Comparative 0295
Medieval 0297
Modern 0298
African 0316
American 0591
Asian 0305
Canadian (English) 0352
Canadian (French) 0355
English 0593
Germanic 0311
Latin American 0312
Middle Eastern 0315
Romance 0313
Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY
Philosophy 0422
Religion
General 0318
Biblical Studies 0321
Clergy 0319
History of 0320
Philosophy of 0322
Theology 0469

SOCIAL SCIENCES
American Studies 0323
Anthropology
Archaeology 0324
Cultural 0326
Physical 0327
Business Administration
General 0310
Accounting 0272
Banking 0770
Management 0454
Marketing 0338
Canadian Studies 0385
Economics
General 0501
Agricultural 0503
Commerce-Business 0505
Finance 0508
History 0509
Labor 0510
Theory 0511
Folklore 0358
Geography 0366
Gerontology 0351
History
General 0578
Ancient 0579
Medieval 0581
Modern 0582
Black 0328
African 0331
Asia, Australia and Oceania 0332
Canadian 0334
European 0335
Latin American 0336
Middle Eastern 0333
United States 0337
History of Science 0585
Law 0398
Political Science
General 0615
International Law and Relations 0616
Public Administration 0617
Recreation 0814
Social Work 0452
Sociology
General 0626
Criminology and Penology 0627
Demography 0938
Ethnic and Racial Studies 0631
Individual and Family Studies 0628
Industrial and Labor Relations 0629
Public and Social Welfare 0630
Social Structure and Development 0700
Theory and Methods 0344
Transportation 0709
Urban and Regional Planning 0999
Women's Studies 0453

HE SCIENCES AND ENGINEERING

OLOGICAL SCIENCES
riculture
General 0473
Agronomy 0285
Animal Culture and Nutrition 0475
Animal Pathology 0476
Food Science and Technology 0359
Forestry and Wildlife 0478
Plant Culture 0479
Plant Pathology 0480
Plant Physiology 0817
Range Management 0777
Wood Technology 0746
logy
General 0306
Anatomy 0287
Biostatistics 0308
Botany 0309
Cell 0379
Ecology 0329
Entomology 0353
Genetics 0369
Limnology 0793
Microbiology 0410
Molecular 0307
Neuroscience 0317
Oceanography 0416
Physiology 0433
Radiation 0821
Veterinary Science 0778
Zoology 0472
physics
General 0786
Medical 0760

RTH SCIENCES
geochemistry 0425
ochemistry 0996
Geodesy 0370
Geology 0372
Geophysics 0373
Hydrology 0388
Mineralogy 0411
Paleobotany 0345
Paleoecology 0426
Paleontology 0418
Paleozoology 0985
Palynology 0427
Physical Geography 0368
Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES
Environmental Sciences 0768
Health Sciences
General 0566
Audiology 0300
Chemotherapy 0992
Dentistry 0567
Education 0350
Hospital Management 0769
Human Development 0758
Immunology 0982
Medicine and Surgery 0564
Mental Health 0347
Nursing 0569
Nutrition 0570
Obstetrics and Gynecology 0380
Occupational Health and Therapy 0354
Ophthalmology 0381
Pathology 0571
Pharmacology 0419
Pharmacy 0572
Physical Therapy 0382
Public Health 0573
Radiology 0574
Recreation 0575
Speech Pathology 0460
Toxicology 0383
Home Economics 0386

PHYSICAL SCIENCES
Pure Sciences
Chemistry
General 0485
Agricultural 0749
Analytical 0486
Biochemistry 0487
Inorganic 0488
Nuclear 0738
Organic 0490
Pharmaceutical 0491
Physical 0494
Polymer 0495
Radiation 0754
Mathematics 0405
Physics
General 0605
Acoustics 0986
Astronomy and Astrophysics 0606
Atmospheric Science 0608
Atomic 0748
Electronics and Electricity 0607
Elementary Particles and High Energy 0798
Fluid and Plasma 0759
Molecular 0609
Nuclear 0610
Optics 0752
Radiation 0756
Solid State 0611
Statistics 0463

Applied Sciences
Applied Mechanics 0346
Computer Science 0984
Engineering
General 0537
Aerospace 0538
Agricultural 0539
Automotive 0540
Biomedical 0541
Chemical 0542
Civil 0543
Electronics and Electrical 0544
Heat and Thermodynamics 0348
Hydraulic 0545
Industrial 0546
Marine 0547
Materials Science 0794
Mechanical 0548
Metallurgy 0743
Mining 0551
Nuclear 0552
Packaging 0549
Petroleum 0765
Sanitary and Municipal 0554
System Science 0790
Geotechnology 0428
Operations Research 0796
Plastics Technology 0795
Textile Technology 0994

PSYCHOLOGY
General 0621
Behavioral 0384
Clinical 0622
Developmental 0620
Experimental 0623
Industrial 0624
Personality 0625
Physiological 0989
Psychobiology 0349
Psychometrics 0632
Social 0451

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

COPYRIGHT PERMISSION

BIVARIATE VERSIONS OF THE POISSON-BINOMIAL DISTRIBUTION

BY

ASMA ALAVI

A Thesis/Practicum submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

TABLE OF CONTENTS

ABSTRACT

Compounding and generalizing are the two well known techniques of forming more complex probability distributions starting with elementary distributions. In the univariate case extensive literature is available discussing these techniques and their applications. Application of these procedures to the bivariate discrete distributions needs some consideration, because not much work has been done in this area. The resultant distributions in the bivariate case turn out to be more complex than in the univariate case. In number of instances entirely new distributions are generated.

In this thesis two types of the bivariate Poisson-binomial distributions, namely, bivariate Poisson distribution and Poisson-double binomial distribution are discussed. Based on the probability generating functions of the these bivariate distributions, we find various properties of the distributions including, joint probability function, marginal and conditional distributions, various types of moments, and limiting distributions.

Three methods of estimation, namely, method of moments, zero-zero cell frequency method and method of maximum likelihood, of the parameters involved are proposed and studied in detail. Due to the complexity of the probability function, iteration procedures are suggested. The asymptotic variances of the estimators along with their efficiencies are examined. It is found that the zero-zero cell frequency method is more efficient as compared to method of moments.

Having decided on the type of the estimates to be used, goodness-of-fit tests for model building are done using simulated as well as the real life data of accidents sustained by bus drivers. A probability distribution is preferable, if along with the theoretical development, it has wide-spread application in real life. For the two distributions under consideration it is found that indeed they do have real life applications.

ACKNOWLEDGMENTS

I would like to express my heartiest thanks to my supervisor Professor Subrahmaniam Kocherlakota and to my advisor Professor Kathleen Kocherlakota for their all around help, advice and encouragement. Because of their constant educational, moral and financial support, I am able to complete the requirements for the master's degree along with my household responsibilities. I have no doubt in saying that, my first ever university experience abroad is remarkable just because of them.

Finally, I thank my husband, my parents, my daughter, my sisters and brothers for their moral support to me.

0

INTRODUCTION AND SUMMARY

Discrete distributions are an essential part of statistical theory and applications. Elementary univariate discrete distributions such as binomial and Poisson are well-known. More complex types of distributions can be obtained from the simple distributions in a number of ways. One way is to consider random sums of independent random variables. This procedure is referred to as a random sum process or generalization. The other way is compounding, often referred to as mixing. In compounding the parameter of a probability distribution is taken to be a random variable rather than as a fixed value. A detailed discussion of these two techniques is given in Chatfield *et al.* (1973). Douglas (1970) and Feller (1943, 1968) also describe these models in detail.

Consider the sum $S_N = X_1 + \ldots + X_N$, where X's are independently identically distributed random variables with probability generating function f(t) and N is a random variable with probability generating function g(t), then the probability generating function of S_N is $g(f(t))$.

The following examples illustrate this procedure: (i) In animal trapping experiments the size of a species is a random variable with the probability function g_n. Let X represents the outcome of a trapping experiment with X = 1 if an animal is trapped and zero otherwise. Then under the assumption that $P\{X = 1\} = p$ and of independence, the probability generating function of the number S_N of animals trapped is $g(q + pt)$, where q = 1 - p. (ii) Let g_n be the probability of an insect laying n eggs. Under different viability of eggs and independence, when the probability of survival of an egg is p, the probability generating function of the number of eggs surviving in an egg mass of size n is $(q+pt)^n$. If the number of egg-masses under discussion has the Poisson distribution with parameter λ, then the probability generating function of the total number of eggs surviving is $\exp[\lambda\{(q+pt)-1\}]$. This distribution is called the Poisson-binomial distribution. A further generalization of this model yields the probability generating function as $\exp[\lambda\{(q+pt)^k-1\}]$. This model is discussed in detail in chapter 1.

These complex models in univariate case, where only one characteristic of the population is taken into account, have been discussed extensively in the literature. In the bivariate case, where two characteristics of the population are considered, there is much room for new work. In this thesis an effort is made to produce some new distributions, using the bivariate distributions available in literature. Bivariate (joint) binomial and double (independent) binomial distributions are very familiar. Two extensions of the Poisson-binomial to the bivariate case are possible. In one case the index parameter of a bivariate binomial distribution is taken to be proportional to the Poisson random variable. In the second case we consider a double binomial distribution, with index parameters each being proportional to a Poisson random variable. Both models are discussed in this thesis.

A brief review of the Poisson-binomial distribution (univariate case) is given in chapter 1. The chapter includes the basic model for the Poisson-binomial distribution, properties of the distribution along with the estimation of the parameters. The efficiencies of the estimators are also given.

Chapter 2 is the main building block of the whole thesis. In this chapter it is shown how new distributions arise, when the compounding process is applied to different distributions. Basically bivariate binomial and double binomial distributions are taken as describing the nature of the population. When the index parameters are taken as random variables (in our case having the Poisson distribution), completely different types of distributions are obtained. Three types of possibilities are discussed in this chapter, which give rise to two distributions. One is the familiar bivariate Poisson distribution with a new set of parameters and the other is the new distribution which is named as the Poisson-double binomial distribution. Various properties of the two distributions including probability generating function, probability function, marginal and conditional distributions and different types of moments are given. Using the probability generating function, difference equations are developed for the probability function. In addition to the usual properties of the distribution, the limiting forms of the distribution are also discussed.

Estimation of the parameters of the two distributions developed in chapter 2 is the object of chapter 3. One section is devoted to obtain estimators and their asymptotic variances and covariances for the bivariate Poisson distribution. This work is unique in the sense that it uses the set of parameters that is different from the set used already in literature. The second section consists of estimation in case of Poisson-double binomial distribution. Three methods of estimation, namely, method of moments, zero-zero cell frequency method and method of maximum likelihood are employed to get estimators and their variances and covariances. In the case of maximum likelihood estimation difference equations are needed for determining the estimators. These equations are developed. Tables and graphs of efficiencies are included.

Chapter 4 deals with the practical application of the two distributions with greater emphasis on the Poisson-double binomial distribution. The estimation procedures are exemplified by the use of simulated and real-life data. Tables of the estimates along with their estimated standard errors are given. The adequacy of the model is tested using two procedures for the goodness-of-fit.

Computational aspect of the problems studied are discussed in appendix. The computer package S-Plus was used to write all of required algorithms. These S-Plus programs cover simulation of the Poisson-double binomial distribution, probability computations, numerical evaluation of the information matrix, iteration process to get maximum likelihood estimates, computation of variance-covariance matrices of method of moments and method of zero-zero cell frequency estimators, computation of Z-statistic for the probability generating function technique used in the goodness-of-fit.

1

UNIVARIATE POISSON-BINOMIAL DISTRIBUTION

1.1 Genesis

Two well known techniques of constructing discrete distributions are generalizing and compounding. The generalized distribution is also called a stopped sum and a compound is also referred to as a mixture. The univariate Poisson-binomial distribution can be obtained in either of these two ways. These techniques are very different, but in certain situations they give rise to the same distribution as in the case of univariate Poisson-binomial distribution. Gurland (1957) describes the relationship between these models. In Gurland's notation, let X_1 be a random variable with probability generating function (pgf)

$$G_1(t) = [h(t)]^{\theta}, \tag{1.1.1}$$

where θ is a given parameter. Let θ be regarded as a random variable X_2 with pgf G_2. Then whatever be X_2

$$X_1 \wedge X_2 \sim X_2 \vee X_1, \tag{1.1.2}$$

where $\wedge$ stands for mixing and $\vee$ stands for generalizing.

In the stopped sum procedure two independent distributions are combined in a particular way. The principal model underlying these distributions can be thought as the sum of the observations from the distribution ψ_2, where the number of the observations in the given summation is a random variable having the distribution ψ_1. Summation of the observations from the distribution ψ_2 is randomly stopped by the value of the observation having ψ_1 distribution. Consider a sequence of independent random variables

$Y_1, Y_2, \ldots, Y_N$ with pgf $G_2(t)$ and size of the sequence N is random variable with pgf $G_1(t)$, then the sum

$$S_N = Y_1 + Y_2 + \ldots + Y_N$$

is a random variable. The pgf of S_N is given by

$$\begin{aligned} E\left[t^{S_N}\right] &= E_N\left[E\left[t^{S_N} | N\right]\right] \\ &= E_N\left[G_2(t)\right] \\ &= G_1\left(G_2(t)\right), \end{aligned} \tag{1.1.3}$$

alternatively,

$$S_N \sim \psi_1 \vee \psi_2.$$

In the particular case of Poisson-binomial distribution

$$\text{Poisson}(\lambda) \vee \text{binomial}(n, p).$$

For example consider n identical coins, with the probability of a head at a single toss being p. These coins are tossed simultaneously N number of times, where N is a Poisson random variable having mean λ, with pgf $G_1(t)$. The pgf $G_2(t)$ of the numbers of heads for each of N trials is $(q + pt)^n$ then the pgf for the total numbers of heads will be

$$G_1\left(G_2(t)\right) = \exp\left[-\lambda + \lambda(q + pt)^n\right], \tag{1.1.4}$$

where q = 1 - p. This distribution is also called Poisson-stopped binomial.

On the other hand the mixture model is defined as follows. Consider the binomial distribution with the parameters nk and p. If we let k vary, while n and p are constant, then depending on the probability distribution of k, the resulting distribution of the observed number of heads will have a certain form. Let the distribution of k be Poisson with parameter λ. Then the resulting distribution of the number of heads will have the pgf defined in (1.1.4). From here it can be seen easily that the Poisson-binomial distribution is a superimposition of binomial distributions with different index parameters, each of which has a Poisson distribution. In other words the mixed distribution is given by

$$\text{binomial}(nk, p) \underset{k}{\wedge} \text{Poisson}(\lambda)$$

or

mixed-binomial(kn,p) on k by Poisson(λ).

The probability generating function (pgf) is given by,

$$\exp\left[\lambda\left\{(q+pt)^n - 1\right\}\right]. \tag{1.1.5}$$

It can be seen that (1.1.4) and (1.1.5) are identical.

The distribution in (1.1.5) has three parameters, namely λ, n and p, which will be represented by Poisson-binomial (λ, n, p). The parameter λ is a positive number, n is positive integer and p lies between 0 and 1. As given in Douglas (1980) and Johnson *et al.* (1992), when n = 2, Poisson-binomial distribution is equivalent to the Hermite distribution.

1.2 Properties

For n = 1 the pgf in (1.1.5) reduces to that of the Poisson with parameter $p\lambda$. The distribution has a reproductive property which means that if $X_1, X_2, \ldots, X_m$ are independently distributed Poisson-binomial (λ_i, n, p), then their sum would also be a Poisson-binomial with parameters $\left(\sum_{i=1}^{m} \lambda_i, n, p\right)$.

Probability function

The probability function (pf) is obtained by expanding the pgf as

$$G(t) = \exp\left[-\lambda + \lambda(q+pt)^n\right]$$

$$= \sum_{k=0}^{\infty} (q+pt)^{kn} \exp(-\lambda) \frac{\lambda^k}{k!}.$$

As in Douglas (1980, p. 260), writing $\mu = \lambda q^n$ and $\rho = \frac{p}{q}$, the pgf will become

$$G(t) = \exp(-\lambda)\sum_{k=0}^{\infty}\frac{\mu^k}{k!}(1+\rho t)^{nk}$$

$$= \exp(-\lambda)\sum_{k}\frac{\mu^k}{k!}\sum_{x=0}^{\infty}\binom{nk}{x}\rho^x t^x$$

where the upper limit of summation for x is nk. Using the fact that binomial coefficient is zero for x greater than nk, we can write the unrestricted summation. Now interchanging the order of summation, we get

$$G(t) = \exp(-\lambda)\sum_{x=0}^{\infty}\left(\sum_{k=0}^{\infty}\binom{nk}{x}\frac{\mu^k}{k!}\right)\rho^x t^x,$$

where now k should be such that $k \geq \frac{x}{n}$. The coefficient of t^x is

$$P(X=x) = \exp(-\lambda)\left(\frac{p}{q}\right)^x \sum_{k\geq\frac{x}{n}}^{\infty}\binom{nk}{x}\frac{\left(\lambda q^n\right)^k}{k!}.$$

A recurrence relation among the probabilities, which is quite useful for numerical evaluation, is also given by Douglas (1980). The form of the relationship is

$$P(X=x+1) = \frac{\lambda np}{x+1}\sum_{s=0}^{x}\frac{(n-1)!}{s!(n-1-s)!}q^{n-1-s}p^s P(X=x-s) \qquad (1.2.1)$$

for x = 0, 1, 2,..., and

$$P(X=0) = \exp\left(-\lambda+\lambda q^n\right). \qquad (1.2.2)$$

Moments and cumulants

According to Johnson *et al.* (1992, p. 365), the cumulants of the Poisson-binomial distribution can be obtained from the raw moments of the binomial distribution. The simplest of all are the factorial cumulants. The factorial cumulant generating function (fcgf) is

$$\ln G(t+1) = \lambda(1+pt)^n - \lambda, \qquad (1.2.3)$$

which gives rise to factorial cumulants in the following form,

$$\kappa_{[r]} = \frac{\lambda p^r n!}{(n-r)!}. \qquad (1.2.4)$$

From this the cumulants can be found easily, using the relationship between factorial cumulants and cumulants. An alternative way of finding cumulants is by using the recurrence relation

$$\kappa_{r+1} = p(1-p)\frac{\partial \kappa_r}{\partial p} + np\kappa_r, \qquad (1.2.5)$$

where $\kappa_1 = \lambda np$. Using these cumulants the first four moments of the distribution are

$$\mu = \lambda np$$

$$\mu_2 = \lambda n^2p^2 + \lambda npq$$

$$\mu_3 = \lambda n^3p^3 + 3\lambda n^2p^2q + \lambda npq(1-2p)$$

$$\mu_4 = \lambda n^4p^4 + 6\lambda n^3p^3q + \lambda n^2p^2q(7-11p) + \lambda npq(1-6pq) + 3\lambda^2(n^2p^2 + npq)^2,$$

where q = 1 - p.

1.3 Limits and approximations of fcgf

Douglas (1980, p. 268) discusses the limiting and approximate forms of the Poisson-binomial distribution. Two cases are of main interest:

(1) $p \to 0$: Suppose λ and n are fixed, then the fcgf in (1.2.3) will become

$$-\lambda + \lambda\left(1 + npt + O(p^2)\right)$$
$$= \lambda npt + O(p^2),$$

where $O(p^2)$ means the term of order zero. The above fcgf corresponds to a Poisson(λnp) distribution. Also if $n \to \infty$ in such a way that np is finite, then fcgf will become

$$-\lambda + \lambda\left(1 + npt\frac{1}{n}\right)^n \to -\lambda + \lambda\exp(npt),$$

which is fcgf of the Neyman Type $A(\lambda, np)$ distribution.

(2) $p \to 1$: The fcgf in (1.2.3) then becomes

$$-\lambda + \lambda(1+t)^n,$$

which is the fcgf of the scaled Poisson(λ) variate with Poisson(λ) probabilities at 0, n, 2n, 3n,... instead of 0, 1, 2, 3,....

1.4 Estimation

The restriction that n should be a positive integer makes the estimation process a bit difficult. However an alternative way of estimating parameters is to fix n and to estimate λ and p. The process is repeated for some specified values of n and a choice is made of that set of values for n, p and λ which gives the best fit based on a goodness-of-fit test.

Douglas (1980, chap. 5) discusses the simultaneous estimation of three parameters and emphasizes that not all n values are possible for certain specified forms of this distribution. Also no two dimensional sufficient estimators exist for this distribution.

Three techniques of estimation, namely the maximum likelihood, the method of moments and zero cell frequency method are discussed by Douglas (1980) and Johnson *et al.* (1992). The efficiencies of method of moments and zero-zero cell frequency method as compared to method of maximum likelihood are also computed.

Method of maximum likelihood

Given a random sample of N observations $x_1, x_2, \ldots, x_N$ from the Poisson-binomial distribution, n being known, the maximum likelihood equations are given in the following form in Sprott (1958)

$$\left.\begin{aligned} \bar{x} = \sum_{j=1}^{N} \frac{x_j}{N} = n\hat{\lambda}\hat{p} \\ \sum_{j=1}^{N} \left(x_j + 1\right) \frac{\hat{p}_{x_{j+1}}}{\hat{p}_{x_j}} = Nn\hat{\lambda}\hat{p} \end{aligned}\right\}, \tag{1.4.1}$$

where $\hat{p}_{x_j}$ is $P(X = x_j)$ with λ and p replaced by $\hat{\lambda}$ and $\hat{p}$. These equations in (1.4.1) have to be solved iteratively. In this respect it is important to start with good initial estimates. Usually the initial estimates are obtained from method of moments or zero-zero cell frequency method.

To get the variances and covariance of maximum likelihood estimators, the information matrix I is used. From the results given by Sprott (1958),

$$\sum xP(x) = \lambda np$$

and

$$\sum x^2 P(x) = \lambda np[1 + (n-1)p] + \lambda^2 n^2 p^2.$$

If N is the sample size, then

$$\frac{I_{\lambda\lambda}}{N} = \sum \frac{1}{P(x)}\left[\frac{\partial P(x)}{\partial \lambda}\right]^2$$
$$= q^2 A + \frac{np + (n-1)pq}{n\lambda},$$

$$\frac{I_{pp}}{N} = \sum \frac{1}{P(x)}\left[\frac{\partial P(x)}{\partial p}\right]^2$$
$$= n^2\lambda^2 A + n\lambda\left(1 - n + \frac{1}{p}\right),$$

and

$$\frac{I_{\lambda p}}{N} = \sum \frac{1}{P(x)}\left[\frac{\partial P(x)}{\partial \lambda}\right]\left[\frac{\partial P(x)}{\partial p}\right]$$
$$= -n\lambda qA + nq + p,$$

where

$$A = -1 + \sum F^2(x)P(x)$$

and

$$F(x) = \frac{(x+1)P(x+1)}{n\hat{\lambda}\hat{p}P(x)}.$$

The determinant of the matrix $\left(\frac{I}{N}\right)$ denoted by $D_{\lambda,p}$ is

$$D_{\lambda,p} = n\lambda\left(n + \frac{p}{q}\right)A - (n-1)^2.$$

Hence the first order asymptotic variances and covariance of the maximum likelihood estimators are

$$\begin{aligned}\mathrm{Var}\left(\hat{\lambda}\right) &= \sigma^2_{\hat{\lambda}} \\ &= \frac{I_{pp}}{N^2 D_{\lambda,p}},\end{aligned}$$

$$\begin{aligned}\mathrm{Var}\left(\hat{p}\right) &= \sigma^2_{\hat{p}} \\ &= \frac{I_{\lambda\lambda}}{N^2 D_{\lambda,p}},\end{aligned}$$

and

$$\begin{aligned}\mathrm{Cov}\left(\hat{\lambda},\hat{p}\right) &= \sigma_{\hat{\lambda}\hat{p}} \\ &= \frac{-I_{\lambda p}}{N^2 D_{\lambda,p}}.\end{aligned}$$

Method of moments

For a given value of n the moment equations for λ and p are given by

$$\left.\begin{aligned}\tilde{p} &= \frac{s^2 - \overline{x}}{(n-1)\overline{x}} \\ \tilde{\lambda} &= \frac{\overline{x}}{n\tilde{p}}\end{aligned}\right\}, \qquad (1.4.3)$$

where $s^2 = \sum_{j=1}^{N} \frac{\left(x_j - \overline{x}\right)^2}{N-1}$.

Solving the two equations in (1.4.3), the method of moments estimators of λ and p are

$$\left.\begin{aligned}\tilde{\lambda} &= \left(\frac{n-1}{n}\right)\frac{\bar{x}^2}{s^2-\bar{x}} \\ \tilde{p} &= \left(\frac{n}{n-1}\right)\frac{s^2-\bar{x}}{\bar{x}^2}\end{aligned}\right\}. \tag{1.4.4}$$

The first order generalized variance, i.e. the determinant of the variance covariance matrix of the method of moments estimators is given by Sprott (1958) as

$$\frac{1}{n^4(n-1)^2\lambda^2p^4N^2}\left\{n^2p^2\lambda^2\left[1+7(n-1)(n-2)p^2+(n-1)(n-2)(n-3)p^3\right]\right.$$
$$\left.\left[1+(n-1)p\right]+2n^3\lambda^3p^3\left[1+(n-1)p\right]^3-n^2\lambda^2p^2\left[1+3(n-1)p+(n-1)(n-2)p^2\right]^2\right\}.$$

Zero cell frequency method

In this method the sample mean and the relative frequency in zero cell are used to estimate the parameters. Two estimating equations are

$$\left.\begin{aligned}\bar{x} &= n\lambda'p' \\ \frac{f_0}{N} &= \exp\left[-\lambda'\left(1-(1-p')^n\right)\right]\end{aligned}\right\}. \tag{1.4.5}$$

Solving for p',

$$\frac{\bar{x}}{\ln\left(\frac{f_0}{N}\right)} = \frac{np'}{(1-p')^n-1} \tag{1.4.6}$$

and

$$\lambda' = \frac{\bar{x}}{np'}, \tag{1.4.7}$$

where f_0 is the relative frequency in zero cell from the observed data. Equation (1.4.6) can be solved iteratively.

Variances and covariance of the estimators from the zero cell frequency method are given in Douglas (1980, p. 285). Applying the Taylor series expansion method, the first order variances and covariance are

$$\mathrm{Var}(p') = \frac{p\left[np + P(0)\left\{\lambda\left(1-q^n\right)\left(1+(n-1)p\right) - 2\lambda np\left(1-q^n\right) - np\right\}\right]}{N\lambda^2 nP(0)\left(1-q^n-npq^{n-1}\right)^2},$$

$$\mathrm{Var}(\lambda') = \frac{1}{N}\left\{\frac{\lambda\left(1+(n-1)p\right)}{np} - \frac{2\lambda}{np}\cdot\frac{\left(1+(n-1)p\right)\left(1-q^n\right)-np}{1-q^n-npq^{n-1}} + \frac{\lambda^2 N}{p^2}\mathrm{Var}(p')\right\},$$

and

$$\mathrm{Cov}(\lambda', p') = \frac{1}{N}\left\{\frac{1}{n}\cdot\frac{\left(1+(n-1)p\right)\left(1-q^n\right)-np}{1-q^n-npq^{n-1}} - \frac{\lambda N}{p}\mathrm{Var}(p')\right\}.$$

Efficiency

The performance of an estimator is evaluated on the basis of its efficiency as compared to that of maximum likelihood estimator. Efficiency can be measured in two ways: first is by comparing individual variances and covariances and second by comparing generalized variances. Considering the three methods described earlier, Katti and Gurland (1962) found that method of zero cell frequency is much more efficient than the method of moments. Following are the tables of efficiencies (ratio of generalized variances) for the two method of estimations. These tables are adapted from Johnson *et al.* (1992).

Table 1.4.1 Efficiency of the Method of Moments for the Poisson-Binomial Distribution

		λ				
n	p	0.1	0.3	0.5	1.0	2.0
2	0.1	0.928	0.865	0.843	0.840	0.870
2	0.3	0.732	0.569	0.525	0.533	0.635
2	0.5	0.494	0.307	0.307	0.267	0.392
3	0.1	0.896	0.823	0.793	0.779	0.810
3	0.3	0.658	0.501	0.452	0.446	0.542
3	0.5	0.426	0.268	0.231	0.231	0.333
5	0.1	0.816	0.726	0.688	0.671	0.715
5	0.3	0.527	0.379	0.337	0.332	0.435
5	0.5	0.345	0.210	0.178	0.176	0.277

Table 1.4.2 Efficiency of the Zero Cell Frequency Method for the Poisson-binomial Distribution

		λ				
n	p	0.1	0.3	0.5	1.0	2.0
2	0.1	0.984	0.974	0.977	0.991	0.947
2	0.3	0.937	0.888	0.883	0.923	0.981
2	0.5	0.862	0.740	0.700	0.717	0.851
3	0.1	0.994	0.986	0.984	0.944	0.994
3	0.3	0.968	0.930	0.918	0.935	0.974
3	0.5	0.896	0.798	0.763	0.765	0.850
5	0.1	0.995	0.987	0.985	0.993	0.989
5	0.3	0.969	0.924	0.905	0.911	0.950
5	0.5	0.889	0.769	0.716	0.690	0.793

Figure 1.4.1 Efficiency Plot for Method of Moments

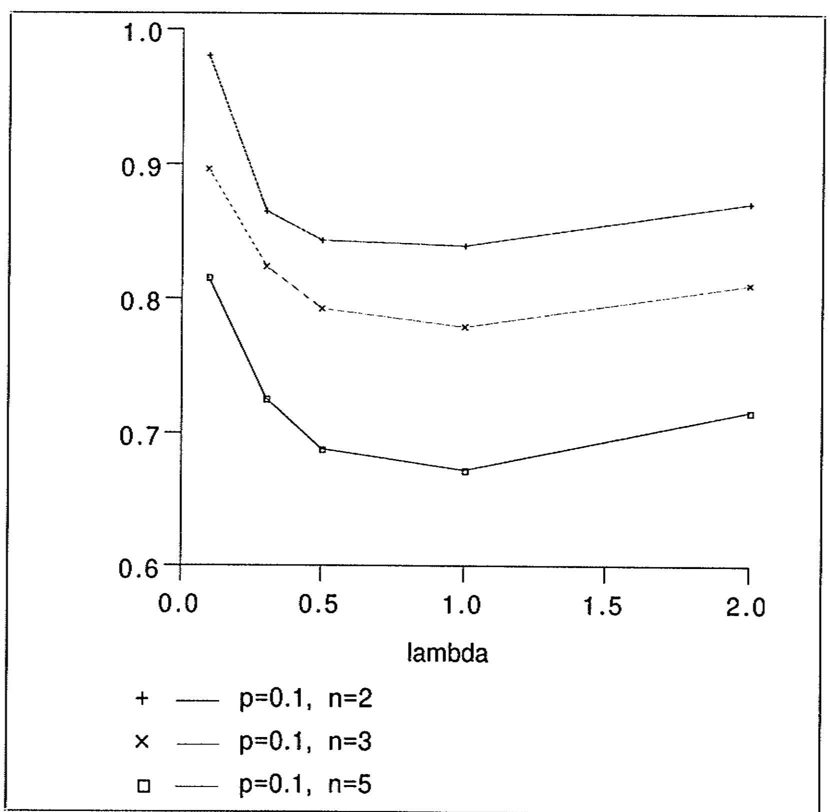
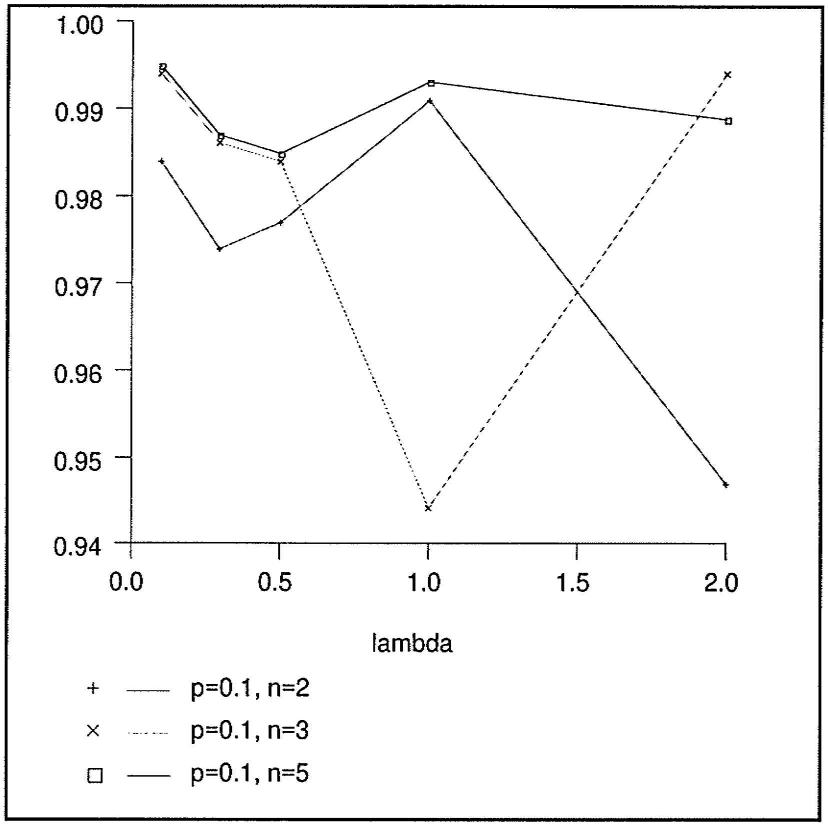


Figure 1.4.2 Efficiency Plot for Method of Zero Cell Frequency



In general it can be seen that efficiency of method of moments and zero cell frequency method is high throughout the region $2 \leq n \leq 5$, $0 \leq \lambda \leq 2$, $0 \leq p \leq 0.5$. Also for $p \leq 0.3$ the efficiency is 90% or more. If λ is not too large and $p > 0.3$, then maximum likelihood estimation should be used. For any given λ, efficiency tends to zero as p approaches 1.

2

BIVARIATE VERSIONS OF POISSON-BINOMIAL DISTRIBUTION

2.1 Introduction

The univariate version of the Poisson-binomial distribution has been discussed in the first chapter. In the bivariate case a number of distributions can be obtained by mixing or generalizing binomial and Poisson distributions. Basically three different cases are considered in this chapter.

Case 1: bivariate binomial distribution with index parameter n having the Poisson distribution.

Case 2: double binomial distribution with index parameters $n_1 = n_2 = n$.having the same Poisson distribution.

Case 3: double binomial distribution with index parameters n_1k and n_2k with $n_1 \neq n_2$ and k having the Poisson distribution.

2.2 Case 1

Let X_1 and X_2 be two random variables having bivariate binomial distribution with the probability generating function (pgf)

$$\pi_2(t_1,t_2) = (q_1q_2 + p_1q_2t_1 + p_2q_1t_2 + p_1p_2t_1t_2)^n. \tag{2.2.1}$$

The distribution has three parameters namely p_1, p_2 and n. Here $q_1 = 1 - p_1$ and $q_2 = 1 - p_2$. If n is not a constant, rather a random variable having Poisson distribution with parameter λ, and the pgf

$$\pi_1(t) = \exp[\lambda(t-1)], \tag{2.2.2}$$

then the pgf of the generalized distribution will be

$$\pi(t_1,t_2)=\pi_1(\pi_2(t_1,t_2))$$

$$=\exp\left[\lambda\left(q_1q_2+p_1q_2t_1+p_2q_1t_2+p_1p_2t_1t_2\right)-\lambda\right]. \quad (2.2.3)$$

The pgf in (2.2.3) can be written as

$$\pi(t_1,t_2)=\exp\left[\lambda\left((q_1q_2-1)+p_1q_2t_1+p_2q_1t_2+p_1p_2t_1t_2\right)\right], \quad (2.2.4)$$

which is equivalent to the pgf of the bivariate Poisson distribution, with the pgf

$$\pi^*(t_1,t_2)=\exp\left[\lambda_1(t_1-1)+\lambda_2(t_2-1)+\lambda_3(t_1t_2-1)\right], \quad (2.2.5)$$

where λ_1, λ_2 and λ_3 are the three parameters of the bivariate Poisson distribution. Comparing the pgfs in (2.2.4) and (2.2.5), we get the relationship among the parameters as

$$\left.\begin{aligned}\lambda_1&=p_1q_2\lambda\\ \lambda_2&=p_2q_1\lambda\\ \lambda_3&=p_1p_2\lambda\end{aligned}\right\}. \quad (2.2.6)$$

Marginal distributions

The marginal generating functions can be found from the joint pgf. For X_1 set $t_2=1$ and $t_1=t$ in the joint pgf in (2.2.4). Hence the marginal pgf for X_1 is

$$g_1(t)=\exp\left[\lambda\left((q_1q_2-1)+p_1q_2t+p_2q_1+p_1p_2t\right)\right]$$

$$=\exp[\lambda(q_1q_2-1)+\lambda p_2q_1+\lambda(p_1q_2+p_1p_2)t],$$

since $q_1q_2+p_2q_1-1=-(p_1q_2+p_1p_2)$,

$$g_1(t)=\exp\left[\lambda(p_1q_2+p_1p_2)t-\lambda(p_1q_2+p_1p_2)\right]$$

$$=\exp\left[\lambda p_1(t-1)\right],$$

which is the pgf of the Poisson variate with parameter λp_1. Similarly the marginal pgf of X_2 is

$$g_2(t)=\exp\left[\lambda p_2(t-1)\right],$$

which is Poisson pgf with parameter λp_2.

Probability function

Using the probability function for the bivariate Poisson distribution given in Kocherlakota and Kocherlakota (1992, p. 92), and reparameterizing in terms of p_1, p_2 and λ the probability function is

$$f(r,s) = e^{(q_1 q_2 - 1)\lambda} \sum_{i=0}^{\min(r,s)} \frac{\lambda^{r+s-i} (p_1 q_2)^{r-i} (p_2 q_1)^{s-i} (p_1 p_2)^i}{(r-i)!(s-i)!i!}$$

or

$$f(r,s) = e^{(q_1 q_2 - 1)\lambda} \sum_{i=0}^{\min(r,s)} \frac{p_1^r p_2^s q_1^{s-i} q_2^{r-i}}{(r-i)!(s-i)!i!} \lambda^{r+s-i}. \tag{2.2.7}$$

Factorial moments and cumulants

The factorial moment generating function (fmgf) is given by

$$G(t_1, t_2) = \exp[\lambda p_1 t_1 + \lambda p_2 t_2 + \lambda p_1 p_2 t_1 t_2] \tag{2.2.8}$$

The (r,s)th factorial moment is the coefficient of $\frac{t_1^r}{r!} \frac{t_2^s}{s!}$ in the expansion of fmgf in (2.2.8); that is,

$$\mu_{[r,s]} = p_1^r p_2^s \sum_{i=0}^{\min(r,s)} \binom{r}{i} \binom{s}{i} \lambda^{r+s-i} i!. \tag{2.2.9}$$

In particular

$$\mu_{[1,0]} = p_1 \lambda$$

$$\mu_{[0,1]} = p_2 \lambda$$

$$\mu_{[1,1]} = p_1 p_2 (\lambda + \lambda^2),$$

$$\mu_{[2,0]} = p_1^2 \lambda^2$$

$$\mu_{[0,2]} = \lambda^2 p_2^2$$

$$\mu_{[1,2]} = p_1 p_2^2 (\lambda^3 + 2\lambda^2)$$

$$\mu_{[2,1]} = p_1^2 p_2\left(\lambda^3 + 2\lambda^2\right)$$

$$\mu_{[2,2]} = p_1^2 p_2^2\left(\lambda^4 + 4\lambda^3 + 2\lambda^2\right).$$

Some of the higher order raw moments are

$$\mu'_{2,0} = p_1^2\lambda^2 + p_1\lambda$$

$$\mu'_{0,2} = p_2^2\lambda^2 + p_2\lambda$$

$$\mu'_{2,1} = p_1^2 p_2\lambda^3 + 2p_1^2 p_2\lambda^2 + p_1p_2\lambda + p_1p_2\lambda^2$$

$$\mu'_{1,2} = p_1p_2^2\lambda^3 + 2p_1p_2^2\lambda^2 + p_1p_2\lambda + p_1p_2\lambda^2$$

$$\mu'_{2,2} = \mu_{[2,2]} + \mu'_{2,1} + \mu'_{1,2} - \mu'_{1,1}$$

Also

$$\mathrm{Cov}(X_1, X_2) = \lambda p_1 p_2.$$

On the other hand, the factorial cumulant generating function (fcgf) is

$$H(t_1,t_2) = \ln G(t_1,t_2)$$
$$= \lambda p_1 t_1 + \lambda p_2 t_2 + \lambda p_1 p_2 t_1 t_2, \qquad (2.2.10)$$

which implies that

$$\kappa_{[1,0]} = \lambda p_1,\ \kappa_{[0,1]} = \lambda p_2 \text{ and } \kappa_{[1,1]} = \lambda p_1 p_2,$$

and all other factorial cumulants are zero.

Conditional distribution and regression

Using the results given in Kocherlakota and Kocherlakota (1992, p. 94), the conditional pgf of X_1 given $X_2 = x_2$ is

$$\pi_{X_1}(t|x_2) = e^{\lambda p_1 q_2 (t-1)}(q_1 + p_1 t)^{x_2}, \qquad (2.2.11)$$

which shows that the conditional distribution is the sum of the independent random variables

$$X \sim \text{Poisson}(\lambda p_1 q_2) \text{ and } Y \sim \text{binomial}(x_2, p_1).$$

Hence the conditional expectation or regression of X_1 on X_2 is

$$E[X_1|x_2] = E(X) + E(Y)$$
$$= \lambda p_1 q_2 + p_1 x_2. \quad (2.2.12)$$

While the conditional variance is

$$Var(X_1|x_2) = Var(X) + Var(Y)$$
$$= \lambda p_1 q_2 + p_1 q_1 x_2. \quad (2.2.13)$$

Both the conditional expectation and variance are linear in x_2.

Similarly, the conditional distribution of X_2 given $X_1 = x_1$ is the sum of independent $\text{Poisson}(\lambda p_2 q_1)$ and $\text{binomial}(x_1, p_2)$. The conditional expectation and variance are

$$E[X_2|x_1] = \lambda p_2 q_1 + p_2 x_1, \quad (2.2.14)$$

$$Var[X_2|x_1] = \lambda p_2 q_1 + p_2 q_2 x_1. \quad (2.2.15)$$

The correlation coefficient between X_1 and X_2 obtained as the square root of the product of the two regression coefficients is

$$\rho_{X_1 X_2} = \sqrt{p_1 p_2}.$$

It is interesting to note that the correlation is independent of λ.

2.3 Case 2

Consider two random variables X_1 and X_2, having double binomial distribution with the pgf

$$\pi_2(t_1, t_2) = (q_1 + p_1 t_1)^n (q_2 + p_2 t_2)^n, \quad (2.3.1)$$

where $q_i = 1 - p_i$; i = 1, 2. In the pgf defined in (2.3.1) the index parameter and probabilities of success for each variable are assumed to be constants. Let us now consider the case when index parameter n is a random variable having the Poisson distribution with parameter λ and the pgf

$$\pi_1(t) = \exp[\lambda(t-1)]. \tag{2.3.2}$$

The resulting joint distribution of X_1 and X_2 will no longer be double binomial, but will be a stopped sum distribution having pgf

$$\pi(t_1,t_2) = \exp[\lambda\{(q_1 + p_1t_1)(q_2 + p_2t_2) - 1\}]. \tag{2.3.3}$$

In this case the pgf is equivalent to that of bivariate Poisson distribution. The results developed in section (2.2) also apply to this case.

2.4 Case 3

Consider the double binomial distribution with the pgf $(q_1 + p_1t_1)^{n_1k}(q_2 + p_2t_2)^{n_2k}$. Let k be a random variable having the Poisson distribution with parameter λ. The resulting stopped sum distribution has the pgf

$$\pi(t_1,t_2) = \exp[\lambda\{(q_1 + p_1t_1)^{n_1}(q_2 + p_2t_2)^{n_2} - 1\}]. \tag{2.4.1}$$

This distribution has been referred to as the Poisson-double binomial distribution. This distribution has five parameters, namely n_1, n_2, p_1, p_2 and λ, with n_1, n_2 being integers, $\lambda > 0$ and $0 \le p_1, p_2 \le 1$.

Marginal Distributions

The marginal generating functions are obtained by using the joint pgf (2.4.1). These are

$$\left.\begin{aligned} \pi_1(t) &= \exp[\lambda\{(q_1 + p_1t)^{n_1} - 1\}] \\ \pi_2(t) &= \exp[\lambda\{(q_2 + p_2t)^{n_2} - 1\}] \end{aligned}\right\}. \tag{2.4.2}$$

Equation (2.4.2) shows that the marginal distributions are Poisson-binomial, i.e.

$$X_1 \sim \text{Poisson-binomial } (\lambda, n_1, p_1)$$

$$X_2 \sim \text{Poisson-binomial } (\lambda, n_2, p_2).$$

Probability function

The probability function (pf) can be obtained from the pgf by expanding it in powers of t_1 and t_2. Due to the complex nature of the pgf it is not possible to give an explicit form for the pf. However we will develop recurrence relationships which can be exploited to determine the probabilities in specific instances.

Form 1: Let

$$T_{n_1} = (q_1 + p_1 t_1)^{n_1}$$
$$R_{n_2} = (q_2 + p_2 t_2)^{n_2}$$

The pgf in (2.4.1) can be written in terms of T_{n_1} and R_{n_2} as

$$\pi(t_1, t_2) = \exp\left[\lambda\left(T_{n_1} R_{n_2} - 1\right)\right]. \tag{2.4.3}$$

Differentiating T_{n_1} r times with respect to (w.r.t.) t_1, we get

$$\frac{d^r T_{n_1}}{dt_1^r} = \frac{n_1!}{(n_1 - r)!} p_1^r (q_1 + p_1 t_1)^{n_1 - r}, \tag{2.4.4}$$

while differentiating R_{n_2} s times w.r.t. t_2, we get

$$\frac{d^s R_{n_2}}{dt_2^s} = \frac{n_2!}{(n_2 - s)!} p_2^s (q_2 + p_2 t_2)^{n_2 - s}. \tag{2.4.5}$$

In what follows we will write

$$\pi^{(r,s)}(t_1, t_2) = \frac{\partial^{r+s}}{\partial t_1^r \partial t_2^s} \pi(t_1, t_2).$$

Using this notation

$$\pi^{(1,0)}(t_1, t_2) = \exp\left[\lambda\left(T_{n_1} R_{n_2} - 1\right)\right] \lambda R_{n_2} n_1 p_1 T_{n_{1-1}}. \tag{2.4.6}$$

Let $\pi = \pi(t_1, t_2)$, so that the equation in (2.4.6) can be written as

$$\pi^{(1,0)}(t_1, t_2) = \lambda n_1 p_1 T_{n_1 - 1} R_{n_2} \pi. \tag{2.4.7}$$

Hence the rth differential coefficient of π w.r.t. t_1 for $r \geq 1$ is given by

$$\pi^{(r,0)}(t_1, t_2) = \frac{\partial^{r-1}}{\partial t_1^{r-1}} \left(\lambda n_1 p_1 T_{n_1 - 1} R_{n_2} \pi\right),$$

which, upon using Leibnitz's theorem, yields

$$\pi^{(r,0)}(t_1,t_2) = \lambda n_1 p_1 R_{n_2} \sum_{z=0}^{r-1} \binom{r-1}{z} \pi^{(r-1-z,0)}(t_1,t_2)$$

$$\frac{(n_1-1)!}{(n_1-1-z)!} p_1^z (q_1 + p_1 t_1)^{n_1-1-z}, \qquad r \geq 1.$$

Rearranging the order of summation on the right hand side

$$\pi^{(r,0)}(t_1,t_2) = \lambda R_{n_2} \sum_{j=1}^{r} \frac{(r-1)!}{(j-1)!(r-j)!} \pi^{(r-j,0)}(t_1,t_2)$$

$$\frac{n_1!}{(n_1-j)!} p_1^j (q_1 + p_1 t_1)^{n_1-j},$$

$$r \geq 1. \qquad (2.4.8)$$

Since the (r,s)th differential coefficient of π is given by

$$\pi^{(r,s)}(t_1,t_2) = \frac{\partial^s}{\partial t_2^s} \pi^{(r,0)}(t_1,t_2),$$

from (2.4 8) we have

$$\pi^{(r,s)}(t_1,t_2) = \lambda \sum_{j=1}^{r} \frac{(r-1)!}{(j-1)!(r-j)!} \frac{n_1!}{(n_1-j)!} p_1^j (q_1 + p_1 t_1)^{n_1-j}$$

$$\sum_{k=0}^{s} \binom{s}{k} \pi^{(r-j,s-k)}(t_1,t_2) \frac{d^k}{dt_2^k} R_{n_2}, \qquad r \geq 1, s \geq 0.$$

But

$$\frac{d^k}{dt_2^k} R_{n_2} = \frac{n_2!}{(n_2-k)!} p_2^k (q_2 + p_2 t_2)^{n_2-k},$$

and hence

$$\pi^{(r,s)}(t_1,t_2) = \lambda \sum_{j=1}^{r} \sum_{k=0}^{s} \frac{(r-1)!s!n_1!n_2!}{(j-1)!(r-j)!k!(s-k)!(n_1-j)!(n_2-k)!} p_1^j p_2^k$$

$$(q_1 + p_1 t_1)^{n_1-j} (q_2 + p_2 t_2)^{n_2-k} \pi^{(r-j,s-k)}(t_1,t_2),$$

$$r \geq 1, s \geq 0. \qquad (2.4.9)$$

Thus

$$\frac{1}{r!s!}\pi^{(r,s)}(t_1,t_2)=\frac{\lambda}{r}\sum_{j=1}^{r}\sum_{k=0}^{s}j\frac{n_1!}{(n_1-j)!j!}p_1^j(q_1+p_1t_1)^{n_1-j}\frac{n_2!}{(n_2-k)!k!}$$

$$p_2^k(q_2+p_2t_2)^{n_2-k}\frac{1}{(r-j)!(s-k)!}\pi^{(r-j,s-k)}(t_1,t_2),$$

$$r\geq 1, s\geq 0. \quad (2.4.10)$$

Setting $t_1 = 0$ and $t_2 = 0$ in (2.4.10), the recurrence relationship for the probability function can be obtained as

$$f(r,s)=\frac{\lambda}{r}\sum_{j=1}^{r}\sum_{k=0}^{s}jb(n_1,j)b(n_2,k)f(r-j,s-k), \quad r\geq 1, s\geq 0, \quad (2.4.11)$$

where b(,) is the binomial probability with the specified parameters. Similarly, differentiating the pgf in (2.4.3) w.r.t. t_2 once, gives

$$\pi^{(0,1)}(t_1,t_2)=\lambda n_2 p_2 T_{n_1} R_{n_2-1}\pi, \quad (2.4.12)$$

and hence, for $s\geq 1$

$$\pi^{(0,s)}(t_1,t_2)=\frac{\partial^{s-1}}{\partial t_2^{s-1}}\left[\lambda n_2 p_2 T_{n_1} R_{n_2-1}\pi\right] \quad (2.4.13)$$

$$=\lambda n_2 p_2 T_{n_1}\sum_{z=0}^{s-1}\binom{s-1}{z}\pi^{(0,s-1-z)}(t_1,t_2)\frac{(n_2-1)!}{(n_2-1-z)!}p_2^z(q_2+p_2t_2)^{n_2-1-z}$$

$$=\lambda T_{n_1}\sum_{k=1}^{s}\frac{(s-1)!}{(s-k)!(k-1)!}\frac{n_2!}{(n_2-k)!}p_2^k(q_2+p_2t_2)^{n_2-k}\pi^{(0,s-k)}(t_1,t_2),$$

$$s\geq 1. \quad (2.4.14)$$

Hence

$$f(r,s)=\frac{\lambda}{s}\sum_{j=0}^{r}\sum_{k=1}^{s}kb(n_1,j)b(n_2,k)f(r-j,s-k), \quad r\geq 0, s\geq 1. \quad (2.4.15)$$

Form 2: Let

$$U_{jk}^{(r,s)}=\lambda\binom{n_1}{r-j}\binom{n_2}{s-k}p_1^{r-j}p_2^{s-k}T_{n_1-r+j}R_{n_2-s+k}, \quad (2.4.16)$$

where $T=(q_1+p_1t_1)$ and $R=(q_2+p_2t_2)$. Writing $\theta_1=\frac{p_1}{T}$, $\theta_2=\frac{p_2}{R}$, we have

$$U_{jk}^{(r,s)} = \lambda \frac{n_1!}{(r-j)!(n_1-r+j)!} \frac{n_2!}{(s-k)!(n_2-s+k)!} \theta_1^{r-j} \theta_2^{s-k} T_{n_1} R_{n_2}. \qquad (2.4.17)$$

It can be seen that

$$U_{00}^{(0,0)} = U = \lambda T_{n_1} R_{n_2}. \qquad (2.4.18)$$

The second recurrence relationship for the pgf is obtained by using the first form given in (2.4.10) and substituting the terms defined above in the right hand side of (2.4.10) as below

$$\frac{1}{r!s!} \pi^{(r,s)}(t_1,t_2) = \frac{n_1 \lambda T_{n_1} R_{n_2}}{r} \sum_{j=1}^{r} \sum_{k=0}^{s} \frac{(n_1-1)!}{(n_1-j)!(j-1)!} \frac{n_2!}{(n_2-k)!k!} p_1^j p_2^k$$
$$(q_1+p_1t_1)^{-j}(q_2+p_2t_2)^{-k} \frac{1}{(r-j)!(s-k)!} \pi^{(r-j,s-k)}(t_1,t_2),$$

which can again be written as

$$\frac{1}{r!s!} \pi^{(r,s)}(t_1,t_2) = \frac{n_1 U \theta_1}{r} \sum_{j=1}^{r} \sum_{k=0}^{s} \frac{(n_1-1)!}{(n_1-j)!(j-1)!} \frac{n_2!}{(n_2-k)!k!} \theta_1^{j-1} \theta_2^k$$
$$\frac{1}{(r-j)!(s-k)!} \pi^{(r-j,s-k)}(t_1,t_2),$$

and finally by changing the variables, above equation can be written as

$$\frac{1}{r!} \frac{1}{s!} \pi^{(r,s)} = \frac{n_1 U \theta_1}{r} \sum_{j=1}^{r} \sum_{k=0}^{s} \binom{n_1-1}{r-j} \binom{n_2}{s-k} \theta_1^{r-j} \theta_2^{s-k} \frac{\pi^{(j-1,k)}}{(j-1)!(k)!},$$
$$r \geq 1, s \geq 0. \qquad (2.4.19)$$

Alternatively,

$$\frac{1}{r!} \frac{1}{s!} \pi^{(r,s)} = \frac{n_2 U \theta_2}{s} \sum_{j=0}^{r} \sum_{k=1}^{s} \binom{n_1}{r-j} \binom{n_2-1}{s-k} \theta_1^{r-j} \theta_2^{s-k} \frac{\pi^{(j,k-1)}}{j!(k-1)!},$$
$$r \geq 0, s \geq 1. \qquad (2.4.20)$$

To get the general form of the probability function, substitute $t_1 = 0$ and $t_2 = 0$ in either of the equations (2.4.19) or (2.4.20). Considering equation (2.4.19), the general form of the probability function is

$$f(r,s)=\frac{\lambda n_1 q_1^{n_1} q_2^{n_2}}{r}\sum_{j=1}^{r}\sum_{k=0}^{s}\binom{n_1-1}{r-j}\binom{n_2}{s-k}\left(\frac{p_1}{q_1}\right)^{r-j}\left(\frac{p_2}{q_2}\right)^{s-k} f(j-1,k),$$

$$r \geq 1, s \geq 0. \qquad (2.4.21)$$

Using equation (2.4.20)

$$f(r,s)=\frac{\lambda n_2 q_1^{n_1} q_2^{n_2}}{s}\sum_{j=0}^{r}\sum_{k=1}^{s}\binom{n_1}{r-j}\binom{n_2-1}{s-k}\left(\frac{p_1}{q_1}\right)^{r-j}\left(\frac{p_2}{q_2}\right)^{s-k} f(j,k-1),$$

$$r \geq 0, s \geq 1. \qquad (2.4.22)$$

Using the recurrence relations in (2.4.11) and (2.4.15) it is possible to write

$$f(1,0)=\lambda f(0,0) b(n_1,1) b(n_2,0)$$

$$f(0,1)=\lambda f(0,0) b(n_1,0) b(n_2,1)$$

$$f(1,1)=\lambda\left[b(n_1,1)b(n_2,0)f(0,1)+b(n_1,1)b(n_2,1)f(0,0)\right]$$

$$f(2,0)=\frac{\lambda}{2}\left[b(n_1,1)b(n_2,0)f(1,0)+2b(n_1,2)b(n_2,0)f(0,0)\right]$$

$$f(0,2)=\frac{\lambda}{2}\left[b(n_1,0)b(n_2,1)f(0,1)+2b(n_1,0)b(n_2,2)f(0,0)\right]$$

$$\begin{aligned}f(2,1)=\frac{\lambda}{2}\big[&b(n_1,1)b(n_2,0)f(1,1)+b(n_1,1)b(n_2,1)f(1,0)\\&+2b(n_1,2)b(n_2,0)f(0,1)+2b(n_1,2)b(n_2,1)f(0,0)\big]\end{aligned}$$

$$\begin{aligned}f(2,2)=\frac{\lambda}{2}\big[&b(n_1,1)b(n_2,0)f(1,2)+b(n_1,1)b(n_2,1)f(1,1)\\&+b(n_1,1)b(n_2,2)f(1,0)+2b(n_1,2)b(n_2,0)f(0,2)\\&+2b(n_1,2)b(n_2,1)f(0,1)+2b(n_1,2)b(n_2,2)f(0,0)\big].\end{aligned}$$

It is also possible to express the probabilities in terms of f(0,0), where $f(0,0)=\exp\left[\lambda\left(q_1^{n_1}q_2^{n_2}-1\right)\right]$. For example

$$f(1,0)=\lambda f(0,0) b(n_1,1) b(n_2,0)$$

$$f(0,1) = \lambda f(0,0) b(n_1,0) b(n_2,1)$$

$$f(1,1) = f(0,0)\left[\lambda b(n_1,1) b(n_2,1) + \lambda^2 b(n_1,1) b(n_1,0) b(n_2,1) b(n_2,0)\right]$$

$$f(2,0) = f(0,0)\left[\lambda b(n_1,2) b(n_2,0) + \lambda^2 b(n_1,2) b(n_1,0) b^2(n_2,0)\right]$$

$$f(2,1) = f(0,0)\left[\lambda b(n_1,2) b(n_2,1) + \lambda^2 b(n_1,2) b(n_1,0) b(n_2,1) b(n_2,0) + \lambda^2 b^2(n_1,1) b(n_2,1) b(n_2,0) + \frac{\lambda^3}{2} b^2(n_1,1) b(n_1,0) b(n_2,1) b^2(n_2,0)\right]$$

$$f(2,2) = f(0,0)\Big[\lambda b(n_1,2) b(n_2,2) + \lambda^2 \Big\{ b(n_1,2) b(n_1,0) b^2(n_2,1) + b(n_1,2) b(n_1,0) b(n_2,2) b(n_2,0) + \frac{1}{2} b^2(n_1,1) b(n_2,2) b(n_2,0) + \frac{1}{2} b^2(n_1,1) b^2(n_2,1) + \frac{1}{2} b^2(n_1,1) b(n_2,2) b(n_2,0) \Big\} + \lambda^3 \Big\{ b(n_1,2) b^2(n_1,0) b^2(n_2,1) b(n_2,0) + b^2(n_1,1) b(n_1,0) b^2(n_2,1) b(n_2,0) + \frac{1}{4} b^2(n_1,1) b(n_1,0) b(n_2,2) b(n_2,0) \Big\} + \frac{\lambda^4}{4} b^2(n_1,1) b^2(n_1,0) b^2(n_2,1) b^2(n_2,0)\Big]$$

$$f(3,1) = f(0,0)\Big[\lambda b(n_1,3) b(n_2,1) + \lambda^2 b(n_1,3) b(n_1,0) b(n_2,1) b(n_2,0) + \frac{5}{3} \lambda^2 b(n_1,2) b(n_1,1) b(n_2,1) b(n_2,0) + \frac{2}{3} \lambda^3 b(n_1,2) b(n_1,1) b(n_1,0) b(n_2,1) b^2(n_2,0) + \frac{1}{2} \lambda^3 b^3(n_1,1) b(n_2,1) b^2(n_2,0) + \frac{1}{6} \lambda^4 b^3(n_1,1) b(n_1,0) b(n_2,1) b^3(n_2,0)\Big],$$

where $b(n_i,j)$ represents the binomial probability of getting a "j" value with parameters n_i and p_i, i = 1, 2. It is of interest to note that the expressions for probabilities have a common feature, that is, they all can be expressed in terms of f(0,0), powers of λ and products of powers of two binomial probabilities.

Following are the tables of pf with $n_1 = n_2 = 2$, $p_1 = p_2 = 0.25$ and for values of $\lambda = 0.5, 1.0, 1.5$. These probabilities are obtained by using the S-plus program pfc931 given in appendix.

Table 2.4.1 $\lambda = 0.5$

x_1	x_2 = 0	1	2	3	4
0	0.7105	0.0749	0.0164	0.0015	0.0002
1	0.0749	0.0579	0.0153	0.0022	0.0003
2	0.0164	0.0153	0.0058	0.0014	0.0002
3	0.0015	0.0022	0.0014	0.0005	0.0001
4	0.0002	0.0033	0.0002	0.0001	0

Table 2.4.2 $\lambda = 1$

x_1	x_2 = 0	1	2	3	4	5
0	0.5048	0.1065	0.0290	0.0045	0.0007	0.0001
1	0.1065	0.0934	0.0329	0.0075	0.0015	0.0002
2	0.0290	0.0329	0.0168	0.0055	0.0013	0.0003
3	0.0045	0.0075	0.0055	0.0024	0.0007	0.0002
4	0.0007	0.0015	0.0013	0.0007	0.0003	0.0001
5	0.0001	0.0002	0.0003	0.0002	0.0001	0

Table 2.4.3 $\lambda = 1.5$

x_1	x_2 = 0	1	2	3	4	5	6
0	0.3587	0.1135	0.036	0.0079	0.0016	0.0003	0
1	0.1135	0.1116	0.0482	0.0143	0.0035	0.0007	0.0001
2	0.036	0.0428	0.0294	0.0117	0.0035	0.0008	0.0002
3	0.0079	0.0143	0.0117	0.0059	0.0022	0.0006	0.0002
4	0.0016	0.0035	0.0035	0.0022	0.0010	0.0003	0.0001
5	0.0003	0.0008	0.0008	0.0006	0.0003	0.0001	0
6	0	0.0002	0.0002	0.0002	0.0001	0	0

It can be seen from the pf tables that as λ becomes large while other parameters remain fixed, the probability of (0,0) cell becomes small and table spreads out more.

Factorial moments and cumulants

The factorial moment generating function is given by

$$G(t_1,t_2)=\exp\left[\lambda\left\{(1+p_1t_1)^{n_1}(1+p_2t_2)^{n_2}-1\right\}\right] \tag{2.4.23}$$

and the factorial moments are obtained by using the relationship

$$\mu_{[r,s]}=\frac{\partial^{r+s}}{\partial t_1^r \partial t_2^s}G(t_1,t_2)\Big|_{t_1=0,t_2=0},$$

or

$$\mu_{[r,s]}=\frac{\partial^{r+s}}{\partial t_1^r \partial t_2^s}\pi(t_1,t_2)\Big|_{t_1=1,t_2=1}. \tag{2.4.24}$$

Using relationship in (2.4.24), in particular

$$\mu_{[1,0]}=\lambda n_1 p_1$$

$$\mu_{[0,1]}=\lambda n_2 p_2$$

$$\mu_{[1,1]}=\lambda n_1 n_2 p_1 p_2+\lambda^2 n_1 n_2 p_1 p_2$$

$$\mu_{[2,0]}=\lambda n_1(n_1-1)p_1^2+\lambda^2 n_1^2 p_1^2$$

$$\mu_{[0,2]}=\lambda n_2(n_2-1)p_2^2+\lambda^2 n_2^2 p_2^2$$

$$\begin{aligned}\mu_{[2,1]}=&\lambda n_1 n_2(n_1-1)p_1^2 p_2+\lambda^2 n_1 n_2(n_1-1)p_1^2 p_2\\&+2\lambda^2 n_1^2 n_2 p_1^2 p_2+\lambda^3 n_1^2 n_2 p_1^2 p_2\end{aligned}$$

$$\begin{aligned}\mu_{[1,2]}=&\lambda n_1 n_2(n_2-1)p_1 p_2^2+\lambda^2 n_1 n_2(n_2-1)p_1 p_2^2\\&+2\lambda^2 n_1 n_2^2 p_1 p_2^2+\lambda^3 n_1 n_2^2 p_1 p_2^2\end{aligned}$$

and

$$\begin{aligned}\mu_{[2,2]}=&\lambda n_1(n_1-1)n_2(n_2-1)p_1^2 p_2^2+\lambda^2\{7n_1^2 n_2^2 p_1^2 p_2^2\\&-3n_1^2 n_2 p_1^2 p_2^2-3n_1 n_2^2 p_1^2 p_2^2+n_1 n_2 p_1^2 p_2^2\}+\lambda^3\{6\\&n_1^2 n_2^2 p_1^2 p_2^2-n_1^2 n_2 p_1^2 p_2^2-n_1 n_2^2 p_1^2 p_2^2\}+\lambda^4 n_1^2 n_2^2 p_1^2 p_2^2.\end{aligned}$$

Using the recurrence relations in (2.4.19) and (2.4.20), we can get the recurrence relation among the factorial moments. From (2.4.19)

$$\mu_{[r,s]} = (r-1)!s!n_1p_1\lambda\sum_{j=1}^{r}\sum_{k=0}^{s}\binom{n_1-1}{r-j}\binom{n_2}{s-k}\frac{p_1^{r-j}p_2^{s-k}}{(j-1)!k!}\mu_{[j-1,k]},$$

for $r \geq 1$ and $s \geq 0$. From (2.4.20)

$$\mu_{[r,s]} = (s-1)!r!n_2\lambda p_2\sum_{j=0}^{r}\sum_{k=1}^{s}\binom{n_1}{r-j}\binom{n_2-1}{s-k}\frac{p_1^{r-j}p_2^{s-k}}{j!(k-1)!}\mu_{[j,k-1]},$$

for $r \geq 0$ and $s \geq 1$.

The factorial cumulant generating function is defined by

$$H(t_1,t_2) = \log G(t_1,t_2)$$

and for this particular distribution factorial cumulant generating function is given by

$$H(t_1,t_2) = \lambda\left[(1+p_1t_1)^{n_1}(1+p_2t_2)^{n_2}-1\right]. \tag{2.4.25}$$

Expanding the right hand side of (2.4.25)

$$H(t_1,t_2) = \lambda\left[\sum_{r=0}^{n_1}\binom{n_1}{r}p_1^rt_1^r\sum_{s=0}^{n_2}\binom{n_2}{s}p_2^st_2^s-1\right]$$

$$= \lambda\left[\sum_{r=0}^{n_1}\sum_{s=0}^{n_2}\binom{n_1}{r}\binom{n_2}{s}p_1^rp_2^st_1^rt_2^s-1\right]. \tag{2.4.26}$$

The coefficient of $\frac{t_1^r}{r!}\frac{t_2^s}{s!}$ in the above expansion gives the (r,s)th factorial cumulant, $\kappa_{[r,s]}$. It can be seen from (2.4.26) that $\kappa_{[r,s]}$ is zero if $r > n_1$ and/or $s > n_2$. Thus for some values of r and s, the factorial cumulants are

$$\kappa_{[1,0]} = \lambda n_1p_1$$

$$\kappa_{[0,1]} = \lambda n_2p_2$$

$$\kappa_{[1,1]} = \lambda n_1n_2p_1p_2$$

$$\kappa_{[2,0]} = \lambda n_1(n_1-1)p_1^2$$

$$\kappa_{[0,2]} = \lambda n_2(n_2-1)p_2^2$$

$$\kappa_{[2,2]} = \lambda n_1(n_1-1)n_2(n_2-1)p_1^2p_2^2.$$

The raw moments can be found using the factorial moments. In special instances we have

$$\mu'_{1,0} = \lambda n_1 p_1$$

$$\mu'_{0,1} = \lambda n_2 p_2$$

$$\mu'_{1,1} = \lambda n_1 n_2 p_1 p_2 + \lambda^2 n_1 n_2 p_1 p_2$$

$$\mu'_{2,0} = \lambda n_1(n_1-1)p_1^2 + \lambda^2 n_1^2 p_1^2 + \lambda n_1 p_1$$

$$\mu'_{0,2} = \lambda n_2(n_2-1)p_2^2 + \lambda^2 n_2^2 p_2^2 + \lambda n_2 p_2$$

$$\mu'_{2,1} = \lambda n_1(n_1-1)n_2 p_1^2 p_2 + \lambda^2 n_1(n_1-1)p_1^2 p_2 + 2\lambda^2 n_1^2 n_2 p_1^2 p_2 + \lambda^3 n_1^2 n_2 p_1^2 p_2 + \lambda n_1 n_2 p_1 p_2 + \lambda^2 n_1 n_2 p_1 p_2$$

$$\mu'_{1,2} = \lambda n_1 n_2(n_2-1)p_1 p_2^2 + \lambda^2 n_1 n_2(n_2-1)p_1 p_2^2 + 2\lambda^2 n_1 n_2^2 p_1 p_2^2 + \lambda^3 n_1 n_2^2 p_1 p_2^2 + \lambda n_1 n_2 p_1 p_2 + \lambda^2 n_1 n_2 p_1 p_2$$

$$\mu'_{2,2} = \lambda n_1(n_1-1)n_2(n_2-1)p_1^2 p_2^2 + \lambda^2 n_1(n_1-1)n_2^2 p_1^2 p_2^2 + \lambda^2 n_1(n_1-1)n_2(2n_2-1)p_1^2 p_2^2 + 2\lambda^2 n_1^2 n_2(2n_2-1)p_1^2 p_2^2 + \lambda^3 n_1^2 n_2(3n_2-1)p_1^2 p_2^2 + \lambda^4 n_1^2 n_2^2 p_1^2 p_2^2 + \mu'_{2,1} + \mu'_{1,2} - \mu'_{1,1}$$

Hence

$$\mu_{1,1} = \lambda n_1 n_2 p_1 p_2$$

$$\mu_{2,1} = \mu'_{2,1} - \mu'_{0,1}\mu'_{2,0} + 2\mu'^2_{1,0}\mu'_{0,1} - 2\mu'_{1,0}\mu'_{1,1}$$

$$\mu_{1,2} = \mu'_{1,2} - \mu'_{1,0}\mu'_{0,2} + 2\mu'_{1,0}\mu'^2_{0,1} - 2\mu'_{0,1}\mu'_{1,1}$$

$$\mu_{2,2} = \mu'_{2,2} - 2\mu'_{2,1}\mu'_{0,1} - 2\mu'_{1,2}\mu'_{1,0} + \mu'^2_{1,0}\mu'_{0,2} + \mu'^2_{0,1}\mu'_{2,0} + 4\mu'_{1,1}\mu'_{1,0}\mu'_{0,1} - 3\mu'^2_{1,0}\mu'^2_{0,1}$$

From earlier results, we have

$$E(X_1) = \lambda n_1 p_1$$

$$E(X_2) = \lambda n_2 p_2$$

$$\text{Var}(X_1) = \lambda[n_1 p_1(n_1 p_1 + q_1)]$$

$$\text{Var}(X_2) = \lambda[n_2 p_2(n_2 p_2 + q_2)]$$

$$\text{Cov}(X_1, X_2) = \lambda n_1 n_2 p_1 p_2,$$

and

$$\rho_{X_1,X_2} = \frac{\sqrt{n_1 n_2 p_1 p_2}}{\sqrt{(n_1 p_1 + q_1)(n_2 p_2 + q_2)}}.$$

It can be seen that in the case of the Poisson-double binomial distribution the correlation coefficient is also independent of λ.

Conditional distribution and regression

To get the conditional distributions of X_2 given $X_1 = r$, Theorem 1.3.1 given in Kocherlakota and Kocherlakota (1992, p. 13) is used. From the second form of the pgf, we can write

$$\pi^{(r,0)}(t_1,t_2) = \lambda R_{n_2} \sum_{j=1}^{r} \frac{(r-1)!}{(j-1)!} \frac{n_1!}{(n_1-j)!} \frac{\pi^{(r-j,0)}(t_1,t_2)}{(r-j)!} p_1^j (q_1 + p_1 t_1)^{n_1-j}, \quad r \geq 1. \quad (2.4.28)$$

Equation (2.4.28), yields

$$\pi^{(r,0)}(0,t) = \lambda(q_2 + p_2 t)^{n_2} \sum_{j=1}^{r} \frac{(r-1)!}{(j-1)!(r-j)!} \frac{n_1!}{(n_1-j)!} \pi^{(r-j,0)}(0,t) p_1^j q_1^{n_1-j}, \quad (2.4.29)$$

and

$$\pi^{(r,0)}(0,1) = \lambda \sum_{j=1}^{r} \frac{(r-1)!}{(j-1)!(r-j)!} \frac{n_1!}{(n_1-j)!} \pi^{(r-j,0)}(0,1) p_1^j q_1^{n_1-j}. \quad (2.4.30)$$

The conditional pgf of X_2 given $X_1 = r$ is given by

$$\pi_{X_2}(t|r) = \frac{\pi^{(r,0)}(0,t)}{\pi^{(r,0)}(0,1)}, \quad (2.4.31)$$

and for the distribution under consideration it becomes

$$\pi_{X_2}(t|r) = \frac{(q_2+p_2t)^{n_2}\sum_{j=1}^{r}\frac{(r-1)!}{(j-1)!(r-j)!}\frac{n_1!}{(n_1-j)!}\pi^{(r-j,0)}(0,t)p_1^j q_1^{n_1-j}}{\sum_{j=1}^{r}\frac{(r-1)!}{(j-1)!(r-j)!}\frac{n_1!}{(n_1-j)!}\pi^{(r-j,0)}(0,1)p_1^j q_1^{n_1-j}}, \quad r \geq 1. \tag{2.4.32}$$

For r = 0, using (2.4.31) directly, we get

$$\pi_{X_2}(t|0) = \frac{\exp\left[\lambda\left\{q_1^{n_1}(q_2+p_2t)^{n_2}-1\right\}\right]}{\exp\left[\lambda\left\{q_1^{n_1}-1\right\}\right]}. \tag{2.4.33}$$

In general for $r \geq 1$

$$\pi_{X_2}(t|r) = (q_2+p_2t)^{n_2}\sum_{j=1}^{r} w_j \pi_{X_2}(t|r-j), \tag{2.4.34}$$

where

$$w_j = \frac{\frac{(r-1)!}{(j-1)!(r-j)!}\frac{n_1!}{(n_1-j)!}\pi^{(r-j,0)}(0,1)p_1^j q_1^{r-j}}{\sum_{i=1}^{r}\frac{(r-1)!}{(i-1)!(r-i)!}\frac{n_1!}{(n_1-i)!}\pi^{(r-i,0)}(0,1)p_1^i q_1^{r-i}}. \tag{2.4.35}$$

It can be seen that the conditional distribution of X_2 given $X_1 = r$ is the sum of two independent random variables Z_1 and Z_2 where $Z_1 \sim \text{binomial}(n_2, p_2)$ and Z_2, that is a mixture of r random variables with pgf $\pi_{X_2}(t|r-j)$ with weights w_j. For $r \geq 1$

$$E(X_2|r) = E(Z_1) + E(Z_2). \tag{2.4.36}$$

Using the result

$$E(X_2|r) = \frac{\pi^{(r,1)}(0,1)}{\pi^{(r,0)}(0,1)},$$

for r = 0

$$E(X_2|0) = \lambda n_2 p_2 q_1^{n_1}. \tag{2.4.37}$$

Conditional expectations for $r \geq 1$ can be obtained by using (2.4.36) and (2.4.37). For example for r = 1

$$E(X_2|1) = n_2p_2 + w_1E(X_2|0)$$

since $w_1 = 1$

$$E(X_2|1) = n_2p_2 + \lambda n_2 p_2 q_1^{n_1}. \tag{2.4.38}$$

Conditional variance of X_2 given $X_1 = r$ is obtained by the relation

$$\mathrm{Var}(X_2|r) = \mathrm{Var}(Z_1) + \mathrm{Var}(Z_2), \tag{2.4.39}$$

where $\mathrm{Var}(Z_1) = n_2p_2q_2$ and

$$\mathrm{Var}(Z_2) = E(Z_2^2) - E^2(Z_2). \tag{2.4.40}$$

To get $E(Z_2^2)$, using the pgf of Z_2, it is known that

$$E[Z_2(Z_2 - 1)] = \frac{d^2}{dt^2}\pi_{Z_2}(t)\big|_{t=1}, \tag{2.4.41}$$

where

$$\pi_{Z_2}(t) = \sum_{j=1}^{r} w_j \pi_{X_2}(t|r-j), \qquad r \geq 1. \tag{2.4.42}$$

Another way of writing pgf in (2.4.42) is

$$\pi_{Z_2}(t) = \sum_{j=0}^{r-1} w_{r-j} \pi_{X_2}(t|j), \qquad r \geq 2. \tag{2.4.43}$$

Differentiating the pgf in (2.4.43) twice with respect to t, the following expression is obtained

$$\begin{aligned}\frac{d^2}{dt^2}\pi_{Z_2}(t) &= \sum_{j=0}^{r-1} w_{r-j}\frac{d^2}{dt^2}\pi_{X_2}(t|j)\\ &= w_r \frac{1}{\pi^{(0,0)}(0,1)}\frac{d^2}{dt^2}\pi^{(0,0)}(0,t)\\ &+\sum_{j=1}^{r-1} w_{r-j}\frac{1}{\pi^{(j,0)}(0,1)}\frac{d^2}{dt^2}\pi^{(j,0)}(0,t), \qquad r \geq 2.\end{aligned}$$

For r = 1 using (2.4.42) and since $w_1 = 1$

$$\frac{d^2}{dt^2}\pi_{Z_2}(t) = \frac{d^2}{dt^2}\frac{\pi^{(0,0)}(0,t)}{\pi^{(0,0)}(0,1)}. \tag{2.4.44}$$

For $r \geq 2$

$$E\left[Z_2(Z_2-1)\right]=w_r\frac{1}{\pi^{(0,0)}(0,1)}\pi^{(0,2)}(0,1)+$$

$$\sum_{j=1}^{r-1}w_{r-j}\frac{1}{\pi^{(j,0)}(0,1)}\lambda\sum_{k=1}^{j}\frac{(j-1)!}{(k-1)!(j-k)!}\frac{n_1!}{(n_1-k)!}p_1^k q_1^{n_1-k}$$

$$\left\{n_2(n_2-1)p_2^2\pi^{(j-k,0)}(0,1)+2n_2p_2\pi^{(j-k,1)}(0,1)+\pi^{(j-k,2)}(0,1)\right\}.$$

Hence

$$E\left(Z_2^2\right)=E\left[Z_2(Z_2-1)\right]+E(Z_2),$$

and variance of Z_2 can be obtained by using (2.4.40). Substituting this value of variance of Z_2 in (2.4.39) $\mathrm{Var}(X_2|r)$ can be obtained for specified r values. Similarly $\mathrm{Var}(X_1|s)$ can be obtained.

Limiting distributions

To get the limiting form of the Poisson-double binomial distribution, three cases are of interest. For simplicity factorial cumulant generating function is used.

(1) $p_1 \to 0$ and $p_2 \to 0$: Suppose λ, n_1 and n_2 are fixed, then the fcgf in (2.4.27) becomes

$$-\lambda+\lambda\left(1+n_1p_1t_1+O\left(p_1^2\right)\right)\left(1+n_2p_2t_2+O\left(p_2^2\right)\right)$$

$$=-\lambda+\lambda(1+n_1p_1t_1)(1+n_2p_2t_2)$$

$$=(\lambda n_1p_1)t_1+(\lambda n_2p_2)t_2+(\lambda n_1n_2p_1p_2)t_1t_2,$$

where $O\left(p_i^2\right)$ i = 1, 2, means the terms of order zero. The resulting fcgf corresponds to the bivariate Poisson distribution with the parametric equivalence $\lambda_1+\lambda_3=\lambda n_1p_1$, $\lambda_2+\lambda_3=\lambda n_2p_2$ and $\lambda_3=\lambda n_1n_2p_1p_2$.

(2) $n_1 \to \infty$ and $n_2 \to \infty$: In this case n_1p_1 and n_2p_2 are assumed to be finite. The fcgf in (2.4.27) will become

$$\lambda\left[\left(1+\frac{n_1p_1t_1}{n_1}\right)^{n_1}\left(1+\frac{n_2p_2t_2}{n_2}\right)^{n_2}-1\right]$$

taking the limit

$$=-\lambda+\lambda\exp(n_1p_1t_1+n_2p_2t_2),$$

which is equivalent to the pgf

$$\exp\left[-\lambda\left\{\exp\left(n_1 p_1 (t_1 - 1) + n_2 p_2 (t_2 - 1)\right) - 1\right\}\right],$$

which is the pgf of the two independent Poisson variates with parameters $n_1 p_1$ and $n_2 p_2$ respectively, compounded by another Poisson variate having parameter λ.

(3) $p_1 \to 1$ and $p_2 \to 1$: In this case the fcgf becomes

$$-\lambda + \lambda(1 + t_1)^{n_1}(1 + t_2)^{n_2},$$

which corresponds to the pgf

$$\exp\left[-\lambda + \lambda t_1^{n_1} t_2^{n_2}\right].$$

3

ESTIMATION

3.1 Introduction

In this chapter we will develop estimation procedures for the parameters of the bivariate discrete distributions described in chapter 2. Various methods of estimation are available for estimating the parameters of bivariate discrete distributions. Among all of these methods, method of maximum likelihood (ML) is the most efficient one, if regularity conditions hold. Maximum likelihood estimators are asymptotically unbiased, normally distributed and have minimum variance. One major problem with ML equations is that; usually they do not provide solution in a closed form, thus leading to the use of iterative procedures. In most instances the method of moments estimation does provide rapid solutions to the estimating equations. Alternative method of estimation uses the zero-zero-cell frequency in addition to the first marginal moments. Two iterative procedures are suggested in literature. These are:

1. Newton-Raphson Method: Let $f_i(\underline{\theta})$, i=1, 2,..., k be a function of k-dimensional vector of parameters $\underline{\theta}$. To solve the equation $f_i(\underline{\theta}) = 0$ iteratively we can expand $f_i(\underline{\theta})$ in a Taylor's series around a trial solution $\underline{\theta}^{(0)}$ as

$$f_i(\underline{\theta}) = f_i\left(\underline{\theta}^{(0)}\right) + \left\{\left(\theta_1 - \theta_1^{(0)}\right)\frac{\partial}{\partial\theta_1} + \left(\theta_2 - \theta_2^{(0)}\right)\frac{\partial}{\partial\theta_2} + \ldots + \left(\theta_k - \theta_k^{(0)}\right)\frac{\partial}{\partial\theta_k}\right\}$$
$$f_i(\underline{\theta})\Big|_{\underline{\theta}=\underline{\theta}^{(0)}}. \tag{3.1.1}$$

If $\underline{\theta}$ is a solution to (3.1.1) then the left hand side of (3.1.1) will be zero. At each iteration the new value $\underline{\theta}^{(j+1)}$ can be found from $\underline{\theta}^{(j)}$ by adding the vector $\underline{\delta}$ to it where

$$\underline{\delta} = -A^{-1}\underline{D}, \tag{3.1.2}$$

In (3.1.2) A^{-1} and $\underline{D}$ are evaluated at $\underline{\theta} = \underline{\theta}^{(j)}$ and are given by

$$A = \begin{bmatrix} \frac{\partial f_1(\underline{\theta})}{\partial \theta_1} & \frac{\partial f_1(\underline{\theta})}{\partial \theta_2} & \cdots & \frac{\partial f_1(\underline{\theta})}{\partial \theta_k} \\ \frac{\partial f_2(\underline{\theta})}{\partial \theta_1} & \frac{\partial f_2(\underline{\theta})}{\partial \theta_2} & \cdots & \frac{\partial f_2(\underline{\theta})}{\partial \theta_k} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_k(\underline{\theta})}{\partial \theta_1} & \frac{\partial f_k(\underline{\theta})}{\partial \theta_2} & \cdots & \frac{\partial f_k(\underline{\theta})}{\partial \theta_k} \end{bmatrix},$$

and

$$D = \begin{bmatrix} f_1(\underline{\theta}) \\ f_2(\underline{\theta}) \\ \cdots \\ f_k(\underline{\theta}) \end{bmatrix}.$$

The iteration process will be continued until $\underline{\delta}$ becomes small. In the case of ML equations $f_i(\underline{\theta})$ is replaced by $\frac{\partial \ln L}{\partial \theta_i}$, i=1, 2,..., k. In this case we have

$$A = \begin{bmatrix} \frac{\partial^2 \ln L}{\partial \theta_1^2} & \frac{\partial^2 \ln L}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial^2 \ln L}{\partial \theta_1 \partial \theta_k} \\ \cdots & \frac{\partial^2 \ln L}{\partial \theta_2^2} & \cdots & \frac{\partial^2 \ln L}{\partial \theta_2 \partial \theta_k} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \frac{\partial^2 \ln L}{\partial \theta_k^2} \end{bmatrix}, \tag{3.1.3}$$

and

$$D = \begin{bmatrix} \frac{\partial \ln L}{\partial \theta_1} \\ \frac{\partial \ln L}{\partial \theta_2} \\ \cdots \\ \frac{\partial \ln L}{\partial \theta_k} \end{bmatrix}. \tag{3.1.4}$$

2. Method of Scoring: This method is applicable only to ML equations. The quantity $\frac{\partial \ln L}{\partial \theta_i}$ is called the efficient score for θ_i. According to this definition the ML

estimator is the value of θ_i for which the efficient score becomes zero. We can expand $\frac{\partial \ln L}{\partial \theta_i}$ around a trial solution as before, but Rao (1973, p. 366) suggests that in order to stabilize the value of $\underline{\delta}$, it is better to replace the matrix A in (3.1.3) by its expected value evaluated at $\underline{\theta}^{(0)}$. The expected value of $-A$ is the information matrix at $\underline{\theta} = \underline{\theta}^{(0)}$

$$\Gamma\left(\underline{\theta}^{(0)}\right) = \left\{ E\left[-\frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j} \right] \right\} \Bigg|_{\underline{\theta}=\underline{\theta}^{(0)}} . \tag{3.1.5}$$

Hence

$$\underline{\delta} = \Gamma^{-1}\left(\underline{\theta}^{(0)}\right)\underline{D}. \tag{3.1.6}$$

Here also the new value $\underline{\theta}^{(j+1)}$ is obtained by adding $\underline{\delta}$ to the value $\underline{\theta}^{(j)}$ obtained at jth iteration.

Following are the variances and covariances of the sample moments which are needed to obtain the variances and covariances of different types of estimators. To order n^{-1}

$$\operatorname{Var}\left(\overline{X}_1\right) = \frac{\mu'_{2,0} - \mu'^2_{1,0}}{n}$$

$$\operatorname{Var}\left(\overline{X}_2\right) = \frac{\mu'_{0,2} - \mu'^2_{0,1}}{n}$$

$$\operatorname{Var}\left(m_{1,1}\right) = \frac{\mu_{2,2} - \mu^2_{1,1}}{n}$$

Also

$$\operatorname{Cov}\left(\overline{X}_1, \overline{X}_2\right) = \frac{\mu'_{1,1} - \mu'_{1,0}\mu'_{0,1}}{n}$$

$$\operatorname{Cov}\left(\overline{X}_1, m_{1,1}\right) = \frac{\mu_{2,1}}{n}$$

$$\operatorname{Cov}\left(\overline{X}_2, m_{1,1}\right) = \frac{\mu_{1,2}}{n}$$

$$\operatorname{Var}\left(\frac{n_{00}}{n}\right) = \frac{1}{n} f(0,0)[1 - f(0,0)]$$

$$\mathrm{Cov}\left(\overline{X}_1, \frac{n_{00}}{n}\right) = \frac{-f(0,0)\mu'_{1,0}}{n}$$

$$\mathrm{Cov}\left(\overline{X}_2, \frac{n_{00}}{n}\right) = \frac{-f(0,0)\mu'_{0,1}}{n}.$$

The covariance matrix for the sample moments to be used to obtain the covariance matrix of method of moments estimators is as follows

$$\Sigma_m = \begin{bmatrix} \mathrm{Var}(\overline{X}_1) & \mathrm{Cov}(\overline{X}_1, \overline{X}_2) & \mathrm{Cov}(\overline{X}_1, m_{1,1}) \\ \mathrm{Cov}(\overline{X}_1, \overline{X}_2) & \mathrm{Var}(\overline{X}_2) & \mathrm{Cov}(\overline{X}_2, m_{1,1}) \\ \mathrm{Cov}(\overline{X}_1, m_{1,1}) & \mathrm{Cov}(\overline{X}_2, m_{1,1}) & \mathrm{Var}(m_{1,1}) \end{bmatrix}.$$

Following is the covariance matrix needed to obtain the covariance matrix of method of zero-zero cell frequency estimators.

$$\Sigma_z = \begin{bmatrix} \mathrm{Var}(\overline{X}_1) & \mathrm{Cov}(\overline{X}_1, \overline{X}_2) & \mathrm{Cov}\left(\overline{X}_1, \frac{n_{00}}{n}\right) \\ \mathrm{Cov}(\overline{X}_1, \overline{X}_2) & \mathrm{Var}(\overline{X}_2) & \mathrm{Cov}\left(\overline{X}_2, \frac{n_{00}}{n}\right) \\ \mathrm{Cov}\left(\overline{X}_1, \frac{n_{00}}{n}\right) & \mathrm{Cov}\left(\overline{X}_2, \frac{n_{00}}{n}\right) & \mathrm{Var}\left(\frac{n_{00}}{n}\right) \end{bmatrix}.$$

3.2 Cases 1 and 2

Consider the distributions described in sections (2.2) and (2.3). As the two distributions are equivalent to the bivariate Poisson distribution with the same set of parameters, estimation of the parameters of any one will lead to those of the other.

Let (x_{1i}, x_{2i}) , i = 1, 2, 3,..., n, be a random sample of size n from the distribution with pgf

$$\pi(t_1, t_2) = \exp\left[\lambda(q_1 q_2 - 1) + p_1 q_2 t_1 + p_2 q_1 t_2 + p_1 p_2 t_1 t_2\right].$$

The frequency of the pair (r,s) is denoted by n_{rs} for r = 0, 1, 2,..., s = 0, 1, 2,.... Also $n = \sum_{r,s} n_{rs}$. Define

$$\overline{x}_1 = \frac{1}{n}\sum_{r,s} m_{rs},$$

$$\bar{x}_2 = \frac{1}{n}\sum_{r,s} s n_{rs},$$

and

$$m_{1,1} = \frac{1}{n}\sum_{r,s} rsn_{rs} - \bar{x}_1\bar{x}_2.$$

Method of moments

Since $E(X_1) = \lambda p_1$, $E(X_2) = \lambda p_2$ and $Cov(X_1, X_2) = \lambda p_1 p_2$, the moment equations are $\bar{x}_1 = \tilde{\lambda}\tilde{p}_1$, $\bar{x}_2 = \tilde{\lambda}\tilde{p}_2$ and $m_{1,1} = \tilde{\lambda}\tilde{p}_1\tilde{p}_2$. From these equations the method of moments estimators are

$$\left.\begin{aligned} \tilde{\lambda} &= \frac{\bar{x}_1\bar{x}_2}{m_{1,1}} \\ \tilde{p}_1 &= \frac{m_{1,1}}{\bar{x}_2} \\ \tilde{p}_2 &= \frac{m_{1,1}}{\bar{x}_1} \end{aligned}\right\}. \qquad (3.2.1)$$

The variance matrix of the estimators is obtained by using the relationship

$$\Sigma_{MM} = T^{-1}\Sigma_m\left(T^{-1}\right)',$$

where

$$T = \begin{bmatrix} \frac{\partial u_1}{\partial \lambda} & \frac{\partial u_1}{\partial p_1} & \frac{\partial u_1}{\partial p_2} \\ \frac{\partial u_2}{\partial \lambda} & \frac{\partial u_2}{\partial p_1} & \frac{\partial u_2}{\partial p_2} \\ \frac{\partial u_3}{\partial \lambda} & \frac{\partial u_3}{\partial p_1} & \frac{\partial u_3}{\partial p_2} \end{bmatrix}$$

$$= \begin{bmatrix} p_1 & \lambda & 0 \\ p_2 & 0 & \lambda \\ p_1 p_2 & p_2\lambda & p_1\lambda \end{bmatrix}, \qquad (3.2.2)$$

with $u_1 = p_1\lambda$, $u_2 = p_2\lambda$ and $u_3 = p_1 p_2 \lambda$. Also the Σ_m matrix is given by

$$\sum_m = \frac{1}{n}\begin{bmatrix} p_1\lambda & p_1p_2\lambda & p_1p_2\lambda \\ p_1p_2\lambda & p_2\lambda & p_1p_2\lambda \\ p_1p_2\lambda & p_1p_2\lambda & p_1p_2\lambda(p_1p_2\lambda+\lambda+1) \end{bmatrix}. \tag{3.2.3}$$

Zero-zero cell frequency method

In this method the (0,0) cell relative frequency is equated to the corresponding population probability. In addition, in the case of three parameter estimation two marginal means are equated to the corresponding population marginal means. The three resulting equations are solved for the parameters in terms of observed data. In the case of bivariate Poisson distribution three estimating equations are

$$\left.\begin{aligned} \bar{x}_1 &= \lambda' p_1' \\ \bar{x}_2 &= \lambda' p_2' \\ \ln\left(\frac{n_{00}}{n}\right) &= \lambda'(p_1'p_2' - p_1' - p_2') \end{aligned}\right\}. \tag{3.2.4}$$

From the equations in (3.2.4), the zero-zero cell frequency estimators are

$$\left.\begin{aligned} \lambda' &= \frac{\bar{x}_1\bar{x}_2}{\left[\ln\left(\frac{n_{00}}{n}\right)+\bar{x}_1+\bar{x}_2\right]} \\ p_1' &= \frac{\left[\ln\left(\frac{n_{00}}{n}\right)+\bar{x}_1+\bar{x}_2\right]}{\bar{x}_2} \\ p_2' &= \frac{\left[\ln\left(\frac{n_{00}}{n}\right)+\bar{x}_1+\bar{x}_2\right]}{\bar{x}_1} \end{aligned}\right\}. \tag{3.2.5}$$

The variance covariance matrix of zero-zero cell frequency method estimators is obtained as below

$$\sum_{ZZ} = T^{-1}\sum_z\left(T^{-1}\right)'.$$

Recalling that $u_1 = p_1\lambda$, $u_2 = p_2\lambda$ and $u_3 = \exp[\lambda(1-p_1)(1-p_2)-\lambda]$, we see that

$$T=\begin{bmatrix} p_1 & \lambda & 0 \\ p_2 & 0 & \lambda \\ e^{(\lambda q_1 q_2-\lambda)}(q_1q_2-1) & -\lambda q_2 e^{(\lambda q_1 q_2-\lambda)} & -\lambda q_1 e^{(\lambda q_1 q_2-\lambda)} \end{bmatrix}.$$

The covariance matrix $\sum_z$ is

$$\sum_z=\frac{1}{n}\begin{bmatrix} p_1\lambda & p_1p_2\lambda & e^{\lambda(q_1q_2-1)}p_1\lambda \\ p_1p_2\lambda & p_2\lambda & e^{\lambda(q_1q_2-1)}p_2\lambda \\ e^{\lambda(q_1q_2-1)}p_1\lambda & e^{\lambda(q_1q_2-1)}p_2\lambda & e^{\lambda(q_1q_2-1)}\left(1-e^{\lambda(q_1q_2-1)}\right) \end{bmatrix}. \quad (3.2.6)$$

Method of maximum likelihood

Consider the pf defined in (2.2.1)

$$f(r,s)=e^{\lambda(p_1p_2-p_1-p_2)}\sum_{i=0}^{\min(r,s)}\frac{\lambda^{r+s-i}p_1^r p_2^s(1-p_1)^{s-i}(1-p_2)^{r-i}}{(r-i)!(s-i)!i!}. \quad (3.2.7)$$

Applying the technique described in Kocherlakota and Kocherlakota (1992, p. 45-47), we can write the likelihood function (L) as

$$L\propto\prod_{r,s}f(r,s)^{n_{rs}},$$

which can be written as

$$\ln L=C+\sum_{r,s}n_{rs}\ln f(r,s),$$

or

$$\ln L=C+\sum_{r,s}n_{rs}\ln\left[e^{\lambda(p_1p_2-p_1-p_2)}\sum_{i=0}^{\min(r,s)}\frac{\lambda^{r+s-i}p_1^r p_2^s(1-p_1)^{s-i}(1-p_2)^{r-i}}{(r-i)!(s-i)!i!}\right], \quad (3.2.8)$$

where C is a constant not involving the parameters. Let $D=(r-i)!(s-i)!i!$, so that (3.2.8) can be written as

$$\ln L=C+\sum_{r,s}n_{rs}\lambda(p_1p_2-p_1-p_2)$$
$$+\sum_{r,s}n_{rs}\ln\left[\sum_{i=0}^{\min(r,s)}\frac{\lambda^{r+s-i}p_1^r p_2^s(1-p_1)^{s-i}(1-p_2)^{r-i}}{D}\right]. \quad (3.2.9)$$

Differentiating (3.2.9) with respect to λ

$$\frac{\partial \ln L}{\partial \lambda} = \sum_{r,s} n_{rs}(p_1p_2 - p_1 - p_2) + \sum_{r,s} n_{rs}\frac{1}{A}\sum_{i=0}^{\min(r,s)}\frac{(r+s-i)\lambda^{r+s-i-1}p_1^r p_2^s q_1^{s-i} q_2^{r-i}}{D}$$

$$= np_1p_2 - np_1 - np_2 + \sum_{r,s} n_{rs}\frac{1}{A}\left[\frac{(r+s)}{\lambda}A - B\right],$$

or

$$\frac{\partial \ln L}{\partial \lambda} = np_1p_2 - np_1 - np_2 + \frac{n\bar{x}}{\lambda} + \frac{n\bar{y}}{\lambda} - \frac{1}{\lambda}\sum_{r,s} n_{rs}\frac{B}{A}, \qquad (3.2.10)$$

where

$$A = \sum_{i=0}^{\min(r,s)}\frac{\lambda^{r+s-i}p_1^r p_2^s (1-p_1)^{s-i}(1-p_2)^{r-i}}{D},$$

and

$$B = \sum_{i=0}^{\min(r,s)}\frac{i\lambda^{r+s-i}p_1^r p_2^s (1-p_1)^{s-i}(1-p_2)^{r-i}}{D}.$$

Differentiating with respect to p_1

$$\frac{\partial \ln L}{\partial p_1} = \sum_{r,s} n_{rs}\lambda(p_2 - 1) + \sum_{r,s} n_{rs}\frac{1}{A}\left[\sum_{i=0}^{\min(r,s)}\frac{\lambda^{r+s-i}p_2^s q_2^{r-i}}{D}\left\{rp_1^{r-1}q_1^{s-i} - p_1^r(s-i)q_1^{s-i-1}\right\}\right],$$

$$= n\lambda p_2 - n\lambda + \frac{n\bar{x}}{p_1} - \frac{n\bar{y}}{(1-p_1)} + \frac{1}{(1-p_1)}\sum_{r,s} n_{rs}\frac{B}{A}. \qquad (3.2.11)$$

Similarly, differentiation of logarithm of likelihood function with respect to p_2 yields

$$\frac{\partial \ln L}{\partial p_2} = n\lambda p_1 - n\lambda + \frac{n\bar{y}}{p_2} - \frac{n\bar{x}}{(1-p_2)} + \frac{1}{(1-p_2)}\sum_{r,s} n_{rs}\frac{B}{A}. \qquad (3.2.12)$$

Solving equations (3.2.10), (3.2.11) and (3.2.12) simultaneously for λ, p_1 and p_2, we get three equations

$$\hat{p}_1 = \frac{\bar{x}_1}{\hat{\lambda}}$$

$$\hat{p}_2 = \frac{\bar{x}_2}{\hat{\lambda}}$$

and

$$\hat{\lambda}Q = \bar{x}_1\bar{x}_2, \qquad (3.2.13)$$

where

$$Q = \frac{1}{n}\sum_{r,s} n_{rs}\frac{B^*}{A^*},$$

with

$$A^* = \sum_{i=0}^{\min(r,s)} \frac{\hat{\lambda}^{r+s-i}\left(\frac{\bar{x}_1}{\hat{\lambda}}\right)^r\left(\frac{\bar{x}_2}{\hat{\lambda}}\right)\left(1-\frac{\bar{x}_1}{\hat{\lambda}}\right)^{s-i}\left(1-\frac{\bar{x}_2}{\hat{\lambda}}\right)^{r-i}}{D},$$

and

$$B^* = \sum_{i=0}^{\min(r,s)} \frac{i\hat{\lambda}^{r+s-i}\left(\frac{\bar{x}_1}{\hat{\lambda}}\right)^r\left(\frac{\bar{x}_2}{\hat{\lambda}}\right)^s\left(1-\frac{\bar{x}_1}{\hat{\lambda}}\right)^{s-i}\left(1-\frac{\bar{x}_2}{\hat{\lambda}}\right)^{r-i}}{D}.$$

We can solve (3.2.13) iteratively for $\hat{\lambda}$. Alternatively, the ML estimators can be found using the method of scoring described in section (3.1). For this purpose the following second order derivatives of the probability function are required.

$$\frac{\partial^2 f(r,s)}{\partial\lambda^2} = f(r,s)\left\{(q_1q_2-1)^2 - \frac{(r+s)}{\lambda^2} + \frac{(r+s)^2}{\lambda^2} + 2\frac{(r+s)}{\lambda}(q_1q_2-1)\right\}$$
$$+\frac{f(0,0)B}{\lambda}\left\{(q_1q_2-1) - 2\frac{(r+s)}{\lambda} + \frac{1}{\lambda} - (q_1q_2-1)\right\} + \frac{f(0,0)}{\lambda}C$$

where

$$f(r,s) = e^{\lambda(q_1q_2-1)}\sum_{i=0}^{\min(r,s)}\frac{p_1^r p_2^s q_1^{s-i} q_2^{r-i}\lambda^{r+s-i}}{D},$$

$$f(0,0) = e^{\lambda(q_1q_1-1)},$$

$$B = \sum_{i=0}^{\min(r,s)}\frac{i p_1^r p_2^s q_1^{s-i} q_2^{r-i}\lambda^{r+s-i}}{D},$$

$$C = \sum_{i=0}^{\min(r,s)}\frac{i^2 p_1^r p_2^s q_1^{s-i} q_2^{r-i}\lambda^{r+s-i}}{D},$$

$$D = (r-i)!(s-i)!i!.$$

$$\frac{\partial^2 f(r,s)}{\partial p_1^2} = \left\{\left(-\lambda q_2 + \frac{r}{p_1} - \frac{s}{q_1}\right)^2 - \left(\frac{r}{p_1^2} + \frac{s}{q_1^2}\right)\right\} f(r,s) + \frac{Bf(0,0)}{q_1}\left\{\left(-\lambda q_2 + \frac{r}{p_1} - \frac{s}{q_1}\right)\right.$$

$$\left. + \frac{1}{q_1} + \frac{r}{p_1} - \frac{s}{q_1} - \lambda q_2\right\} + \frac{Cf(0,0)}{q_1^2},$$

$$\frac{\partial^2 f(r,s)}{\partial p_2^2} = \left\{\left(-\lambda q_1 + \frac{s}{p_2} - \frac{r}{q_2}\right)^2 - \left(\frac{s}{p_2^2} + \frac{r}{q_2^2}\right)\right\} f(r,s) + \frac{Bf(0,0)}{q_2}\left\{\left(-\lambda q_1 + \frac{s}{p_2} - \frac{r}{q_2}\right)\right.$$

$$\left. + \frac{1}{q_2} + \frac{s}{p_2} - \frac{r}{q_2} - \lambda q_1\right\} + \frac{Cf(0,0)}{q_2^2}.$$

$$\frac{\partial^2 f(r,s)}{\partial \lambda \partial p_1} = f(r,s)\left\{-q_2 + \left(q_1 q_2 - 1 + \frac{(r+s)}{\lambda}\right)\left(-\lambda q_2 + \frac{r}{p_1} - \frac{s}{q_1}\right)\right\}$$

$$+B\left\{\frac{1}{q_1}\left(q_1 q_2 - 1 + \frac{(r+s)}{\lambda}\right) + q_2 f(0,0) - \frac{rf(0,0)}{\lambda p_1} + \frac{sf(0,0)}{\lambda q_1}\right\}$$

$$-\frac{f(0,0)C}{\lambda q_1},$$

$$\frac{\partial^2 f(r,s)}{\partial \lambda \partial p_2} = f(r,s)\left\{-q_1 + \left(q_1 q_2 - 1 + \frac{(r+s)}{\lambda}\right)\left(-\lambda q_1 + \frac{s}{p_2} - \frac{r}{q_2}\right)\right\}$$

$$+B\left\{\frac{1}{q_2}\left(q_1 q_2 - 1 + \frac{(r+s)}{\lambda}\right) + q_1 f(0,0) - \frac{sf(0,0)}{\lambda p_2} + \frac{rf(0,0)}{\lambda q_2}\right\}$$

$$-\frac{f(0,0)C}{\lambda q_2},$$

$$\frac{\partial f(r,s)}{\partial p_1 \partial p_2} = f(r,s)\left\{\lambda + \left(-\lambda q_2 + \frac{r}{p_1} - \frac{s}{q_1}\right)\left(-\lambda q_1 + \frac{s}{p_2} - \frac{r}{q_2}\right)\right\}$$

$$+f(0,0)B\left\{-\lambda + \frac{1}{q_2}\left(-\lambda q_2 + \frac{r}{p_1} - \frac{s}{q_1}\right) + \frac{s}{q_1 p_2} - \frac{r}{q_1 q_2}\right\}$$

$$+\frac{f(0,0)C}{q_1 q_2}.$$

The numerical evaluation of these derivatives can be done using S-plus programs dlbvp1, dp1bvp1 and dp2bvp1.

3.3 Case 3

Poisson-double binomial distribution has five parameters. In order to estimate the parameters it is a common practice to assume that n_1 and n_2 are known. Under this assumption the estimation procedure becomes less complex, and number of parameters to be estimated reduces to three. Various methods of estimation are employed, and their efficiencies relative to the method of maximum likelihood are obtained.

Let (x_{1i}, x_{2i}), i = 1, 2, 3,..., n, be a random sample of size n from the distribution having pgf

$$\pi(t_1,t_2) = \exp\left[\lambda\left\{(q_1 + p_1t_1)^{n_1}(q_2 + p_2t_2)^{n_2} - 1\right\}\right].$$

The frequency of the pair (r,s) is denoted by n_{rs} for r = 0, 1, 2,..., s = 0, 1, 2,.... Also $n = \sum_{r,s} n_{rs}$.

Method of moments

Since $E(X_1) = \lambda n_1 p_1$, $E(X_2) = \lambda n_2 p_2$, and $Cov(X_1, X_2) = \lambda n_1 n_2 p_1 p_2$, the moments estimators are obtained by solving the three equations

$$\left.\begin{aligned} \bar{x}_1 &= \tilde{\lambda} n_1 \tilde{p}_1 \\ \bar{x}_2 &= \tilde{\lambda} n_2 \tilde{p}_2 \\ m_{1,1} &= \tilde{\lambda} n_1 n_2 \tilde{p}_1 \tilde{p}_2 \end{aligned}\right\}. \qquad (3.3.1)$$

The moments estimators are given by

$$\left.\begin{aligned} \tilde{p}_1 &= \frac{m_{1,1}}{n_1 \bar{x}_2} \\ \tilde{p}_2 &= \frac{m_{1,1}}{n_2 \bar{x}_1} \\ \tilde{\lambda} &= \frac{\bar{x}_1 \bar{x}_2}{m_{1,1}} \end{aligned}\right\}. \qquad (3.3.2)$$

The variance matrix of the estimators is obtained by using the relationship

$$\Sigma_{MM} = T^{-1} \Sigma_m \left(T^{-1}\right)',$$

where

$$T = \begin{bmatrix} n_1p_1 & n_1\lambda & 0 \\ n_2p_2 & 0 & n_2\lambda \\ n_1n_2p_1p_2 & n_1n_2p_2\lambda & n_1n_2p_1\lambda \end{bmatrix}, \tag{3.3.3}$$

with $u_1 = n_1p_1\lambda$, $u_2 = n_2p_2\lambda$ and $u_3 = n_1n_2p_1p_2\lambda$. Also the $\sum_m$ matrix is given by

$$\sum_m = \frac{1}{n}\begin{bmatrix} \lambda n_1p_1(n_1p_1+q_1) & \lambda n_1n_2p_1p_2 & \lambda n_1n_2p_1p_2(n_1p_1+q_1) \\ \cdots & \lambda n_2p_2(n_2p_2+q_2) & \lambda n_1n_2p_1p_2(n_2p_2+q_2) \\ \cdots & \cdots & n\mathrm{Var}(m_{1,1}) \end{bmatrix}, \tag{3.3.4}$$

and

$$\begin{aligned} \mathrm{Var}(m_{1,1}) = {} & \frac{1}{n}\lambda^2\big[n_1n_2p_1^2p_2^2\{2n_1n_2 - n_1 - n_2 + 1\} + n_1n_2p_1p_2 \\ & \{1 + n_1p_1 + n_2p_2 - p_1 - p_2\}\big] + \frac{1}{n}\lambda\big[n_1^{[2]}n_2^{[2]}p_1^2p_2^2 + \\ & n_1^{[2]}n_2p_1^2p_2 + n_1n_2^{[2]}p_1p_2^2 + n_1n_2p_1p_2\big], \end{aligned}$$

where $a^{[r]} = a(a-1)(a-2)\ \ldots\ (a-r+1)$.

Zero-zero cell frequency method

The $(0,0)$ cell relative frequency along with the two marginal means can also be used to estimate the parameters rather quickly as compared to maximum likelihood method. Three estimating equations are

$$\left.\begin{aligned} \bar{x}_1 &= \lambda' n_1 p_1' \\ \bar{x}_2 &= \lambda' n_2 p_2' \\ \log\left(\frac{n_{00}}{n}\right) &= \lambda'\left[\left(1 - \frac{\bar{x}_1}{n_1\lambda'}\right)^{n_1}\left(1 - \frac{\bar{x}_2}{n_2\lambda'}\right)^{n_2} - 1\right] \end{aligned}\right\}, \tag{3.3.5}$$

yielding $p_1' = \dfrac{\bar{x}_1}{n_1\lambda'}$ and $p_2' = \dfrac{\bar{x}_2}{n_2\lambda'}$.

The variance matrix of the estimators is found in a similar manner to moments estimators. That is the variance covariance matrix of estimators is given by

$$\sum_{ZZ}=T^{-1}\sum_{z}\left(T^{-1}\right)',$$

with $u_1 = n_1 p_1 \lambda$, $u_2 = n_2 p_2 \lambda$ and $u_3 = \exp\left[\lambda(1-p_1)^{n_1}(1-p_2)^{n_2} - \lambda\right]$. Thus the matrix T is

$$T=\begin{bmatrix} n_1 p_1 & n_1\lambda & 0 \\ n_2 p_2 & 0 & n_2\lambda \\ f(0,0)\left(q_1^{n_1}q_2^{n_2}-1\right) & -\lambda n_1 q_1^{n_1-1}q_2^{n2} f(0,0) & -\lambda n_2 q_2^{n_2-1} q_1^{n_1} f(0,0) \end{bmatrix}$$

where $f(0,0)=\exp\left[\lambda\left(q_1^{n_1}q_2^{n_2}-1\right)\right]$. The $\sum_z$ matrix is of the following form,

$$\sum_z=\frac{1}{n}\begin{bmatrix} \lambda n_1 p_1(n_1 p_1 + q_1) & \lambda n_1 n_2 p_1 p_2 & -f(0,0)\lambda n_1 p_1 \\ \cdots & \lambda n_2 p_2(n_2 p_2 + q_2) & -f(0,0)\lambda n_2 p_2 \\ \cdots & \cdots & f(0,0)(1-f(0,0)) \end{bmatrix}. \quad (3.3.6)$$

Method of maximum likelihood

Method of maximum likelihood usually does not provide the estimators of the parameters in a closed form, and one has to use iterative techniques to get a solution. The technique described by Kocherlakota and Kocherlakota (1992) in section (2.1.4) is being used to get maximum likelihood estimates and variance covariance matrix of estimators of the parameters of the Poisson-double binomial distribution.

In order to get the maximum likelihood estimators of the parameters of Poisson-double binomial distribution and their variances and covariances, differentiation of probability function with respect to parameters is needed. There are two different forms which give rise to the same results.

Form 1 We know that

$$\pi(t_1,t_2)=\sum_{x_1}\sum_{x_2} t_1^{x_1} t_2^{x_2} f(x_1,x_2), \quad (3.3.7)$$

differentiating both sides of (3.3.7) with respect to λ yields

$$\frac{\partial\pi(t_1,t_2)}{\partial\lambda}=\sum_{x_1}\sum_{x_2} t_1^{x_1} t_2^{x_2}\frac{\partial}{\partial\lambda} f(x_1,x_2). \quad (3.3.8)$$

Consider left hand side of (3.3.8)

$$\frac{\partial \pi(t_1,t_2)}{\partial \lambda} = \pi(t_1,t_2)\left[(q_1+p_1t_1)^{n_1}(q_2+p_2t_2)^{n_2}-1\right]$$

which on using binomial expansion and (3.3.7) becomes

$$= \sum_{x_1}\sum_{x_2} t_1^{x_1}t_2^{x_2} f(x_1,x_2)\left[\sum_{r=0}^{n_1}\binom{n_1}{r}q_1^{n_1-r}p_1^r t_1^r \sum_{s=0}^{n_2}\binom{n_2}{s}q_2^{n_2-s}p_2^s t_2^s - 1\right]$$

$$\begin{aligned} &= \sum_{x_1}\sum_{x_2}\sum_{r}\sum_{s}\binom{n_1}{r}\binom{n_2}{s} f(x_1,x_2) t_1^{x_1+r} t_2^{x_2+s} q_1^{n_1-r} q_2^{n_2-s} p_1^r p_2^s \\ &-\sum_{x_1}\sum_{x_2} t_1^{x_1}t_2^{x_2} f(x_1,x_2), \end{aligned} \tag{3.3.9}$$

equating (3.3.8) and (3.3.9) and comparing the powers $t_1^j t_2^k$, we have in the first summation on the right hand side of (3.3.9) $x_1+r=j$ or $x_1=j-r$, $x_2+s=k$ or $x_2=k-s$. Hence $t_1^j t_2^k$ has the coefficient in the first summation

$$\sum_{r}\sum_{s}\binom{n_1}{r}\binom{n_2}{s} q_1^{n_1-r} q_2^{n_2-s} p_1^r p_2^s f(j-r,k-s).$$

In the second summation on the right hand side of (3.3.9) The coefficient is f(j,k). Hence

$$\begin{aligned} \frac{\partial f(j,k)}{\partial \lambda} &= \sum_{r}\sum_{s}\binom{n_1}{r}\binom{n_2}{s} p_1^r q_1^{n_1-r} p_2^s q_2^{n_2-s} f(j-r,k-s) - f(j,k) \\ &= \sum_{r}\sum_{s} b(n_1,r)b(n_2,s)f(j-r,k-s) - f(j,k), \end{aligned}$$

where if $j<n_1, k<n_2$ then r = 0, 1, 2,..., j and s = 0, 1, 2,..., k. If $j \geq n_1$ and/or $k \geq n_2$ then $r=0,1,\ldots,n_1$ and $s=0,1,\ldots,n_2$. For example,

$$\begin{aligned} \frac{\partial f(0,0)}{\partial \lambda} &= \left[b(n_1,0)b(n_2,0)-1\right]f(0,0) \\ &= \left(q_1^{n_1}q_2^{n_2}-1\right)f(0,0). \end{aligned} \tag{3.3.10}$$

$$\begin{aligned} \frac{\partial f(1,1)}{\partial \lambda} &= b(n_1,0)b(n_2,0)f(1,1) + b(n_1,0)b(n_2,1)f(1,0) + \\ & b(n_1,1)b(n_2,0)f(0,1) + b(n_1,1)b(n_2,1)f(0,0) - f(1,1) \end{aligned}$$

$$= f(1,1)\left[b(n_1,0)b(n_2,0)+\frac{1}{\lambda}-1\right]+b(n_1,1)b(n_2,0)f(0,1).$$

Differentiating both sides of (3.3.7) with respect to p_1, yields

$$\frac{\partial\pi(t_1,t_2)}{\partial p_1}=\sum_{x_1}\sum_{x_2}t_1^{x_1}t_2^{x_2}\frac{\partial}{\partial p_1}f(x_1,x_2). \qquad (3.3.11)$$

Consider the left hand side of (3.3.11)

$$\frac{\partial\pi(t_1,t_2)}{\partial p_1}=\pi(t_1,t_2)\left[\lambda(q_2+p_2t_2)^{n_2}\frac{\partial}{\partial p_1}(q_1+p_1t_1)^{n_1}\right].$$

Since

$$\frac{\partial}{\partial p_1}(q_1+p_1t_1)^{n_1}=n_1(t_1-1)(q_1+p_1t_1)^{n_1-1}$$

$$=n_1(t_1-1)\sum_{r=0}^{n_1-1}\binom{n_1-1}{r}q_1^{n_1-1-r}p_1^rt_1^r$$

$$=n_1\sum_{r=0}^{n_1-1}\binom{n_1-1}{r}q_1^{n_1-1-r}p_1^rt_1^{r+1}-n_1\sum_{r=0}^{n_1-1}\binom{n_1-1}{r}q_1^{n_1-1-r}p_1^rt_1^r.$$

Hence the left hand side of (3.3.11) becomes

$$\frac{\partial\pi(t_1,t_2)}{\partial p_1}=n_1\lambda\sum_{x_1}\sum_{x_2}t_1^{x_1}t_2^{x_2}f(x_1,x_2)\sum_{s=0}^{n_2}\binom{n_2}{s}q_2^{n_2-s}p_2^st_2^s$$

$$\left[\sum_{r=0}^{n_1-1}\binom{n_1-1}{r}q_1^{n_1-1-r}p_1^rt_1^{r+1}-\sum_{r=0}^{n_1-1}\binom{n_1-1}{r}q_1^{n_1-1-r}p_1^rt_1^r\right]$$

$$=n_1\lambda\sum_{x_1}\sum_{x_2}\sum_{r=0}^{n_1-1}\sum_{s=0}^{n_2}f(x_1,x_2)\binom{n_1-1}{r}q_1^{n_1-1-r}p_1^r\binom{n_2}{s}q_2^{n_2-s}p_2^st_1^{x_1+r+1}t_2^{x_2+s}$$

$$-n_1\lambda\sum_{x_1}\sum_{x_2}\sum_{r=0}^{n_1-1}\sum_{s=0}^{n_2}f(x_1,x_2)\binom{n_1-1}{r}q_1^{n_1-1-r}p_1^r\binom{n_2}{s}q_2^{n_2-s}p_2^st_1^{x_1+r}t_2^{x_2+s}.$$

In the first summation for finding coefficient of $t_1^jt_2^k$, we have $x_1+r+1=j\Rightarrow x_1=j-r-1$ and $x_2+s=k\Rightarrow x_2=k-s$, where r = 0, 1, ..., j - 1, s = 0, 1, ..., k. In the second summation $x_1+r=j\Rightarrow x_1=j-r$ and $x_2+s=k\Rightarrow x_2=k-s$, where r = 0, 1, ..., j and s = 0, 1, ..., k. Comparing coefficients of $t_1^j t_2^k$ on both sides of (3.3.11), we have for $j \geq 1, k \geq 0$

$$\frac{\partial f(j,k)}{\partial p_1} = n_1\lambda \sum_{r=0}^{j-1}\sum_{s=0}^{k}\binom{n_1-1}{r} q_1^{n_1-1-r} p_1^r \binom{n_2}{s} q_2^{n_2-s} p_2^s f(j-r-1,k-s)$$

$$-n_1\lambda \sum_{r=0}^{j}\sum_{s=0}^{k}\binom{n_1-1}{r} q_1^{n_1-1-r} p_1^r \binom{n_2}{s} q_2^{n_2-s} p_2^s f(j-r,k-s).$$

$$= n_1\lambda \sum_{r=0}^{j-1}\sum_{s=0}^{k} b(n_1-1,r)b(n_2,s)f(j-r-1,k-s)$$

$$-n_1\lambda \sum_{r=0}^{j}\sum_{s=0}^{k} b(n_1-1,r)b(n_2,s)f(j-r,k-s). \qquad (3.3.12)$$

For $j = 0, k \geq 1$

$$\frac{\partial f(0,k)}{\partial p_1} = -n_1\lambda \sum_{s=0}^{k} b(n_1-1,0)b(n_2,s)f(0,k-s),$$

and for $j \geq 1, k = 0$

$$\frac{\partial f(j,0)}{\partial p_1} = n_1\lambda \sum_{r=0}^{j-1} b(n_1-1,r)b(n_2,0)f(j-r-1,0)$$

$$-n_1\lambda \sum_{r=0}^{j} b(n_1-1,r)b(n_2,0)f(j-r,0).$$

Differentiating $f(0,0) = \exp\left[\lambda\left(q_1^{n_1} q_2^{n_2} - 1\right)\right]$ directly with respect to p_1, yields

$$\frac{\partial f(0,0)}{\partial p_1} = -n_1\lambda q_1^{n_1-1} q_2^{n_2} f(0,0)$$

$$= -n_1\lambda b(n_1-1,0)b(n_2,0)f(0,0).$$

For example

$$\frac{\partial f(1,0)}{\partial p_1} = n_1\lambda b(n_1-1,0)b(n_2,0)f(0,0) - n_1\lambda b(n_1-1,0)b(n_2,0)f(1,0)$$

$$-n_1\lambda b(n_1-1,1)b(n_2,0)f(0,0).$$

The formulae for p_2 are defined in a similar fashion, except the rolls of j and k are interchanged, i.e., differentiating both sides of (3.3.7) with respect to p_2, yields

$$\frac{\partial \pi(t_1,t_2)}{\partial p_2} = \sum_{x_1}\sum_{x_2} t_1^{x_1} t_2^{x_2} \frac{\partial}{\partial p_2} f(x_1,x_2). \qquad (3.3.13)$$

Consider the left hand side of (3.3.13)

$$\frac{\partial \pi(t_1,t_2)}{\partial p_2} = \pi(t_1,t_2)\left[\lambda(q_1+p_1t_1)^{n_1}\frac{\partial}{\partial p_2}(q_2+p_2t_2)^{n_2}\right].$$

Since

$$\frac{\partial}{\partial p_2}(q_2+p_2t_2)^{n_2} = n_2(t_2-1)(q_2+p_2t_2)^{n_2-1}$$

$$= n_2(t_2-1)\sum_{s=0}^{n_2-1}\binom{n_2-1}{s} q_2^{n_2-1-s}p_2^s t_2^s$$

$$= n_2\sum_{s=0}^{n_2-1}\binom{n_2-1}{s} q_2^{n_2-1-s}p_2^s t_2^{s+1} - n_2\sum_{s=0}^{n_2-1}\binom{n_2-1}{s} q_2^{n_2-1-s}p_2^s t_2^s.$$

Hence the left hand side of (3.3.13) becomes

$$\frac{\partial \pi(t_1,t_2)}{\partial p_2} = n_2\lambda\sum_{x_1}\sum_{x_2} t_1^{x_1}t_2^{x_2} f(x_1,x_2)\sum_{r=0}^{n_1}\binom{n_1}{r} q_1^{n_1-r}p_1^r t_1^r$$
$$\left[\sum_{s=0}^{n_2-1}\binom{n_2-1}{s} q_2^{n_2-1-s}p_2^s t_2^{s+1} - \sum_{s=0}^{n_2-1}\binom{n_2-1}{s} q_2^{n_2-1-s}p_2^s t_2^s\right]$$

$$= n_2\lambda\sum_{x_1}\sum_{x_2}\sum_{r=0}^{n_1}\sum_{s=0}^{n_2-1} f(x_1,x_2)\binom{n_1}{r} q_1^{n_1-r}p_1^r\binom{n_2-1}{s} q_2^{n_2-1-s}p_2^s t_1^{x_1+r}t_2^{x_2+s+1}$$
$$-n_2\lambda\sum_{x_1}\sum_{x_2}\sum_{r=0}^{n_1}\sum_{s=0}^{n_2-1} f(x_1,x_2)\binom{n_1}{r} q_1^{n_1-r}p_1^r\binom{n_2-1}{s} q_2^{n_2-1-s}p_2^s t_1^{x_1+r}t_2^{x_2+s}.$$

Hence in the first summation for finding coefficient of $t_1^j t_2^k$, we have $x_2+s+1=k \Rightarrow x_2=k-s-1$ and $x_1+r=j \Rightarrow x_1=j-r$, where r = 0, 1, ..., j, s = 0, 1, ..., k - 1. In the second summation $x_2+s=k \Rightarrow x_2=k-s$ and $x_1+r=j \Rightarrow x_1=j-r$, where r = 0, 1, ..., j and s = 0, 1, ..., k. Comparing coefficients of $t_1^j t_2^k$ on both sides of (3.3.13), we have for $j \ge 0, k \ge 1$

$$\frac{\partial f(j,k)}{\partial p_2} = n_2\lambda \sum_{r=0}^{j}\sum_{s=0}^{k-1}\binom{n_1}{r} q_1^{n_1-r} p_1^r \binom{n_2-1}{s} q_2^{n_2-1-s} p_2^s f(j-r,k-s-1)$$

$$-n_2\lambda \sum_{r=0}^{j}\sum_{s=0}^{k}\binom{n_1}{r} q_1^{n_1-r} p_1^r \binom{n_2-1}{s} q_2^{n_2-1-s} p_2^s f(j-r,k-s).$$

$$= n_2\lambda \sum_{r=0}^{j}\sum_{s=0}^{k-1} b(n_1,r)b(n_2-1,s)f(j-r,k-s-1)$$

$$-n_2\lambda \sum_{r=0}^{j}\sum_{s=0}^{k} b(n_1,r)b(n_2-1,s)f(j-r,k-s).$$

For $j \geq 1, k = 0$

$$\frac{\partial f(j,0)}{\partial p_2} = -n_2\lambda \sum_{r=0}^{j} b(n_1,r)b(n_2-1,0)f(j-r,0),$$

and for $j = 0, k \geq 1$

$$\frac{\partial f(0,k)}{\partial p_2} = n_2\lambda \sum_{s=0}^{k-1} b(n_1,0)b(n_2-1,s)f(0,k-s-1)$$

$$-n_2\lambda \sum_{s=0}^{k} b(n_1,0)b(n_2-1,s)f(0,k-s).$$

Differentiating $f(0,0) = \exp\left[\lambda\left(q_1^{n_1} q_2^{n_2} - 1\right)\right]$ directly with respect to p_2, yields

$$\frac{\partial f(0,0)}{\partial p_2} = -n_2\lambda q_2^{n_2-1} q_1^{n_1} f(0,0)$$

$$= -n_2\lambda b(n_1,0)b(n_2-1,0)f(0,0).$$

For example

$$\frac{\partial f(0,1)}{\partial p_2} = n_2\lambda b(n_1,0)b(n_2-1,0)f(0,0) - n_2\lambda b(n_1,0)b(n_2-1,0)f(0,1)$$

$$-n_2\lambda b(n_1,0)b(n_2-1,1)f(0,0).$$

Form 2

An alternative approach for finding the partial derivatives of the probability function with respect to the parameters is the use of the difference equations for probability function given in (2.4.21) and (2.4.22). In the following we will use the notation

$$b(n_i,r)=\binom{n_i}{r}p_i^r q_i^{n_i-r}, \quad i=1,2.$$

Differentiation with respect to λ, yields for $j\geq 1, k\geq 0$

$$\frac{\partial f(j,k)}{\partial\lambda}=\frac{f(j,k)}{\lambda}+\frac{\lambda}{j}\sum_{r=1}^{j}\sum_{s=0}^{k} rb(n_1,r)b(n_2,s)\frac{\partial f(j-r,k-s)}{\partial\lambda}.$$

If $j\geq 0, k\geq 1$, then

$$\frac{\partial f(j,k)}{\partial\lambda}=\frac{f(j,k)}{\lambda}+\frac{\lambda}{k}\sum_{r=0}^{j}\sum_{s=1}^{k} sb(n_1,r)b(n_2,s)\frac{\partial f(j-r,k-s)}{\partial\lambda}.$$

For example

$$\frac{\partial f(1,0)}{\partial\lambda}=\frac{f(1,0)}{\lambda}+\lambda b(n_1,1)b(n_2,0)\frac{\partial}{\partial\lambda}f(0,0)$$

$$=f(1,0)\left[\frac{1}{\lambda}+q_1^{n_1}q_2^{n_2}-1\right].$$

$$\frac{\partial f(1,1)}{\partial\lambda}=\frac{f(1,1)}{\lambda}+\lambda b(n_1,1)\left[b(n_2,0)\frac{\partial f(0,1)}{\partial\lambda}+b(n_2,1)\frac{\partial f(0,0)}{\partial\lambda}\right]$$

$$=\frac{f(1,1)}{\lambda}+\lambda b(n_1,1)\left[b(n_2,0)f(0,1)\left\{\frac{1}{\lambda}+q_1^{n_1}q_2^{n_2}-1\right\}\right.$$
$$\left.+b(n_2,1)f(0,0)\left\{q_1^{n_1}q_2^{n_2}-1\right\}\right]$$

$$=f(1,1)\left[\frac{1}{\lambda}+q_1^{n_1}q_2^{n_2}-1\right]+b(n_1,1)b(n_2,0)f(0,1).$$

Differentiating with respect to p_1 results in

$$\frac{\partial f(j,k)}{\partial \lambda} = \sum_{r=0}^{j}\sum_{s=0}^{k} b(n_1,r)b(n_2,s)f(j-r,k-s) - f(j,k), \quad j \geq 0, k \geq 0. \qquad \mathbf{I}_{jk}$$

$$\begin{aligned}\frac{\partial f(j,k)}{\partial p_1} &= n_1\lambda \sum_{r=0}^{j-1}\sum_{s=0}^{k} b(n_1-1,r)b(n_2,s)f(j-r-1,k-s) \\ &\quad - n_1\lambda \sum_{r=0}^{j}\sum_{s=0}^{k} b(n_1-1,r)b(n_2,s)f(j-r,k-s), \quad j \geq 1, k \geq 0.\end{aligned} \qquad \mathbf{II}_{jk}$$

$$\frac{\partial f(0,k)}{\partial p_1} = -n_1\lambda b(n_1-1,0)\sum_{s=0}^{k} b(n_2,s)f(0,k-s), \quad k \geq 0. \qquad \mathbf{III}_{0k}$$

$$\begin{aligned}\frac{\partial f(j,k)}{\partial p_2} &= n_2\lambda \sum_{r=0}^{j}\sum_{s=0}^{k-1} b(n_1,r)b(n_2-1,s)f(j-r,k-s-1) \\ &\quad - n_2\lambda \sum_{r=0}^{j}\sum_{s=0}^{k} b(n_1,r)b(n_2-1,s)f(j-r,k-s), \quad j \geq 0, k \geq 1.\end{aligned} \qquad \mathbf{IV}_{jk}$$

$$\frac{\partial f(j,0)}{\partial p_2} = -n_2\lambda b(n_2-1,0)\sum_{r=0}^{j} b(n_1,r)f(j-r,0), \quad j \geq 0. \qquad \mathbf{V}_{j0}$$

The elements of the information matrix I, which is 3×3, are defined as below

$$I_{11} = -\frac{1}{n}\left[\sum_{j\geq 0}\sum_{k\geq 0} n_{jk}\left\{\frac{1}{f(j,k)}\mathbf{I}_{jk}\right\}^2\right]$$

$$I_{22} = -\frac{1}{n}\left[\sum_{j\geq 1}\sum_{k\geq 0} n_{jk}\left\{\frac{1}{f(j,k)}\mathbf{II}_{jk}\right\}^2 + \sum_{k\geq 0} n_{0k}\left\{\frac{1}{f(0,k)}\mathbf{III}_{0k}\right\}^2\right]$$

$$I_{33} = -\frac{1}{n}\left[\sum_{j\geq 0}\sum_{k\geq 1} n_{jk}\left\{\frac{1}{f(j,k)}\mathbf{IV}_{jk}\right\}^2 + \sum_{j\geq 0} n_{j0}\left\{\frac{1}{f(j,0)}\mathbf{V}_{j0}\right\}^2\right]$$

$$I_{12} = -\frac{1}{n}\left[\sum_{j\geq 1}\sum_{k\geq 0} n_{jk}\left\{\frac{1}{f(j,k)}\mathbf{I}_{jk}\right\}\left\{\frac{1}{f(j,k)}\mathbf{II}_{jk}\right\} + \sum_{k\geq 0} n_{0k}\left\{\frac{1}{f(0,k)}\mathbf{I}_{0k}\right\}\left\{\frac{1}{f(0,k)}\mathbf{III}_{0k}\right\}\right]$$

$$I_{13} = -\frac{1}{n}\left[\sum_{j\geq 0}\sum_{k\geq 1} n_{jk}\left\{\frac{1}{f(j,k)}\mathbf{I}_{jk}\right\}\left\{\frac{1}{f(j,k)}\mathbf{IV}_{jk}\right\} + \sum_{j\geq 0} n_{j0}\left\{\frac{1}{f(j,0)}\mathbf{I}_{j0}\right\}\left\{\frac{1}{f(j,0)}\mathbf{V}_{j0}\right\}\right]$$

$$I_{23} = -\frac{1}{n}\left[\sum_{j\geq 1}\sum_{k\geq 1} n_{jk}\left\{\frac{1}{f(j,k)}\mathbf{II}_{jk}\right\}\left\{\frac{1}{f(j,k)}\mathbf{IV}_{jk}\right\} + \sum_{k\geq 1} n_{0k}\left\{\frac{1}{f(0,k)}\mathbf{III}_{0k}\right\}\left\{\frac{1}{f(0,k)}\mathbf{IV}_{0k}\right\}\right.$$

$$\left. + \sum_{j\geq 1} n_{j0}\left\{\frac{1}{f(j,0)}\mathbf{II}_{j0}\right\}\left\{\frac{1}{f(j,0)}\mathbf{V}_{j0}\right\} + n_{00}\left\{\frac{1}{f(0,0)}\frac{\partial f(0,0)}{\partial p_1}\right\}\left\{\frac{1}{f(0,0)}\frac{\partial f(0,0)}{\partial p_2}\right\}\right].$$

Using the notation of section (3.1), we have

$$\underline{\theta} = \begin{bmatrix} \lambda \\ p_1 \\ p_2 \end{bmatrix},$$

and

$$\Gamma(\underline{\theta}) = I,$$

which yields the increment vector

$$\underline{\delta} = I^{-1}\left(\underline{\theta}^{(0)}\right)\underline{D}.$$

For grouped data

$$\underline{D} = \begin{bmatrix} \sum_j \sum_k \frac{n_{jk}}{f(j,k)} \frac{\partial f(j,k)}{\partial \lambda} \\ \sum_j \sum_k \frac{n_{jk}}{f(j,k)} \frac{\partial f(j,k)}{\partial p_1} \\ \sum_j \sum_k \frac{n_{jk}}{f(j,k)} \frac{\partial f(j,k)}{\partial p_2} \end{bmatrix},$$

with $\underline{\theta} = \underline{\theta}^{(0)}$, where $\underline{\theta}^{(0)}$ is a trial or starting solution, usually taken to be the method of moments estimates. The iterations are stopped when $\underline{\theta}$ becomes stationary. The asymptotic variance matrix of the maximum likelihood estimator $\hat{\underline{\theta}}$ is $I^{-1}(\underline{\theta})$ and is estimated by evaluating I^{-1} at $\hat{\underline{\theta}}$, where $\hat{\underline{\theta}}$ is the vector of maximum likelihood estimates.

3.4 Efficiency

The precision of an estimator is assesed as the ratio of the generalized variance of maximum likelihood estimators to those of the estimators under consideration. The asymptotic relative efficiencies for MM and ZZ estimators relative to the ML estimator are obtained for various set of parameters. Computer package S-Plus has been used to do all the necessary computations. The tables and graphs of efficiencies of MM and ZZ estimators are given below:

Table 3.4.1 Efficiency of the Method of Moments for the Poisson-double binomial distribution

		λ				
$n_1 = n_2$	$p_1 = p_2$	0.1	0.3	0.5	1.0	1.5
2	0.1	0.6338	0.6389	0.6474	0.6764	0.7074
2	0.3	0.5364	0.4704	0.4491	0.4604	0.5027
2	0.5	0.4257	0.2943	0.2534	0.2374	0.2576
3	0.1	0.5367	0.5264	0.5272	0.5520	0.5884
3	0.3	0.4422	0.3616	0.3364	0.3446	0.3894
3	0.5	0.3504	0.2247	0.1906	0.1808	0.2030

Table 3.4.2 Efficiency of the Method of Zero-Zero Cell Frequency for the Poisson-double binomial distribution

		λ				
$n_1 = n_2$	$p_1 = p_2$	0.1	0.3	0.5	1.0	1.5
2	0.1	0.9890	0.9708	0.9587	0.9399	0.9218
2	0.3	0.9472	0.8840	0.8515	0.8187	0.7970
2	0.5	0.8528	0.7015	0.6277	0.5510	0.5171
3	0.1	0.9934	0.9801	0.9735	0.9698	0.9642
3	0.3	0.9554	0.8938	0.8620	0.8390	0.8347
3	0.5	0.8335	0.6728	0.5989	0.5308	0.5104

3.5 Discussion

While examining the tables and graphs of efficiencies, several prominent features of the Poisson-double binomial distribution are revealed. We will consider them below.

(a) Effect of change in n_1 and n_2, for fixed λ, p_1 and p_2: In this case for both types of the estimators (MM and ZZ) the resultant patterns of the efficiency tables and graphs are almost alike. It shows that with increased index values while varying other parameters, the efficiencies tend to decrease in relatively same order, as for the low values of the index parameters.

Figure 3.4.1a

Efficiency Plot for MM Method with $n_1 = n_2 = 2$.

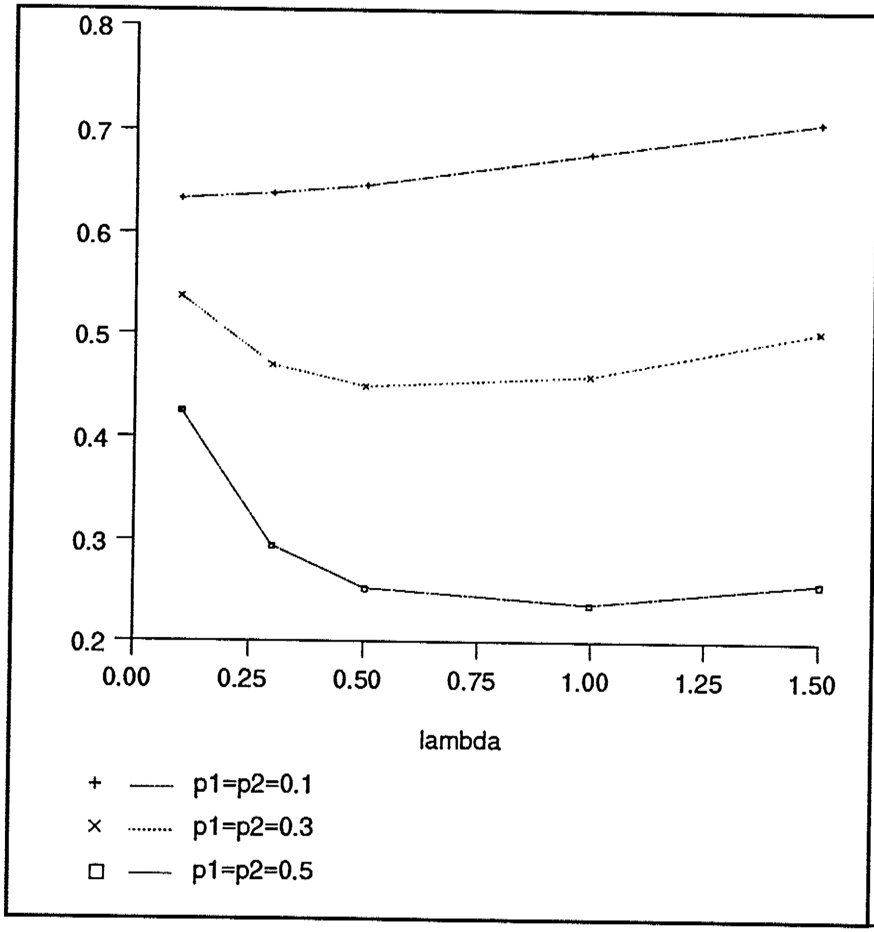


Figure 3.4.1b

Efficiency Plot for MM Method with $n_1 = n_2 = 3$.

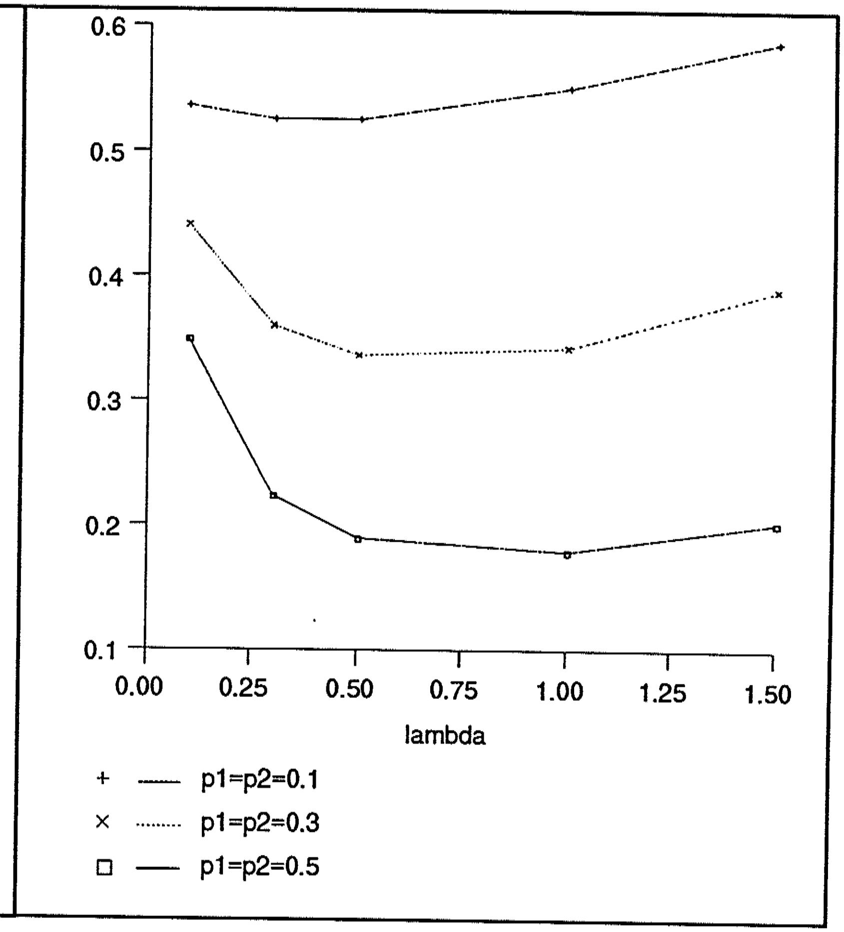


Figure 3.4.2a

Efficiency Plot for ZZ Method with $n_1 = n_2 = 2$.

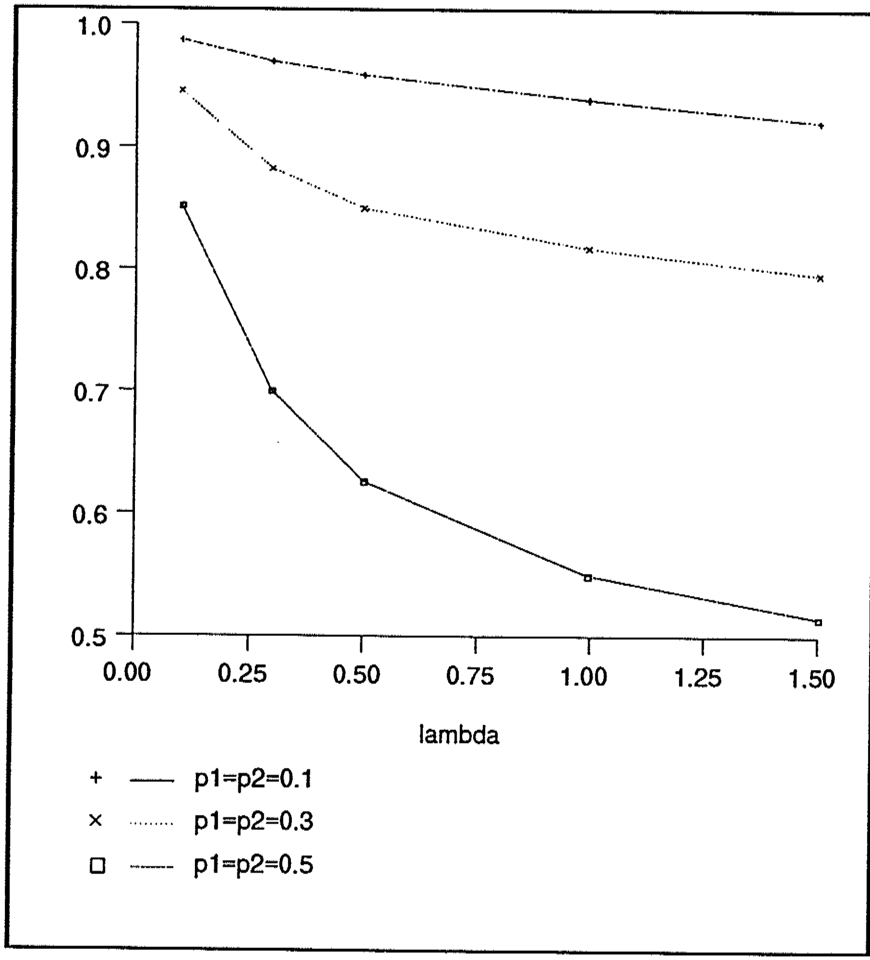
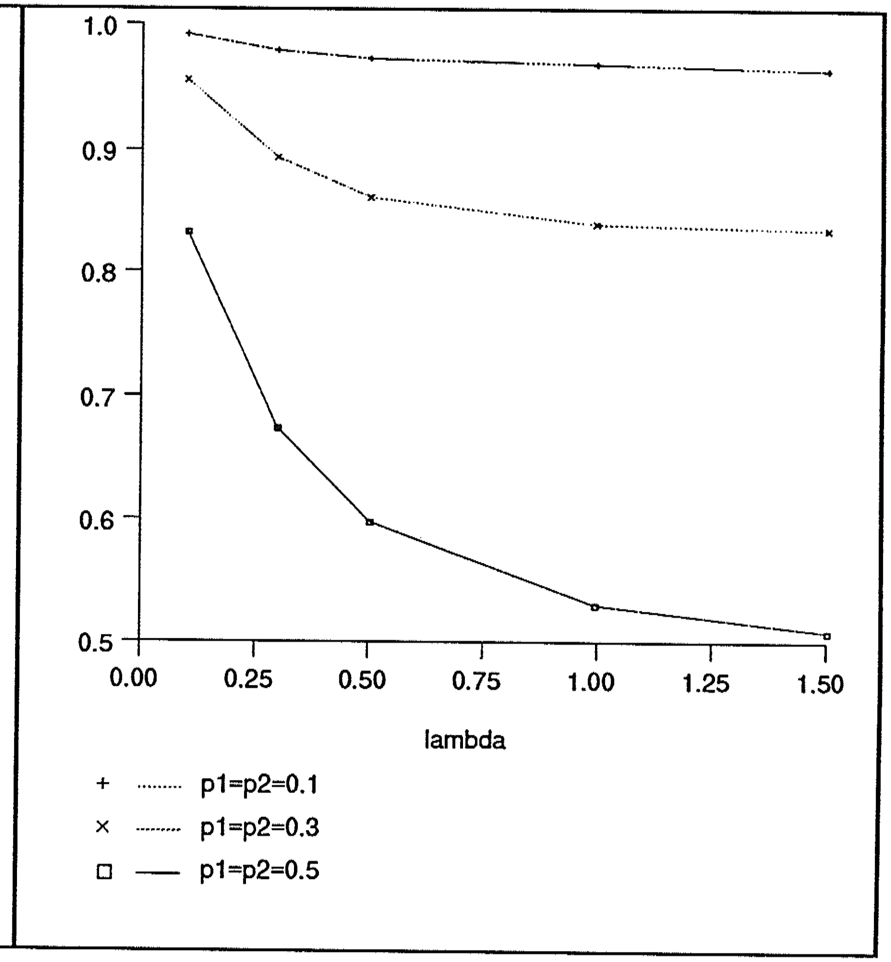
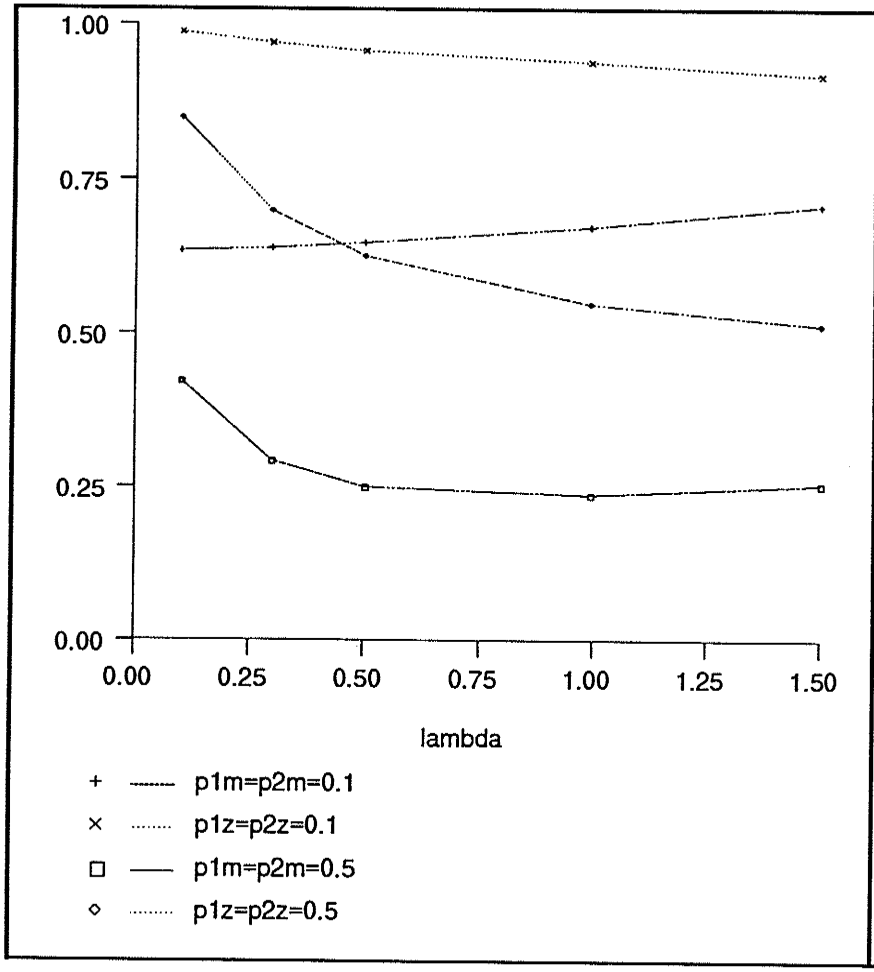


Figure 3.4.2b

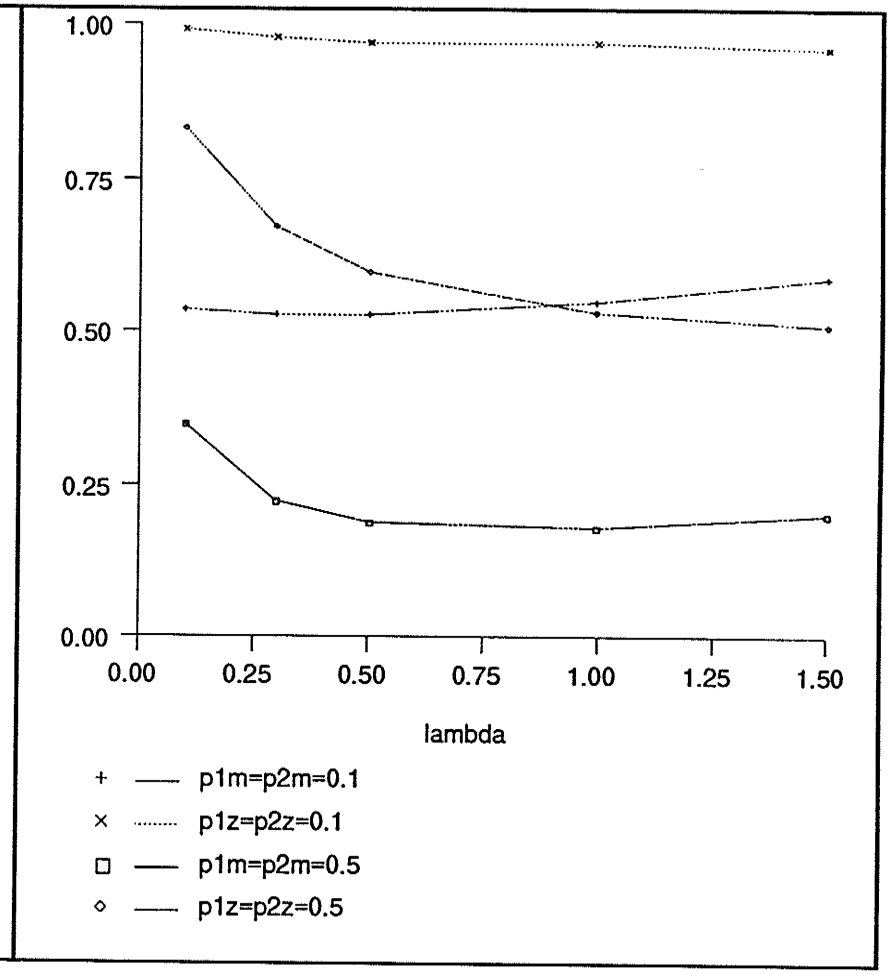
Efficiency Plot for ZZ Method with $n_1 = n_2 = 3$.



Efficiency Plot for MM and ZZ Methods for $n_1 = n_2 = 2$.



Efficiency Plot for MM and ZZ Methods for $n_1 = n_2 = 3$



(b) Effect of change in λ for fixed n_1, n_2, p_1 and p_2: The dominent parameter in the Poisson-double binomial distribution is λ. It is evident from the graphs 3.4.1a and b that, as λ increases for fixed values of the other parameters, the efficiencies first fall until λ reaches the value 1.5, there the efficiency tends to increase again. In the ZZ case considering the graphs 3.4.2a and b, it can be seen that the pattern is different from MM case. Here with large values of λ, the efficiency becomes smaller and smaller.

(c) Effect of change in p_1 and p_2 for fixed λ, n_1 and n_2: In this case the efficiencies become small with increased p_1 and p_2 values for both types of estimators. There is a consistency in the pattern of efficiencies.

(d) Overall comparison of MM and ZZ methods: In general the efficiency patterns for the two methods are not exactly alike. For a fixed set of parameters ZZ method has high efficiency as compared to MM method. One feature that is common for both methods is that, efficiency is highest for the parameter set having smallest values and lowest for the one with highest values.

4

APPLICATIONS

4.1 Introduction

Simulation or computer generation of certain specified bivariate discrete distributions has been discussed in detail in Kocherlakota and Kocherlakota (1992). In this chapter the computer generation of Poisson-double binomial distribution will be presented. Also various methods of estimation and goodness-of-fit will be applied to data for Poisson-double binomial and bivariate Poisson distributions.

4.2 Poisson-double binomial distribution

Simulation

To simulate the Poisson-double binomial distribution, we use the stochastic nature of the Poisson-double binomial distribution. Since there was no standard algorithm available, a new technique was developed to obtain random observations from the Poisson-double binomial distribution. Using the options available in S-plus, it is possible to generate random observations from Poisson and binomial distributions. This simulation program is given in appendix under the name of si.mat5. In this algorithm first a realization k from a Poisson distribution is generated for a specified value of the Poisson parameter. In the second step, this generated value of the Poisson observation is used to get two independent binomial observations with index parameter determined by the formerly realized Poisson observation. The procedure is repeated a desired, say n numbers of times, to obtain n random pairs of observations from the two binomial distribution.

These observations can then be arranged into a frequency table to give the n random observations from the Poisson-double binomial distribution.

Examples

Various estimation techniques and goodness- of-fit procedures are illustrated on the following three data sets. First two data sets were simulated using si.mat5, while the third one is real life data.

Data set I: $n = 200$, $\lambda = 0.5$, $n_1 = 2$, $n_2 = 2$, $p_1 = 0.25$ and $p_2 = 0.25$.

	x_2			
x_1	0	1	2	3
0	150	7	4	0
1	16	11	2	1
2	2	4	1	1
3	0	1	0	0

Data set II: $n = 200$, $\lambda = 1$, $n_1 = n_2 = 2$, $p_1 = p_2 = 0.25$.

	x_2				
x_1	0	1	2	3	4
0	111	15	4	1	0
1	20	21	4	2	1
2	2	5	6	1	0
3	3	1	1	1	0
4	1	0	0	0	0

The following data set is adapted from Cresswell *et al.* (1963). The data represent the observed frequency distribution of the accidents sustained by bus drivers in two consecutive 2-year time periods. Total number of observations is 708, and it is assumed that the data are from the Poisson-double binomial distribution with $n_1 = n_2 = 2$.

Data set III:

x_1	x_2 0	1	2	3	4	5	6	7
0	117	96	55	19	2	2	0	0
1	61	69	47	27	8	5	1	0
2	34	42	31	13	7	2	3	0
3	7	15	16	7	3	1	0	0
4	3	3	1	1	2	1	1	1
5	2	1	0	0	0	0	0	0
6	0	0	0	0	1	0	0	0
7	0	0	0	1	0	0	0	0

Estimation:

In the following zero-zero cell frequency estimate of λ is obtained by using the computer package Mathematica. The ML estimates are obtained by using the S-plus program iter1, which is given in appendix. This program uses the iteration procedure described in section (3.1), which is called the method of scoring. The method of moments estimates are taken as the initial values.

Data set	Summary Statistics $\bar{x}_1$	$\bar{x}_2$	$m_{1,1}$
I	0.2450	0.2150	0.1423
II	0.4900	0.4550	0.2621
III	1.0014	1.2910	0.3259

Estimates of the parameters along with their estimated standard errors using three methods of estimation are given in the following table. For calculating the estimates of the standard errors of the various types of estimators, S-Plus programs try, vmom, vzero1, comp, dvmom, dvzero1 were used. These programs are given in appendix.

Table of values of Estimators along with their Standard Errors

Method	Parameter	Data I	Data II	Data III
MM	λ	0.3701	0.8508	3.9669
		(0.0888)	(0.1851)	(0.6573)
	p_1	0.3310	0.2880	0.1262
		(0.0795)	(0.0636)	(0.0212)
	p_2	0.2905	0.2674	0.1627
		(0.0715)	(0.0595)	(0.0272)
ZZ	λ	0.3728	0.7603	3.5790
		(0.0690)	(0.1185)	(0.5752)
	p_1	0.3286	0.3222	0.1399
		(0.0611)	(0.0506)	(0.0228)
	p_2	0.2884	0.2992	0.1804
		(0.0558)	(0.0477)	(0.0293)
ML	λ	0.3557	0.8157	3.9983
		(0.0627)	(0.1234)	(0.5999)
	p_1	0.3374	0.3004	0.1252
		(0.0592)	(0.0465)	(0.0192)
	p_2	0.2882	0.2789	0.1614
		(0.0534)	(0.0438)	(0.0245)

The numbers in parentheses are the respective standard errors.

Tests of goodness-of-fit

Two testing procedures, Pearson's chi-square $\left(\chi^2\right)$ goodness-of-fit test and empirical pgf technique test, are used. While using these techniques it is necessary to have the data in a two way array. The frequency in (r, s)th cell is denoted by n_{rs} while the corresponding probability is denoted by $f(r,s)$. The hypothesis of interest is

$$H_0:\ f(r,s) = f^0(r,s)$$

where $f^0(r,s)$ is specified and $\sum_r \sum_s f^0(r,s) = 1$. Generally the specified functional form of the probability function (pf) involves many parameters, say k, which have to be estimated from the given data set. Using the estimates of these k parameters it is always possible to find the estimated probabilities $\hat{f}^0(r,s)$. The testing procedure is to test whether the data conform to the prescribed distribution estimated by $\hat{f}^0(r,s)$. The details of the two testing procedures named before are given below:

<u>Pearson's χ^2 test</u>

The test statistic is

$$\underline{X}^2 = \sum_r \sum_s \frac{\left[n_{rs} - n\hat{f}^0(r,s)\right]^2}{n\hat{f}^0(r,s)},$$

which has the χ^2 distribution for large n. The degrees of freedom (df) are: number of cells in the array - 1 - number of parameters estimated. In order to use this statistic properly it is essential that the expected frequency in each cell should be large. Most of the statisticians are agreed on the use of minimum expected frequency taken to be 5. In case of a large data set the minimum expected frequency can be taken as small as 1. Cochran (1954) suggests a bit different approach. According to him if relatively few expectations are less than 5, then a minimum expectation of 1 can be used in calculation of $\underline{X}^2$ for any contigency table with df greater than 1. To ensure a minimum expectation of 1, the data have to be grouped in a condensed form. Grouping is rather easier in univariate case, but for bivariate case there is no hard and fast rule. Loukas and Kemp (1986) give three systematic grouping procedures: (i) row based, (ii) column based and (iii) reordering the classes so that they may be treated as a J-shaped univariate distribution. Depending on the data either of these procedures should be used. In the following we will use a combination of procedures (i) and (ii).

Probability generating function technique

This technique was developed by Kocherlakota and Kocherlakota (1986). The advantage of this method is that it does not require pooling of the cells, which is subjective in most cases. This technique is based on the probability generating function of discrete random variables to test goodness-of-fit hypotheses. Considering the bivariate case, let $f(x_1,x_2;\underline{\theta})$ be the joint probability function of two random variables X_1 and X_2. By definition the joint probability generating function of X_1 and X_2 is given by $\pi(t_1,t_2;\underline{\theta}) = E\left(t_1^{X_1}t_2^{X_2}\right)$, where $\underline{\theta}$ is a k element vector of parameters. Let the sample consists of n observations with frequency in (r, s)th cell being n_{rs}, with $n = \sum_r \sum_s n_{rs}$.

Based on this sample the empirical (sample) pgf is defined as

$$P(t_1,t_2) = \frac{1}{n}\sum_r \sum_s n_{rs} t_1^r t_2^s,$$

where $|t_i| < 1$, i =1, 2. It can be seen that

$$E[P(t_1,t_2)] = \pi(t_1,t_2).$$

Let $\hat{\underline{\theta}}$ be the maximum likelihood estimator of $\underline{\theta}$. Then the maximum likelihood estimator of $\pi(t_1,t_2;\underline{\theta})$ is $\hat{\pi}(t_1,t_2;\hat{\underline{\theta}})$. Since $\hat{\underline{\theta}}$ is the maximum likelihood estimator, we can expand $\hat{\pi}(t_1,t_2;\hat{\underline{\theta}})$. Retaining only the first order terms

$$\hat{\pi}(t_1,t_2;\hat{\underline{\theta}}) = \pi(t_1,t_2;\underline{\theta}) + \sum_{i=1}^{k}\left(\hat{\theta}_i - \theta_i\right)\frac{\partial \pi(t_1,t_2;\underline{\theta})}{\partial \theta_i}. \qquad (4.2.1)$$

Consider now the random variable

$$\xi(t_1,t_2) = P(t_1,t_2) - \hat{\pi}(t_1,t_2;\hat{\underline{\theta}}),$$

using (4.2.1), we have

$$\xi(t_1,t_2) = \{P(t_1,t_2) - \pi(t_1,t_2;\underline{\theta})\} - \sum_{i=1}^{k}(\hat{\theta}_i - \theta_i)\frac{\partial\pi(t_1,t_2;\underline{\theta})}{\partial\theta_i}.$$

Using the asymptotic properties of ML estimators

$$\xi(t_1,t_2) \xrightarrow{d} N(0,\sigma^2),$$

where

$$\sigma^2 = \frac{1}{n}\left[\pi(t_1^2,t_2^2;\underline{\theta}) - \pi^2(t_1,t_2;\underline{\theta})\right] - \sum_{i}^{k}\sum_{j}^{k}\sigma_{ij}\frac{\partial\pi(t_1,t_2;\underline{\theta})}{\partial\theta_i}\frac{\partial\pi(t_1,t_2;\underline{\theta})}{\partial\theta_j} \tag{4.2.2}$$

with $\{\sigma_{ij}\}$ as the inverse of the information matrix. In the case of Poisson-double binomial distribution

$$\underline{\theta} = \begin{bmatrix} \lambda \\ p_1 \\ p_2 \end{bmatrix},$$

and

$$\pi(t_1,t_2) = \exp\left[\lambda\left\{(q_1 + p_1t_1)^{n_1}(q_2 + p_2t_2)^{n_2} - 1\right\}\right].$$

The partial derivatives of pgf with respect to the parameters are given as

$$\frac{\partial\pi(t_1,t_2)}{\partial\lambda} = \pi(t_1,t_2)\left[(q_1 + p_1t_1)^{n_1}(q_2 + p_2t_2)^{n_2} - 1\right]$$

$$\frac{\partial\pi(t_1,t_2)}{\partial p_1} = n_1\lambda(t_1 - 1)(q_1 + p_1t_1)^{n_1-1}(q_2 + p_2t_2)^{n_2}\pi(t_1,t_2)$$

$$\frac{\partial\pi(t_1,t_2)}{\partial p_2} = n_2\lambda(t_2 - 1)(q_1 + p_1t_1)^{n_1}(q_2 + p_2t_2)^{n_2-1}\pi(t_1,t_2).$$

The test statistic to be used is

$$Z = \frac{P(t_1,t_2) - \pi(t_1,t_2;\hat{\underline{\theta}})}{\sigma},$$

which is approximately N(0,1) under the null hypothesis. To perform the test σ^2 in (4.2.2) is estimated by replacing the parameters by their maximum likelihood estimates. If the parameters are assumed known, then

$$\sigma^2 = \frac{1}{n}\left[\pi\left(t_1^2, t_2^2; \underline{\theta}\right) - \pi^2\left(t_1, t_2; \underline{\theta}\right)\right].$$

The null hypothesis is rejected if

$$|Z| > z_{\frac{\alpha}{2}},$$

where α is the level of significance and z_γ being the upper 100 γ percent point of the standard normal distribution. To calculate Z statistic in both cases (parameters estimated and parameters known), we used four S-plus programs; namely, fit1, pgf1, epgf1 and fit2, which are given in appendix.

Examples

In order to ensure that the expected frequency in each cell is at least 1, the original set of data has been condensed in the following form. The numbers in parentheses are the expected frequencies. For each of the testing procedure, it is concluded that the data do fit the Poisson-double binomial distribution. For the chi-square and pgf techniques the ML estimates were used. For the pgf technique value of the test statistics Z depends on the choice of t_1 and t_2. In the following table both are taken to be 0.1. The results are given in the following tables.

Data set I:

	x_2		
x_1	0	1	≥ 2
0	150 (151.6750)	7 (9.7181)	4 (2.4293)
1	16 (12.2219)	11 (10.6802)	3 (3.1525)
≥ 2	2 (3.9136)	5 (4.0915)	2 (2.1177)

Statistic	Value	df	Prob-value
Chi-square	4.1231	5	0.5318
Z(parameters est)	-1.3436	_	0.1791
Z(parameters known)	1.1846	_	0.2362

Data set II:

x_1	x_2 = 0	1	≥ 2
0	111 (108.4848)	15 (17.5752)	5 (5.5860)
1	20 (19.3686)	21 (18.2022)	7 (7.7675)
≥ 2	6 (6.8931)	6 (8.7696)	9 (7.3312)

Statistic	Value	df	Prob-value
Chi-square	2.3939	5	0.7924
Z(parameters est)	1.0750	_	0.2824
Z(parameters known)	1.3678	_	0.1714

Data set III:

x_1	x_2 = 0	1	2	3	4
0	117 (111.7146)	96 (92.5319)	55 (47.2260)	19 (17.9559)	4 (7.5327)
1	61 (68.8079)	69 (83.4788)	47 (53.5746)	27 (24.3673)	14 (12.3137)
2	34 (26.1141)	42 (39.8387)	31 (31.0135)	13 (16.5511)	12 (10.0216)
3	7 (7.3833)	15 (13.4741)	16 (12.3076)	7 (7.5648)	4 (5.4471)
4	5 (2.1232)	4 (4.6134)	1 (4.9852)	2 (3.5943)	6 (3.2334)

Statistic	Value	df	Prob-value
Chi-square	23.7144	21	0.3071
Z(parameters est)	1.2573	_	0.2086

4.3 Bivariate Poisson distribution

Example

Consider the data set III of section (4.2). Assuming the data are from the bivariate Poisson distribution, parameters estimation and goodness of fit tests will be done in this section. The data is given below

	x_2							
x_1	0	1	2	3	4	5	6	7
0	117	96	55	19	2	2	0	0
1	61	69	47	27	8	5	1	0
2	34	42	31	13	7	2	3	0
3	7	15	16	7	3	1	0	0
4	3	3	1	1	2	1	1	1
5	2	1	0	0	0	0	0	0
6	0	0	0	0	1	0	0	0
7	0	0	0	1	0	0	0	0

Estimation

The summary statistics for this data set are: $\bar{x}_1 = 1.0014$, $\bar{x}_2 = 1.2910$, $m_{1,1} = 0.3259$.

The method of moments estimates are $\tilde{\lambda} = 3.9669$, $\tilde{p}_1 = 0.2524$ and $\tilde{p}_2 = 0.3254$. The zero-zero cell frequency estimates are $\lambda' = 2.6271$, $p_1' = 0.3812$ and $p_2' = 0.4914$. Using the results given in Kemp *et al.* (1983, Table 2) the maximum likelihood estimates are found to be $\hat{\lambda} = 5.6432$, $\hat{p}_1 = 0.1774$ and $\hat{p}_2 = 0.2288$. In tabular form

Method	Estimates		
	λ	p_1	p_2
MM	3.9669	0.2524	0.3254
ZZ	2.6271	0.3812	0.4914
ML	5.6432	0.1774	0.2288

To apply the pgf technique for estimated parametrs we need the partial derivatives of the pgf with respect to parameters of the bivariate Poisson distribution. For the bivariate Poisson distribution the pgf is given as

$$\pi(t_1,t_2)=\exp[\lambda\{q_1q_2-1+p_1q_2t_1+p_2q_1t_2+p_1p_2t_1t_2\}].$$

The partial derivatives with respect to the parameters are given as

$$\frac{\partial\pi(t_1,t_2)}{\partial\lambda}=\pi(t_1,t_2)\{q_1q_2-1+p_1q_2t_1+p_2q_1t_2+p_1p_2t_1t_2\}$$

$$\frac{\partial\pi(t_1,t_2)}{\partial p_1}=\pi(t_1,t_2)[q_2(t_1-\lambda)+p_2t_2(t_1-1)]$$

$$\frac{\partial\pi(t_1,t_2)}{\partial p_2}=\pi(t_1,t_2)[q_1(t_2-\lambda)+p_1t_1(t_2-1)].$$

In order to ensure that the expected frequency in each cell is at least 1, the original set of data has been condensed in the following form. The numbers in parentheses are the expected frequencies. For each of the testing procedure, it is concluded that the data do not fit the bivariate Poisson distribution. For the chi-square and pgf techniques the ML estimates were used. For the pgf technique value of the test statistics Z depends on the choice of t_1 and t_2. In the following table both are taken to be 0.1. The results are given in the following tables.

	x_2				
x_1	0	1	2	3	4
0	117 (88.5476)	96 (94.1884)	55 (50.0943)	19 (17.7618)	4 (5.9334)
1	61 (69.2180)	69 (94.3030)	47 (61.1515)	27 (25.5813)	14 (10.1646)
2	34 (27.0540)	42 (44.9396)	31 (34.9109)	13 (17.1378)	12 (8.1388)
3	7 (7.0494)	15 (13.8154)	16 (12.5944)	7 (7.1823)	4 (4.0827)
4	5 (1.6242)	4 (3.7522)	1 (4.0455)	2 (2.7284)	6 (1.9471)

Statistic	Value	df	Prob-value
Chi-square	47.1695	21	0.0009
Z(parameters est)	3.8570	–	0.0001

4.4 Conclusion

In sections (4.2) and (4.3) we have applied two types of goodness-of-fit techniques on data sets for Poisson-double binomial and bivariate Poisson distribution. In (4.2) two simulated and one real data set were used. All of three do fit the Poisson-double binomial distribution. For simulated data sets it is natural to expect that they must fit the Poisson-double binomial distribution. We tried to fit the same data set (data III) for the two distributions under consideration. The data supply the strong evidence that they came from the Poisson-double binomial distribution, but fail to do so for the bivariate Poisson distribution. This result reveals a very important aspect of the different theoretical development of the two distributions.

APPENDIX: S-PLUS PROGRAMS

(1) si.mat5: This program is written to get random observations from the Poisson-double binomial distribution. Poisson-double binomial distribution has five parameters. In this program n is the total sample size.

```
function(n, lambda, n1, n2, p1, p2)
{
        result <- matrix(0, n, 3)
        x1 <- matrix(0, 1, 1)
        x2 <- matrix(0, 1, 1)
        for(i in 1:n) {
                k <- rpois(1, lambda)
                if(k == 0) {
                        x1 == 0
                        x2 == 0
}
                else {
                        x1 <- rbinom(1, n1 * k, p1)
                        x2 <- rbinom(1, n2 * k, p2)
                        result[i, ] <- c(x1, x2, k)
}
}
        list(result1 <- result, result2 <- table(result[, 1],
result[, 2]))
}
```

(2) pfc911: The following program is written as a part of the next program i.e. program (3). The purpose of this program is to provide basis for the computations of probability function at various X_1 and X_2 values for the Poisson-double binomial distribution. In the following program r and c are the limits for X_1 and X_2 values.

```
function(lam, n1, n2, p1, p2, r, c)
{
        ff <- matrix(0, r, c)
        ff[1, 1] <- exp(lam * (dbinom(0, n1, p1) * dbinom(0,
n2, p2) - 1))
        for(j in 2:r) {
                ps <- 0
                a <- (lam/(j - 1)) * dbinom(0, n2, p2)
                for(h in 1:(j - 1)) {
                        ps <- ps + a * h * dbinom(h, n1, p1) *
ff[j - h, 1]
}
                ff[j, 1] <- ps
```

```
}
        for(l in 2:c) {
                ps1 <- 0
                a1 <- (lam/(l - 1)) * dbinom(0, n1, p1)
                for(m in 1:(l - 1)) {
                    ps1 <- ps1 + a1 * m * dbinom(m, n2, p2) *
ff[1, l - m]
}
                ff[1, l] <- ps1
}
ff
}
```

(3) pfc931: This is the actual program for the computation of the Poisson-double binomial probabilities.

```
function(lam, n1, n2, p1, p2, r, c)
{
        ff1 <- pfc911(lam, n1, n2, p1, p2, r, c)
        for(jj in 2:r) {
                for(ll in 2:c) {
                        ps2 <- 0
                        a2 <- (lam/(jj - 1))
                        for(hh in 1:(jj - 1)) {
                                for(mm in 0:(ll - 1)) {
                                        ps2 <- ps2 + a2 * hh *
dbinom(hh, n1, p1) * dbinom(mm, n2, p2) * ff1[jj - hh, ll -
mm]}}
ff1[jj, ll] <- ps2}}
ff1
}
```

(4) nb1: The following S-plus program is used to differentiate the Poisson-double binomial pf with respect to λ. In the following program symbol lam is used for λ

```
function(lam, n1, n2, p1, p2, j, k)
{
        ff <- pfc931(lam, n1, n2, p1, p2, j, k)
        dfl <- matrix(0, j, k)
            dfl[1, 1] <- ff[1, 1] * (dbinom(0, n1, p1) *
dbinom(0, n2, p2) - 1)
        for(t in 2:(n1 + 1)) {
                cs <- 0
                for(r in 0:(t - 1)) {
                        cs <- cs + (1 - p2)^n2 * dbinom(r,
n1, p1) * ff[t - r, 1] }
                dfl[t, 1] <- cs - ff[t, 1] }
        for(m in 2:(n2 + 1)) {
                cs1 <- 0
```

```
                  for(s in 0:(m - 1)) {
                          cs1 <- cs1 + (1 - p1)^n1 * dbinom(s,
n2, p2) * ff[1, m - s] }
                  df1[1, m] <- cs1 - ff[1, m] }
            for(t in (n1 + 2):j) {
                  cs <- 0
                  for(r in 0:n1) {
                          cs <- cs + (1 - p2)^n2 * dbinom(r,
n1, p1) * ff[t - r, 1] }
                  df1[t, 1] <- cs - ff[t, 1] }
        for(m in (n2 + 2):k) {
                      cs1 <- 0
                  for(s in 0:n2) {
                          cs1 <- cs1 + (1 - p1)^n1 * dbinom(s,
n2, p2) * ff[1, m - s] }
                  df1[1, m] <- cs1 - ff[1, m] }
        for(t in 2:(n1 + 1)) {
                  for(m in 2:(n2 + 1)) {
                          cs <- 0
                          for(r in 0:(t - 1)) {
                                  for(s in 0:(m - 1)) {
                                      cs <- cs + dbinom(r, n1,
p1) * dbinom(s, n2,p2)*ff[t - r, m - s] } }
                  df1[t, m] <- cs - ff[t, m] } }
        for(t in 2:(n1 + 1)) {
                  for(m in (n2 + 2):k) {
                          cs <- 0
                          for(r in 0:(t - 1)) {
                                  for(s in 0:n2) {
                                      cs <- cs + dbinom(r, n1,
p1) * dbinom(s, n2, p2)*ff[t - r, m - s] } }
                  df1[t, m] <- cs - ff[t, m] } }
      for(t in (n1 + 2):j) {
                  for(m in 2:(n2 + 1)) {
                          cs <- 0
                          for(r in 0:n1) {
                                  for(s in 0:(m - 1)) {
                                      cs <- cs + dbinom(r, n1,
p1) * dbinom(s, n2, p2) * ff[t - r, m - s] } }
                  df1[t, m] <- cs - ff[t, m] } }
        for(t in (n1 + 2):j) {
                  for(m in (n2 + 2):k) {
                          cs <- 0
                          for(r in 0:n1) {
                                  for(s in 0:n2) {
                                      cs <- cs + dbinom(r, n1,
p1) * dbinom(s, n2, p2)*ff[t - r, m - s] } }
                  df1[t, m] <- cs - ff[t, m] } }
df1
}
```

(5) nb3: Following program is written to get the differential coefficients when Poisson-double binomial pf is differentiated with respect to p_1.

```
function(lam, n1, n2, p1, p2, j, k)
{
        ff <- pfc931(lam, n1, n2, p1, p2, j, k)
        dfp1 <- matrix(0, j, k)
        dfp1[1, 1] <- ( - n1) * lam * dbinom(0, n1 - 1, p1) *
dbinom(0, n2, p2)*ff[1,1]
        for(t in 1:1) {
                for(m in 2:k) {
                        cs <- 0
                        for(s in 0:(m - 1)) {
                             cs <- cs + dbinom(0, n1 - 1, p1)
* dbinom(s, n2  p2) * ff[1, m-s] }
         dfp1[1, m] <- ( - n1) * lam * cs } }
        for(t in 2:j) {
                for(m in 1:k) {
                     cs1 <- 0
                        cs2 <- 0
                        for(r in 0:(t - 2)) {
                                for(s in 0:(m - 1)) {
                              cs1 <- cs1 + dbinom(r, n1 - 1,
p1) * dbinom(s, n2, p2) * ff[t - r - 1, m - s] } }
       for(r in 0:(t - 1)) {
               for(s in 0:(m - 1)) {
                                   cs2 <- cs2 + dbinom(r, n1 -
1, p1) * dbinom(s, n2, p2) * ff[t - r, m - s] } }
       dfp1[t, m] <- n1 * lam * cs1 - n1 * lam * cs2 } }
       dfp1
}
```

(6) nb5: The following program is written to get the differential coefficients when Poisson-double binomial pf is differetiated with respect to p_2.

```
function(lam, n1, n2, p1, p2, j, k)
{
        ff <- pfc931(lam, n1, n2, p1, p2, j, k)
        dfp2 <- matrix(0, j, k)
          dfp2[1, 1] <-   - n2 * lam * dbinom(0, n1, p1) *
dbinom(0, n2 - 1, p2) *ff[1, 1]
     for(t in 2:j) {
                for(m in 1:1) {
                        cs <- 0
                        for(r in 0:(t - 1)) {
                                cs <- cs + dbinom(0, n2 - 1,
p2) * dbinom(r, n1, p1) * ff[t - r, 1] }
                        dfp2[t, 1] <-   - n2 * lam * cs } }
     for(t in 1:j) {
                for(m in 2:k) {
```

```
                        cs1 <- 0
                        cs2 <- 0
                        for(r in 0:(t - 1)) {
                                for(s in 0:(m - 2)) {
                                   cs1 <- cs1 + dbinom(r, n1,
p1) * dbinom(s, n2-1, p2) * ff[t - r, m - s - 1] } }
                        for(r in 0:(t - 1)) {
                                for(s in 0:(m - 1)) {
                                   cs2 <- cs2 + dbinom(r, n1,
p1) * dbinom(s, n2-1,p2) * ff[t - r, m - s] } }
                         dfp2[t, m] <- n2 * lam * cs1 - n2 *
lam * cs2 } }
        dfp2
}
```

(7) try: Following program is to get the gamma matrix defined in Kocherlakota and Kocherlakota (1992, p. 46), which is used to obtain maximum likelihood estimators.

```
function(n, lam, n1, n2, p1, p2, j, k)
{
        ff <- pfc931(lam, n1, n2, p1, p2, j, k)
        dl <- nb1(lam, n1, n2, p1, p2, j, k)
        dp1 <- nb3(lam, n1, n2, p1, p2, j, k)
        dp2 <- nb5(lam, n1, n2, p1, p2, j, k)
        gamt <- matrix(0, 3, 3)
        cs <- 0
        for(r in 1:j) {
                for(s in 1:k) {
                        cs <- cs + (1/ff[r, s]) * dl[r, s] *
dl[r, s]
}
}
        gamt[1, 1] <- n * cs
        cs <- 0
        for(r in 1:j) {
                for(s in 1:k) {
                        cs <- cs + (1/ff[r, s]) * dp1[r, s] *
dp1[r, s]
}
}
        gamt[2, 2] <- n * cs
        cs <- 0
        for(r in 1:j) {
                for(s in 1:k) {
                        cs <- cs + (1/ff[r, s]) * dp2[r, s] *
dp2[r, s]
}
}
        gamt[3, 3] <- n * cs
        cs <- 0
        for(r in 1:j) {
```

```
                for(s in 1:k) {
                        cs <- cs + (1/ff[r, s]) * dl[r, s] *
dp1[r, s]
                }
        }
        gamt[1, 2] <- n * cs
        gamt[2, 1] <- gamt[1, 2]
        cs <- 0
        for(r in 1:j) {
                for(s in 1:k) {
                        cs <- cs + (1/ff[r, s]) * dl[r, s] *
dp2[r, s]
}
}
        gamt[1, 3] <- n * cs
        gamt[3, 1] <- gamt[1, 3]
        cs <- 0
        for(r in 1:j) {
                for(s in 1:k) {
                        cs <- cs + (1/ff[r, s]) * dp1[r, s] *
dp2[r, s]
}
}
        gamt[2, 3] <- n * cs
        gamt[3, 2] <- gamt[2, 3]
gamt
}
```

(8) try1: Following program is written to get the D matrix, which is needed to obtain the iteration vector for maximum likelihood estimates in case of Poisson-double binomial distribution. This matrix is also defined in Kocherlakota and Kocherlakota (1992, p. 46). nrs is the observed data set.

```
function(lam, n1, n2, p1, p2, j, k)
{
        ff <- pfc931(lam, n1, n2, p1, p2, j, k)
        dl <- nb1(lam, n1, n2, p1, p2, j, k)
        dp1 <- nb3(lam, n1, n2, p1, p2, j, k)
        dp2 <- nb5(lam, n1, n2, p1, p2, j, k)
        fr <- nrs
        dss <- matrix(0, 3, 1)
        dss[1, 1] <- sum(fr * (1/ff) * dl)
        dss[2, 1] <- sum(fr * (1/ff) * dp1)
        dss[3, 1] <- sum(fr * (1/ff) * dp2)
        dss
}
```

(9) iter1: Following program is used for iterations which are necessary to maximum likelihood estimates. In this program lami, p1i, p2i are the starting values of the three parameters to be estimated. Usually these initial values are either method of moments or zero-zero cell frequency estimates.

```
function(lami,p1i,p2i,n,n1,n2,j,k,st)
{
lam<-lami
p1<-p1i
p2<-p2i
it<-0
del<-matrix(0,3,1)
vcov<-matrix(0,3,3)
while((it<-it+1)<st){del<-
solve(try(n,lam,n1,n2,p1,p2,j,k))%*%try1(lam,n1,n2,p1,p2,j,k)
lam<-lam+del[1,1]
p1<-p1+del[2,1]
p2<-p2+del[3,1]
vcov<-solve(try(n,lam,n1,n2,p1,p2,j,k))
cat(it,lam,p1,p2,"\n")
print(del)
print(vcov)} }
```

(10) gam1: Purpose of this program is the same as of program (7), except the difference that following program is for the estimation of p_1 and p_2, while λ is assumed known.

```
function(n,lam,n1,n2,p1,p2,j,k)
{
ff<-pfc931(lam,n1,n2,p1,p2,j,k)
dp1<-nb3(lam,n1,n2,p1,p2,j,k)
dp2<-nb5(lam,n1,n2,p1,p2,j,k)
gamt<-matrix(0,2,2)
cs<-0
for(r in 1:j) {
                for(s in 1:k){
                        cs<-cs+(1/ff[r,s])*dp1[r,s]*dp1[r,s]}}
gamt[1,1]<-n*cs
cs<-0
for(r in 1:j) {
                for(s in 1:k) {
                        cs<-cs+(1/ff[r,s])*dp2[r,s]*dp2[r,s]}}
gamt[2,2]<-n*cs
cs<-0
for(r in 1:j) {
                for(s in 1:k) {
                        cs<-cs+(1/ff[r,s])*dp1[r,s]*dp2[r,s]}}
gamt[1,2]<-n*cs
gamt[2,1]<-gamt[1,2]
```

```
gamt
}
```

(11) ds1: Purpose is the same as of program (8), except the difference that this program is for the estimation of p_1 and p_2, while λ is assumed known. nrs is the observed frequency distribution.

```
function(lam,n1,n2,p1,p2,j,k)
{
ff<-pfc931(lam,n1,n2,p1,p2,j,k)
dp1<-nb3(lam,n1,n2,p1,p2,j,k)
dp2<-nb5(lam,n1,n2,p1,p2,j,k)
fr<-nrs
dss<-matrix(0,2,1)
dss[1,1]<-sum(fr*(1/ff)*dp1)
dss[2,1]<-sum(fr*(1/ff)*dp2)
dss
}
```

(12) iter2: Purpose is the same as of program (9), except the difference that this program is for the estimation of p_1 and p_2, while λ is assumed known.

```
function(lam,p1i,p2i,n,n1,n2,j,k,st)
{
lam<-lam
p1<-p1i
p2<-p2i
it<-0
del<-matrix(0,2,1)
vcov<-matrix(0,2,2)
while((it<-it+1)<st)
{del<-
solve(gam1(n,lam,n1,n2,p1,p2,j,k))%*%ds1(lam,n1,n2,p1,p2,j,k)
p1<-p1+del[1,1]
p2<-p2+del[2,1]
vcov<-solve(gam1(n,lam,n1,n2,p1,p2,j,k))
cat(it,lam,p1,p2,"\n")
print(del)
print(vcov)}}
```

(13) mom: This program results in numerical values of the population raw moments and moments about mean, when parametric values are supplied. In this program symbol m is used for λ.

```
function(m, n1, n2, p1, p2)
```

```
{
        result <- matrix(0, 12, 1)
        mp10 <- m * n1 * p1
        result[1, 1] <- mp10
        mp01 <- m * n2 * p2
        result[2, 1] <- mp01
        mp11 <- m * n1 * n2 * p1 * p2 + m^2 * n1 * n2 * p1 *
p2
        result[3, 1] <- mp11
        mp20 <- m * n1 * (n1 - 1) * p1^2 + m^2 * n1^2 * p1^2
+ m * n1 * p1
        result[4, 1] <- mp20
        mp02 <- m * n2 * (n2 - 1) * p2^2 + m^2 * n2^2 * p2^2
+ m * n2 * p2
        result[5, 1] <- mp02
        mp21 <- m * n1 * (n1 - 1) * n2 * p1^2 * p2 + m^2 * n1
* (n1 - 1) * n2 *p1^2 * p2 + 2 * m^2 * n1^2 * n2 * p1^2 * p2
+ m^3 * n1^2 * n2 * p1^2 * p2 + m * n1 * n2 * p1 * p2 + m^2 *
n1 * n2 * p1 * p2
        result[6, 1] <- mp21
        mp12 <- m * n1 * (n2 - 1) * n2 * p1 * p2^2 + m^2 * n1
* n2^2 * p1 * p2^2 + m^2 * n1 * n2 * (2 * n2 - 1) * p1 * p2^2
+ m^3 * n1 * n2^2 * p1 * p2^2 + m * n1 * n2 * p1 * p2 + m^2 *
n1 * n2 * p1 * p2
        result[7, 1] <- mp12
        mp22 <- m * (n1 - 1) * n1 * (n2 - 1) * n2 * p1^2 *
p2^2 + m^2 * n1 * (n1 - 1) * n2^2 * p1^2 * p2^2 + m^2 * n1 *
(n1 - 1) * n2 * (2 * n2 - 1) * p1^2 * p2^2 + 2 * m^2 * n1^2 *
n2 * (2 * n2 - 1) * p12 * p2^2 + m^3 * n1 * (n1 - 1) * n2^2 *
p1^2 * p2^2 + m^3 * n1 * (n1 - 1) * n2^2 * p1^2 * p2^2 + 2 *
m^3 * n1^2 * n2^2 * p1^2 * p2^2 + m^3 * n1^2 * n2 * (3 * n2 -
1) * p1^2 * p2^2 + m^4 * n1^2 * n2^2 * p1^2 * p2^2 + mp21 +
mp12 - mp11
        result[8, 1] <- mp22
        m11 <- m * n1 * n2 * p1 * p2
        result[9, 1] <- m11
        m21 <- mp21 - mp01 * mp20 + 2 * mp10^2 * mp01 - 2 *
mp10 * mp11
        result[10, 1] <- m21
        m12 <- mp12 - mp10 * mp02 + 2 * mp10 * mp01^2 - 2 *
mp01 * mp11
        result[11, 1] <- m12
        m22 <- mp22 - 2 * mp21 * mp01 - 2 * mp12 * mp10 +
mp10^2 * mp02 + mp01^2 * mp20 + 4 * mp11 * mp10 * mp01 - 3 *
mp10^2 * mp01^2
        result[12, 1] <- m22
        result
}
```

(14) vmom: Purpose of this program is to calculate numerical values of the variances and covariances of the method of moments estimators. Here n is the total sample size.

```
function(n, m, n1, n2, p1, p2)
{
        mm <- mom(m, n1, n2, p1, p2)
        vcv <- matrix(0, 3, 3)
        tt <- matrix(0, 3, 3)
        result <- matrix(0, 3, 3)
        tt[1, 1] <- n1 * p1
        tt[1, 2] <- n1 * m
        tt[1, 3] <- 0
        tt[2, 1] <- n2 * p2
        tt[2, 2] <- 0
        tt[2, 3] <- n2 * m
        tt[3, 1] <- n1 * n2 * p1 * p2
        tt[3, 2] <- n1 * n2 * p2 * m
        tt[3, 3] <- n1 * n2 * p1 * m
        vcv[1, 1] <- (mm[4, 1] - mm[1, 1]^2)/n
        vcv[1, 2] <- (mm[3, 1] - mm[1, 1] * mm[2, 1])/n
        vcv[1, 3] <- mm[10, 1]/n
        vcv[2, 1] <- vcv[1, 2]
        vcv[2, 2] <- (mm[5, 1] - mm[2, 1]^2)/n
        vcv[2, 3] <- mm[11, 1]/n
        vcv[3, 1] <- vcv[1, 3]
        vcv[3, 2] <- vcv[2, 3]
      vcv[3, 3] <- (mm[12, 1] - mm[9, 1]^2)/n
        result <- solve(tt) %*% vcv %*% t(solve(tt))
        result
}
```

(15) vzero1: Following program is used to get the numerical values of the variances and covariances of the zero-zero cell frequency method estimators.

```
function(n, m, n1, n2, p1, p2)
{
        mm <- mom(m, n1, n2, p1, p2)
        vcv <- matrix(0, 3, 3)
        tt <- matrix(0, 3, 3)
        result <- matrix(0, 3, 3)
        tt[1, 1] <- n1 * p1
        tt[1, 2] <- n1 * m
        tt[1, 3] <- 0
        tt[2, 1] <- n2 * p2
        tt[2, 2] <- 0
        tt[2, 3] <- n2 * m
        tt[3, 1] <- ((1 - p1)^n1 * (1 - p2)^n2 - 1) * exp(m *
(1 - p1)^n1 * (1 - p2)^n2 - m)
        tt[3, 2] <-  - m * n1 * (1 - p2)^n2 * (1 - p1)^(n1 -
1) * exp(m * (1 - p1)^n1 * (1 - p2)^n2 - m)
        tt[3, 3] <-  - m * n2 * (1 - p1)^n1 * (1 - p2)^(n2 -
1) * exp(m * (1 - p1)^n1 * (1 - p2)^n2 - m)
        vcv[1, 1] <- (mm[4, 1] - mm[1, 1]^2)/n
        vcv[1, 2] <- (mm[3, 1] - mm[1, 1] * mm[2, 1])/n
```

```
        vcv[1, 3] <-   - (mm[1, 1]/n) * exp(m * (1 - p1)^n1 *
(1 - p2)^n2 - m)
        vcv[2, 1] <- vcv[1, 2]
   vcv[2, 2] <- (mm[5, 1] - mm[2, 1]^2)/n
        vcv[2, 3] <-   - (mm[2, 1]/n) * exp(m * (1 - p1)^n1 *
(1 - p2)^n2 - m)
        vcv[3, 1] <- vcv[1, 3]
        vcv[3, 2] <- vcv[2, 3]
          vcv[3, 3] <- (1/n) * exp(m * (1 - p1)^n1 * (1 -
p2)^n2 - m) * (1 - exp(m * (1 - p1)^n1 * (1 - p2)^n2 - m))
        result <- solve(tt) %*% vcv %*% t(solve(tt))
        result
}
```

(16) det: This program is used to get the determinant of a 3 by 3 matrix. Here x is the name of the matrix whose determinant is needed.

```
function(x)
{
        det <- x[1, 1] * (x[2, 2] * x[3, 3] - x[2, 3] * x[3,
2]) - x[1, 2] * (x[2, 1] * x[3, 3] - x[2, 3] * x[3, 1]) +
x[1, 3] * (x[2, 1] * x[3,2] - x[2, 2] * x[3, 1])
}
```

(17) comp: This program is used to get the variance covariance matrix of the maximum likelihood estimators, along with their respective determinants. Resultant vector consists of the parameter values and the determinant of the the corresponding variance covariance matrix.

```
function(n, m, n1, n2, p1, p2, r1, r2)
{
        result <- matrix(0, 1, 6)
        a <- matrix(0, 3, 3)
        a <- solve(try(n, m, n1, n2, p1, p2, r1, r2))
        dd1 <- det(a)
        print(a)
        result[, 1] <- m
        result[, 2] <- n1
        result[, 3] <- n2
        result[, 4] <- p1
        result[, 5] <- p2
        result[, 6] <- dd1
        result
}
```

(18) dvmom: Purpose of this program is the same as of program (17), except this program results in the determinant of the variance covariance matrix of the method of moments estimators.

```
function(n, m, n1, n2, p1, p2)
{
        result <- matrix(0, 1, 6)
        a <- matrix(0, 3, 3)
        a <- vmom(n, m, n1, n2, p1, p2)
        dd <- det(a)
        print(a)
        result[, 1] <- m
        result[, 2] <- n1
        result[, 3] <- n2
        result[, 4] <- p1
        result[, 5] <- p2
        result[, 6] <- dd
        result
}
```

(19) dvzero1: It serves the same purpose as program (18) for the estimators obtained from zero-zero cell frequency method.

```
function(n, m, n1, n2, p1, p2)
{
        result <- matrix(0, 1, 6)
        a <- matrix(0, 3, 3)
        a <- vzero1(n, m, n1, n2, p1, p2)
        dd <- det(a)
        print(a)
        result[, 1] <- m
        result[, 2] <- n1
        result[, 3] <- n2
        result[, 4] <- p1
        result[, 5] <- p2
        result[, 6] <- dd
        result
}
```

(20) fit1: This program is used to get σ^2 for the test of goodness of fit while using probability generating function technique when estimates of parameters are obtained. This technique is defined in Kocherlakota and Kocherlakota (1992, p.49).

```
function(n, m, n1, n2, p1, p2, j, k)
{
        dpi <- matrix(0, 3, 3)
```

```
        sigij <- solve(try(n, m, n1, n2, p1, p2, j, k))
        q1 <- 1 - p1
        q2 <- 1 - p2
        t1 <- 0.1
        t2 <- 0.1
        aa1 <- (q1 + p1 * t1)
        aa2 <- (q2 + p2 * t2)
        pii <- exp(m * ((q1 + p1 * t1)^n1 * (q2 + p2 * t2)^n2
- 1))
        pii2 <- pii^2
        pii22 <- exp(m * ((q1 + p1 * t1^2)^n1 * (q2 + p2 *
t2^2)^n2 - 1))
        dpi[1, 1] <- pii^2 * (-1 + aa1^n1 * aa2^n2)^2
        dpi[1, 2] <- pii^2 * (-1 + aa1^n1 * aa2^n2) * n1 * m
* aa1^(n1 - 1) * aa2^n2 * (t1 - 1)
        dpi[1, 3] <- pii^2 * (-1 + aa1^n1 * aa2^n2) * n2 * m
* aa1^n1 * aa2^(n2 - 1) * (t2 - 1)
        dpi[2, 1] <- dpi[1, 2]
        dpi[2, 2] <- pii^2 * n1^2 * m^2 * (t1 - 1)^2 * aa2^(2
* n2) * aa1^(2 * n1 - 2)
        dpi[2, 3] <- pii^2 * n1 * n2 * m^2 * (t1 - 1) * (t2 -
1)* aa1^(2 * n1 - 1) * aa2^( 2 * n2 - 1)
        dpi[3, 1] <- dpi[1, 3]
        dpi[3, 2] <- dpi[2, 3]
        dpi[3, 3] <- pii^2 * n2^2 * m^2 * (t2 - 1)^2 * aa1^(2
* n1) * aa2^(2 * n2 - 2)
        dpi
        s1 <- sum(sigij * dpi)
        sig2 <- ((pii22 - pii2)/n) - s1
        print(dpi)
        print(s1)
        print(pii2)
        print(pii22)
        print(sig2)
}
```

(21) fit2: This program calculates σ^2 for pgf technique when parameters are known, in case of Poisson-double binomial distribution.

```
function(n, m, n1, n2, p1, p2, j, k)
{
        q1 <- 1 - p1
        q2 <- 1 - p2
        t1 <- 0.1
        t2 <- 0.1
        pii <- exp(m * ((q1 + p1 * t1)^n1 * (q2 + p2 * t2)^n2
- 1))
        pii2 <- pii^2
        pii22 <- exp(m * ((q1 + p1 * t1^2)^n1 * (q2 + p2 *
t2^2)^n2 - 1))
        sig2 <- ((pii22 - pii2)/n)
```

```
        print(pii2)
        print(pii22)
        print(sig2)
}
```

(22) pgf1: It calculates the numerical value of the pgf for the Poisson-double binomial distribution when parametric values are supplied. we need this in applying pgf technique in test of goodness-of-fit.

```
function(m, n1, n2, p1, p2)
{
        q1 <- 1 - p1
        q2 <- 1 - p2
        t1 <- 0.1
        t2 <- 0.1
        pii <- exp(m * ((q1 + p1 * t1)^n1 * (q2 + p2 * t2)^n2
- 1))
}
```

(23) epgf1: This calculates numerical value of the empirical (sample) pgf.

```
function(n, j, k)
{
        freq <- nrs
        t1 <- 0.1
        t2 <- 0.1
        cs <- 0
        for(r in 1:j) {
                for(s in 1:k) {
                        cs <- cs + (freq[r, s]/n) * t1^(r -
1) * t2^(s - 1)
}
}
cs
}
```

(24) pfbvp1: This program calculates the bivariate Poisson probabilities when parametric values are supplied.

```
function(lam, p1, p2, j, k)
{
        ff <- matrix(0, j, k)
        q1 <- 1 - p1
        q2 <- 1 - p2
        ff[1, 1] <- exp(lam * (q1 * q2 - 1))
        for(r in 2:j) {
```

```
                ps <- (exp(lam * (q1 * q2 - 1)) * p1^(r - 1)
* q2^(r - 1) * lam^(r - 1))/(gamma(r))
        ff[r, 1] <- ps
}
        for(s in 2:k) {
                ps1 <- (exp(lam * (q1 * q2 - 1)) * p2^(s - 1)
* q1^(s - 1) * lam^(s - 1))/(gamma(s))
        ff[1, s] <- ps1
}
        for(r in 2:j) {
                for(s in 2:k) {
                        ps2 <- 0
                        ff[1, 1] <- exp(lam * (q1 * q2 - 1))
                        for(rr in 0:min(r - 1, s - 1)) {
                                ps2 <- ps2 + (p1^(r - 1) *
p2^(s - 1) * q1^(s - 1 - rr) * q2^(r - 1 - rr) * lam^(r + s -
rr - 2))/(gamma(r - rr) * gamma(s - rr) * gamma(rr + 1))
}
ff[r, s] <- ff[1, 1] * ps2
}
}
ff
}
```

(25) dlbvp1: This program finds the numerical values of the derivatives of the bivariate Poisson probability function with respect to λ.

```
function(lam, p1, p2, j, k)
{
        ff <- pfbvp1(lam, p1, p2, j, k)
        dfl <- matrix(0, j, k)
        q1 <- 1 - p1
        q2 <- 1 - p2
        for(r in 1:j) {
                for(s in 1:k) {
                        ps <- 0
                        for(rr in 0:min(r - 1, s - 1)) {
                                ps <- ps + (rr * p1^(r - 1) *
p2^(s - 1) * q2^(s - 1 - rr) * q2^(r - rr - 1) * lam^(r + s -
rr - 2))/(gamma(r - rr) * gamma(s - rr) * gamma(rr + 1))
}
dfl[r, s] <- (q1 * q2 - 1) * ff[r, s] + ((r + s - 2)/lam) *
ff[r, s] - (exp(lam * (q1 * q2 - 1))/lam) * ps
}
}
dfl
}
```

(26) dp1bvp1: This program calculates the numerical values of the derivatives of bivariate Poisson probability function with respect to p_1.

```
function(lam, p1, p2, j, k)
{
        ff <- pfbvp1(lam, p1, p2, j, k)
        dfp1 <- matrix(0, j, k)
        q1 <- 1 - p1
        q2 <- 1 - p2
        for(r in 1:j) {
                for(s in 1:k) {
                        ps <- 0
                        for(rr in 0:min(r - 1, s - 1)) {
                                ps <- ps + (rr * p1^(r - 1) *
p2^(s - 1) * q2^(s - 1 - rr) * q2^(r - rr - 1) * lam^(r + s -
rr - 2))/(gamma(r - rr) * gamma(s - rr) * gamma(rr + 1))
}
        dfp1[r, s] <-  - lam * q2 * ff[r, s] + ((r - 1)/p1) *
ff[r, s] - ((s - 1)/q1) * ff[r, s] + (exp(lam * (q1 * q2 -
1))/q1) * ps
}
}
dfp1
}
```

(27) dp2bvp1: Purpose of this program is same as of (25) except the derivatives of pf are with respect to p_2.

```
function(lam, p1, p2, j, k)
{
        ff <- pfbvp1(lam, p1, p2, j, k)
        dfp2 <- matrix(0, j, k)
        q1 <- 1 - p1
        q2 <- 1 - p2
        for(r in 1:j) {
                for(s in 1:k) {
                        ps <- 0
                        for(rr in 0:min(r - 1, s - 1)) {
                                ps <- ps + (rr * p1^(r - 1) *
p2^(s - 1) * q2^(s - 1 - rr) * q2^(r - rr - 1) * lam^(r + s -
rr - 2))/(gamma(r - rr) * gamma(s - rr) * gamma(rr + 1))
}
dfp2[r, s] <-  - lam * q1 * ff[r, s] + ((s - 1)/p2) * ff[r,
s] - ((r - 1)/q2) * ff[r, s] + (exp(lam * (q1 * q2 - 1))/q2)
* ps
}
}
dfp2
}
```

(28) gambvp1: This program finds the information matrix for the bivariate Poisson distribution.

```
function(n, lam, p1, p2, j, k)
{
        ff <- pfbvp1(lam, p1, p2, j, k)
        dl <- dlbvp1(lam, p1, p2, j, k)
        dp1 <- dp1bvp1(lam, p1, p2, j, k)
        dp2 <- dp2bvp1(lam, p1, p2, j, k)
        gamt <- matrix(0, 3, 3)
        cs <- 0
        for(r in 1:j) {
                for(s in 1:k) {
                        cs <- cs + (1/ff[r, s]) * dl[r, s] *
dl[r, s]
}}
        gamt[1, 1] <- n * cs
        cs <- 0
        for(r in 1:j) {
                for(s in 1:k) {
                        cs <- cs + (1/ff[r, s]) * dp1[r, s] *
dp1[r, s]
}}
        gamt[2, 2] <- n * cs
        cs <- 0
        for(r in 1:j) {
                for(s in 1:k) {
                        cs <- cs + (1/ff[r, s]) * dp2[r, s] *
dp2[r, s]
}}
        gamt[3, 3] <- n * cs
        cs <- 0
        for(r in 1:j) {
                for(s in 1:k) {
                        cs <- cs + (1/ff[r, s]) * dl[r, s] *
dp1[r, s]
}}
        gamt[1, 2] <- n * cs
        gamt[2, 1] <- gamt[1, 2]
        cs <- 0
        for(r in 1:j) {
                for(s in 1:k) {
                        cs <- cs + (1/ff[r, s]) * dl[r, s] *
dp2[r, s]
}}
        gamt[1, 3] <- n * cs
        gamt[3, 1] <- gamt[1, 3]
        cs <- 0
        for(r in 1:j) {
                for(s in 1:k) {
                        cs <- cs + (1/ff[r, s]) * dp1[r, s] *
dp2[r, s]
```

```
}}
        gamt[2, 3] <- n * cs
        gamt[3, 2] <- gamt[2, 3]
        gamt
}
```

(29) dsbvp1: This program is written to get the D matrix, which is needed to obtain the iteration vector for ML estimates for the parameters of the bivariate Poisson distribution.

```
function(lam, p1, p2, j, k)
{
        ff <- pfbvp1(lam, p1, p2, j, k)
        dl <- dlbvp1(lam, p1, p2, j, k)
        dp1 <- dp1bvp1(lam, p1, p2, j, k)
        dp2 <- dp2bvp1(lam, p1, p2, j, k)
        fr <- death
        dss <- matrix(0, 3, 1)
        dss[1, 1] <- sum((fr) * (1/ff) * dl)
        dss[2, 1] <- sum((fr) * (1/ff) * dp1)
        dss[3, 1] <- sum((fr) * (1/ff) * dp2)
        dss
}
```

(30) pgfbvp: To apply pgf technique for goodness-of-fit for bivariate Poisson distribution, we need to calculate the numerical value of the pgf when parametric values are supplied. This program serves this purpose.

```
function(m, p1, p2)
{
        q1 <- 1 - p1
        q2 <- 1 - p2
        t1 <- 0.1
        t2 <- 0.1
        pii <- exp(m * (q1 * q2 - 1 + p1 * q2 * t1 + p2 * q1
* t2 + p1 * p2 * t1 * t2))
}
```

(31) epgfbvp: Thids priogram calculates the numerical value of the empirical pgf.

```
function(n, j, k)
{
        freq <- death
        t1 <- 0.1
        t2 <- 0.1
        cs <- 0
        for(r in 1:j) {
                for(s in 1:k) {
```

```
                        cs <- cs + (freq[r, s]/n) * t1^(r -
1) * t2^(s - 1)
}
}
cs
}
```

(32) fitbvp1: This program finds σ^2 for the pgf technique (when estimates of the parameters are used) for bivaraite Poisson distribution.

```
function(n, m, p1, p2, j, k)
{
        dpi <- matrix(0, 3, 3)
        sigij <- solve(gambvp1(n, m, p1, p2, j, k))
        q1 <- 1 - p1
        q2 <- 1 - p2
        t1 <- 0.1
        t2 <- 0.1
        pii <- exp(m * (q1 * q2 - 1 + p1 * q2 * t1 + p2 * q1
* t2 + p1 * p2 * t1 * t2))
        pii2 <- pii^2
        pii22 <- exp(m * (q1 * q2 - 1 + p1 * q2 * t1^2 + p2 *
q1 * t2^2 + p1 * p2 * t1^2 * t2^2))
        dpi[1, 1] <- pii^2 * (q1 * q2 - 1 + p1 * q2 * t1 + p2
* q1 * t2 + p1 * p2 * t1 * t2)^2
        dpi[1, 2] <- pii^2 * (q1 * q2 - 1 + p1 * q2 * t1 + p2
* q1 * t2 + p1 * p2 * t1 * t2) * ( - m * q2 + q2 * t1 - p2 *
t2 + p2 * t1 * t2)
        dpi[1, 3] <- pii^2 * (q1 * q2 - 1 + p1 * q2 * t1 + p2
* q1 * t2 + p1 * p2 * t1 * t2) * ( - m * q1 + q1 * t2 - p1 *
t1 + p1 * t1 * t2)
        dpi[2, 1] <- dpi[1, 2]
        dpi[2, 2] <- pii^2 * ( - m * q2 + q2 * t1 - p2 * t2 +
p2 * t1 * t2)^2
        dpi[2, 3] <- pii^2 * ( - m * q2 + q2 * t1 - p2 * t2 +
p2 * t1 * t2) * (- m * q1 + q1 * t2 - p1 * t1 + p1 * t1 * t2)
        dpi[3, 1] <- dpi[1, 3]
        dpi[3, 2] <- dpi[2, 3]
        dpi[3, 3] <- pii^2 * ( - m * q1 + q1 * t2 - p1 * t1 +
p1 * t1 * t2)^2
        dpi
        s1 <- sum(sigij * dpi)
        sig2 <- ((pii22 - pii2)/n) - s1
        print(dpi)
        print(s1)
        print(pii2)
        print(pii22)
        print(sig2)
}.
```

Bibliography

Chatfield, C. and Theobald, C. M. (1973). Mixtures and random sums. *The Statistician,* **22**, 281-287.

Cochran, W. G. (1954). Some methods for strengthening the common χ^2 tests. , *Biometrics*, **10**, 417-451.

Douglas, J. B. (1970). "Statistical models in discrete distributions." In: Random Counts in Scientific Work. Vol. 3, 203-232. Ed. G. P. Patil. University Park: Penn. State University Press.

Douglas, J. B. (1980). *Analysis with Standard Contagious Distributions,* Vol. 4. International Co-operative Publishing House, Fairland, Maryland.

Feller, W. (1943). "On a general class of contagious distributions." Ann. Math. Statist., **14**, 389-400.

Feller, W. (1968). *An introduction to Probability Theory and its Applications* (third edition). Vol.1. New York: Wiley.

Gurland, J. (1957). Some interrelations among compound and generalized distributions. *Biometrika*, **44**, 265-268.

Johnson, N. L., Kotz, S. and Kemp, A. W. (1992). *Univariate Discrete Distributions*, Second Edition. John Wiley and Sons, New York.

Katti, S. K., and Gurland, J. (1962). Some Methods of Estimation for the Poisson-binomial Distribution, *Biometrics*, **18**, 42-51.

Kemp, C. D. and Papageorgiou, H. (1983). Bivariate Hermite distribution. *Sankhya*, **44**, 269-280.

Kocherlakota, K. and Kocherlakota, S. (1986). Goodness of fit tests for discrete distributions. *Communications in Statistics - Theory and Methods*, **15**, 815-829.

Kocherlakota, K. and Kocherlakota, S. (1992). *Bivariate discrete distributions*. Marcel Dekker, Inc., New York.

Loukas, S. and Kemp, C. D. (1986). On the chi-square goodness-of-fit statistics for bivariate discrete distributions. The *Statistician*, **35**, 525-529.

Rao, C. R. (1973). *Linear Statistical inference and its applications* (Second edition), John Wiley and Sons, New York.

S-Plus (1995). MathSoft, Inc., Seattle, Washington.

Sprott, D. A. (1958). The Method of Maximum Likelihood Applied to the Poisson-binomial Distribution, *Biometrics*, **14**, 97-106.

Venables, W. N. and Ripley, B. D. (1996). *Modern Applied Statistics with S-plus*. Springer.