Statistics & Probability Letters 60 (2002) 231–240

STATISTICS & PROBABILITY LETTERS

www.elsevier.com/locate/stapro

A new bivariate binomial distribution

Atanu Biswas[a,*], Jing-Shiang Hwang[b]

[a] *Applied Statistics Unit, Indian Statistical Institute, 203 B.T. Road, Calcutta 700035, India*

[b] *Institute of Statistical Science, Academia Sinica, Nan-Kang, Taipei 115, Taiwan, ROC*

Received June 2001; received in revised form August 2002

Abstract

To bring correlation between binomial random variables is an important statistical problem with a lot of theoretical and practical applications. In this paper we provide a new formulation of bivariate binomial distribution in the sense that marginally each of the two random variables has a binomial distribution and they have some non-zero correlation in the joint distribution. A 2×2 contingency table is the immediate application of the proposed model.

Keywords: Bivarite binomial distribution; 2×2 contingency table; Trinomial distribution; Quasi-binomial distribution; Maximum likelihood estimate; Odds ratio

1. Introduction

This short article is motivated to provide some kind of solution to what may be an elementary statistical problem—to give a joint distribution of two marginally binomially distributed random variables. Quite often, in practice, it may be reasonable to assume a correlation between two binomial random variables. A 2×2 contingency table can be an obvious motivation of the present paper and that may be an immediate application too. Although the problem can be generalized to a multivariate setup, in the present paper we focus our attention to the bivariate version only.

A large part of the literature on contingency tables have been concerned with measuring the degree of association between two dichotomized quantitative characters. Pearson (1900) introduced the tetrachoric correlation as an estimate of the correlation coefficient in a bivariate normal model underlying the 2×2 table. For 2×2 tables resulting from such discrete characters with no margin fixed, the exact model is the four-cell multinomial distribution (see Fisher, 1950). A bivariate binomial

* Corresponding author. Tel.: +91-33-577-8085x2818; fax: +91-33-577-6680.

E-mail addresses: atanu@isical.ac.in (A. Biswas), jshwang@stat.sinica.edu.tw (J.-S. Hwang).

PII: S0167-7152(02)00323-1

distribution was attempted mainly to explain a 2×2 contingency tables formed in terms of two characters A and B. Writing the probabilities of the four cells AB, AB^c, A^cB and A^cB^c as p_1, p_2, p_3 and $p_4 = 1 - p_1 - p_2 - p_3$, and fixing only the total sample size combining the four cells as n, a bivariate binomial distribution was first introduced by Aitken and Gonin (1935) in connection with a fourfold sampling scheme with replacement. The distribution and its properties have been studied in detail by Hamdan in a series of papers. In fact, this model is referred as the bivariate binomial distribution in Patil and Joshi (1968). The canonical representation of the probability function of this model can be given as a series bilinear in Krawtchouk's polynomials (see Szego, 1959). See also Hamdan (1972) in this context. Note that when $p_3 = 0$, the distribution reduces to the standard trinomial distribution (see Mardia, 1970, p. 82). Hamdan (1975) has considered the relationship of this distribution with the trinomial distribution. Also Marshall and Olkin (1985) related the genesis of the distribution with the trinomial. In fact, the probability function of the model given by Aitken and Gonin (1935) can be written as a sum of trinomial probability functions (see Hamdan and Nasro, 1986). The model is then studied in different directions by several authors. For example, Hamdan and Martinson (1971) derived the maximum likelihood estimates, conditional distributions are studied by Kocherlakota (1989) in the most general form of the bivariate binomial distribution, and Hamdan and Jensen (1976) discussed some possible applications. Loukas and Kemp (1986) and Oluyede (1994) also considered the same distribution of Aitken and Gonin. Ong (1992) suggested some mixture models to derive the distribution of Aitken and Gonin (1935).

There are also some attempts of finding bivariate binomial distributions in other directions. Hamdan and Tsokos (1971) introduced a bivariate binomial distribution (which is, actually, a bivariate compound Poisson distribution). A symmetric bivariate binomial distribution was proposed by Le (1984) to analyze clustered samples in medical research. Papageorgiou and David (1994) examined mixtures of bivariate binomial distributions which were derived from bivariate-compounded Poisson distribution. Ling and Tai (1990) discussed bivariate binomial distributions from extension of classes of univariate discrete distributions of order k. Takeuchi and Takemura (1987) obtained the sum of 0–1 random variables in the multivariate setup. A bivariate generalization of the three parameter quasi-binomial distribution of Consul (1974) has been obtained by Mishra (1996). Crowder and Sweeting (1989) carried out Bayesian inference, and they defined the bivariate binomial distribution in a different sense. They defined a two-fold binomial model like $Y_1|m \sim \mathrm{Bin}(m, p)$ and $Y_2|Y_1, m \sim \mathrm{Bin}(Y_1, q)$. A different kind of prior in the same setting were derived by Polson and Wasserman (1990). For some discussions on bivariate binomial distributions see Kocherlakota and Kocherlakota (1992) and Johnson et al. (1997).

None of the existing works deal with the following simple problem. Suppose that one index population of size n_1 exposed to A and one other index population of size n_2 exposed to B, where both n_1 and n_2 are known constants. Let Y_1 and Y_2 be two jointly distributed random outcome variables such that marginally Y_i, $i = 1, 2$, follows a binomial distribution with parameters (n_i, p_i), where n_i's are fixed, and there is correlation between Y_1 and Y_2 due to some other common exposures of the two populations.

No solution of the problem is available in the statistical literature, so far. One possible approach is to consider a random component b (may be a vector) and to write p_i as a function of b. Then, conditionally on b, Y_i's are independent binomials. Correlation will be incorporated if we integrate the joint distribution of Y_1 and Y_2 over b. But, unfortunately the unconditional marginals of Y_i's will never remain binomials. Moreover, in that case, due to the common effect b, the correlation always becomes positive. Again, too many parameters in that model makes the estimation problem more complicated. Here we propose a model by assuming $n_2 \geqslant n_1$. The other case can be similarly dealt with. In this present paper we propose the following probability model:

$$\Pr(Y_1 = y_1, Y_2 = y_2) = \binom{n_1}{y_1} p_1^{y_1}(1-p_1)^{n_1-y_1} \times f(y_2|y_1), \tag{1.1}$$

where

$$\begin{aligned} f(y_2|y_1) = (1+\alpha)^{-n_1} \sum_{(j_1,j_2,j_3)\in S} & \binom{y_1}{j_1}\binom{n_1-y_1}{j_2}\binom{n_2-n_1}{j_3}\{p_2+\alpha(p_2-p_1)+\alpha\}^{j_1} \\ & \times\{1-p_2-\alpha(p_2-p_1)\}^{y_1-j_1}\{p_2+\alpha(p_2-p_1)\}^{j_2} \\ & \times\{1-p_2-\alpha(p_2-p_1)+\alpha\}^{n_1-y_1-j_2} p_2^{j_3}(1-p_2)^{n_2-n_1-j_3}, \end{aligned} \tag{1.2}$$

with $S=\{(j_1,j_2,j_3): j_1+j_2+j_3=y_2; j_1=0,1,\ldots,y_1; j_2=0,1,\ldots,n_1-y_1; j_3=0,1,\ldots,n_2-n_1\}$. Although the model is complicated in its expression, it can be derived from a simple conditioning mechanism. From the discussions of Section 2, it will follow that for model (1.1) and (1.2), marginally Y_1 and Y_2 are simple binomials, and they have a correlation coefficient

$$\rho = \sqrt{\frac{m}{M}}\left(\frac{\alpha}{1+\alpha}\right)\sqrt{\frac{p_1(1-p_1)}{p_2(1-p_2)}}, \tag{1.3}$$

where $m = \min(n_1, n_2)$ and $M = \max(n_1, n_2)$. Note that, if $f(y_2|y_1) = \binom{n_2}{y_2} p_2^{y_2}(1-p_2)^{n_2-y_2}$, (1.1) represents the product of two independent binomial probability functions, which is obtained for $\alpha = 0$ in (1.2). The simple case of common p_i is also discussed as a special case. Again, it will be quite easy to simulate samples from the bivariate distribution. The theoretical expression of the joint probability generating function is also derived. In Section 3, we obtain the maximum likelihood estimates of the parameters. Finally, in Section 4 we provide some concluding remarks.

2. The model

First we consider the simple case where both Y_1 and Y_2 follow binomial with parameters $(n, p = 0.5)$. In this case we can define $Y_2 = Y_1$ with probability α and $Y_2 = n - Y_1$ with probability $(1-\alpha)$, and set α accordingly to get a desired level of correlation. But, given Y_1, in this case Y_2 can take at most two distinct values, Y_1 and $n - Y_1$. As a generalization where $Y_1 \sim \text{Bin}(n_1, 0.5)$ and $Y_2 \sim \text{Bin}(n_2, 0.5)$ with $n_1 < n_2$, we can set $Y_2 = Y_1 + W$ with probability α and $Y_2 = n - Y_1 + W$ with probability $1-\alpha$, where $W \sim \text{Bin}(n_2 - n_1, 0.5)$, independent of Y_1. In this case also, given Y_1, Y_2 cannot take all possible values $0, 1, \ldots, n_2$. Again, even with such restrictions, the approach will be intractable if Y_1 and Y_2 have common success probabilities other than 0.5. The situation will be more difficult if $p_1 \neq p_2$. In fact, we need one model which will allow all the $(n_1+1)(n_2+1)$ possible combinations of (Y_1, Y_2) with some desired correlation.

We present the model in the following manner. We represent $Y_i = \sum_{j=1}^{n_i} Y_{ij}$, $i = 1, 2$, where for a given i, Y_{ij}'s will be marginally independently and identically distributed (i.i.d.) Bernoulli (p_i) random variables. We, at first, generate Y_{1j}'s, $j = 1, \ldots, n_1$, as i.i.d. Bernoulli (p_1). Then, for $j = 1, \ldots, m$, we generate Y_{2j} in the following way. Here Y_{2j} will depend only on Y_{1j} such that

$$\Pr(Y_{2j} = 1 | Y_{1j}) = \frac{p_2 + \alpha(p_2 - p_1) + \alpha Y_{1j}}{1 + \alpha}, \tag{2.1}$$

for some α. Clearly, taking expectation in (2.1), the unconditional probability distribution of Y_{2j}, $j = 1, \ldots, m$, will be Bernoulli (p_2). If $n_2 > m$, we generate $Y_{2,m+1}, \ldots, Y_{2n_2}$ as i.i.d. Bernoulli (p_2). The parameter α can be interpreted as the mixing parameter. Clearly, the value of α must be such that the right hand side of (2.1) belongs to [0,1] for both $Y_{1j} = 0, 1$. Thus, for $p_1 > p_2$, the possible range of α is $[-((1 - p_2)/(1 + p_1 - p_2)), p_2/(p_1 - p_2)]$; and for $p_2 > p_1$, the possible range is $[-((1 - p_2)/(1 + p_1 - p_2)), (1 - p_2)/(p_2 - p_1)]$. In the very special case when $p_1 = p_2 = p$, we have $\alpha \in (-(1 - p), \infty)$.

From (2.1), we have for $j = 1, \ldots, m$,

$$\begin{aligned} E(Y_{1j}Y_{2j}) = E\{Y_{1j}E(Y_{2j}|Y_{1j})\} &= \left(\frac{p_2 + \alpha(p_2 - p_1) + \alpha}{1 + \alpha}\right) E(Y_{1j}) \\ &= \left(\frac{p_2 + \alpha(p_2 - p_1) + \alpha}{1 + \alpha}\right) p_1, \end{aligned}$$

as $Y_{1j}^2 = Y_{1j}$. Consequently, we have the covariance between Y_{1j} and Y_{2j}, $j = 1, \ldots, m$, as $(\alpha/(1 + \alpha))p_1(1 - p_1)$ resulting the covariance between Y_1 and Y_2 being

$$m\left(\frac{\alpha}{1 + \alpha}\right) p_1(1 - p_1).$$

Hence, the correlation coefficient between Y_1 and Y_2 will be as given in (1.3). In the simple particular case where $n_1 = n_2$ and $p_1 = p_2$ the correlation has an easy form, namely $\alpha/(1 + \alpha)$. The correlation can be increased towards +1 taking α large enough.

Fig. 1 provides the bivariate distribution with $n_1 = n_2 = 10$, with different (p_1, p_2) values and some possible correlation values (including the case of independence). Fig. 2 provides the corresponding contours. From the figures we observe that the probability distribution with some non-zero correlation changes dramatically from the independent case.

Note that, from the above formulation the simulation of bivariate samples is an easy task. The joint probability distribution (1.1) and (1.2) can be derived from the formulation (2.1). In fact, we derive the joint probability generating function $P(t_1, t_2)$ of (Y_1, Y_2). Note that, from the independence of the pairs $(Y_{11}, Y_{21}), \ldots, (Y_{1m}, Y_{2m})$, we have

$$\begin{aligned} P(t_1, t_2) &= E(t_1^{Y_1} t_2^{Y_2}) \\ &= \left\{\prod_{i=1}^{m} E(t_1^{Y_{1i}} t_2^{Y_{2i}})\right\} \left\{\prod_{i=m+1}^{M} E(t_1^{Y_{1i}})\right\}^{\delta(n_1 - m)} \left\{\prod_{i=m+1}^{M} E(t_2^{Y_{2i}})\right\}^{\delta(n_2 - m)}, \end{aligned}$$

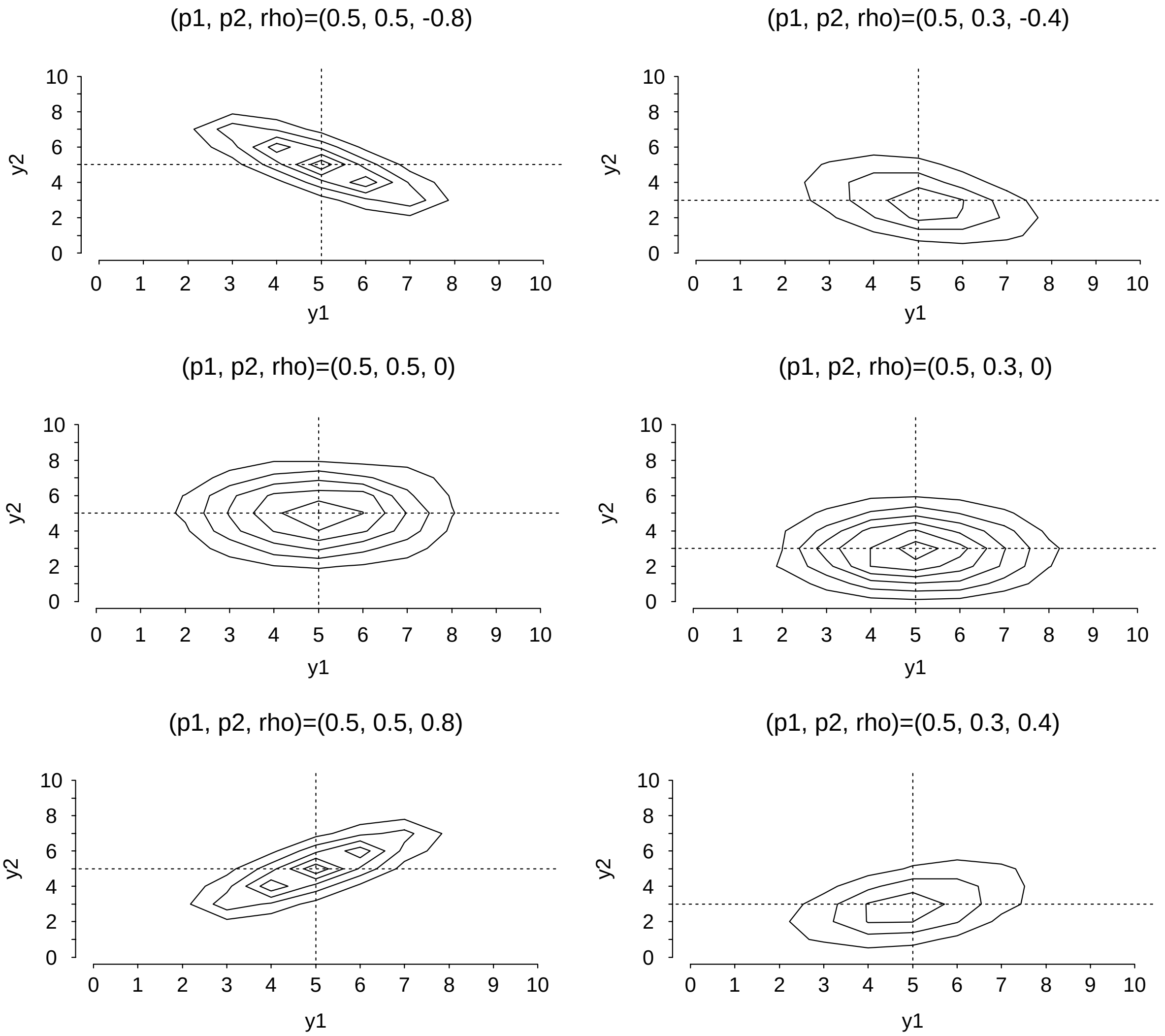

Fig. 1. Bivariate distribution with $n_1 = n_2 = 10$, with different (p_1, p_2) values and some possible correlation values (including the case of independence).

where $\delta(x) = 1$ if $x > 0$, and $\delta(x) = 0$ otherwise. Note that, $E(t_1^{Y_{1i}}) = p_1 t_1 + (1 - p_1)$ for $i > m$. Now,

$$
\begin{aligned}
E(t_1^{Y_{1i}} t_2^{Y_{2i}}) &= E\{t_1^{Y_{1i}} E(t_2^{Y_{2i}} | Y_{1i})\} \\
&= E\left[t_1^{Y_{1i}} \left\{ t_2 \left(\frac{p_2 + \alpha(p_2 - p_1) + \alpha Y_{1i}}{1 + \alpha} \right) + \left(\frac{1 - p_2 - \alpha(p_2 - p_1) + \alpha(1 - Y_{1i})}{1 + \alpha} \right) \right\} \right] \\
&= \left\{ t_1 t_2 \left(\frac{p_2 + \alpha(p_2 - p_1) + \alpha}{1 + \alpha} \right) + t_1 \left(\frac{1 - p_2 - \alpha(p_2 - p_1)}{1 + \alpha} \right) \right\} p_1 \\
&\quad + \left\{ t_2 \left(\frac{p_2 + \alpha(p_2 - p_1)}{1 + \alpha} \right) + \left(\frac{1 - p_2 - \alpha(p_2 - p_1) + \alpha}{1 + \alpha} \right) \right\} (1 - p_1).
\end{aligned}
$$

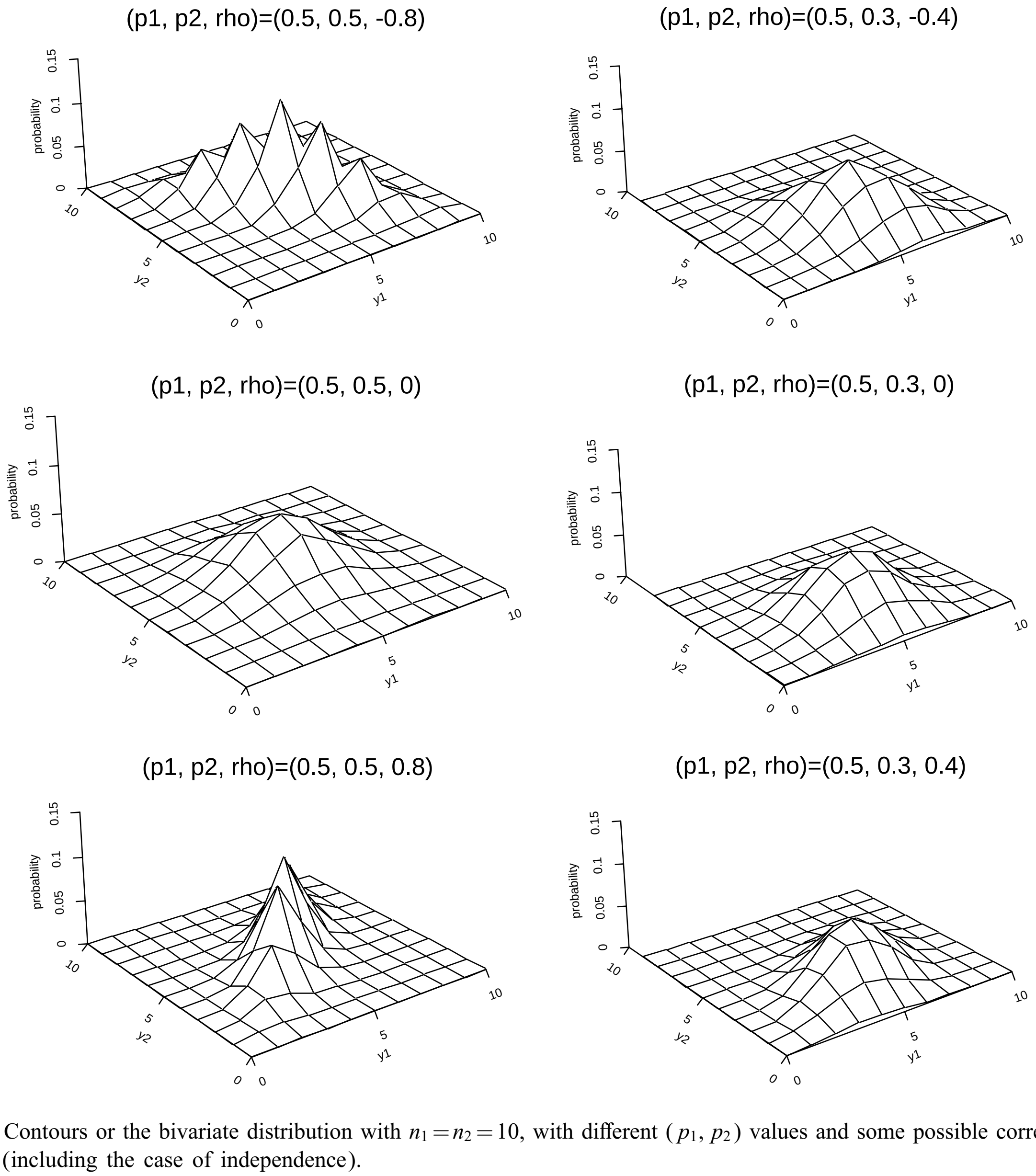

Fig. 2. Contours or the bivariate distribution with $n_1 = n_2 = 10$, with different (p_1, p_2) values and some possible correlation values (including the case of independence).

Consequently, we have

$$P(t_1,t_2) = \left[\left\{t_1t_2\left(\frac{p_2+\alpha(p_2-p_1)+\alpha}{1+\alpha}\right)+t_1\left(\frac{1-p_2-\alpha(p_2-p_1)}{1+\alpha}\right)\right\}p_1\right.$$
$$\left.+\left\{t_2\left(\frac{p_2+\alpha(p_2-p_1)}{1+\alpha}\right)+\left(\frac{1-p_2-\alpha(p_2-p_1)+\alpha}{1+\alpha}\right)\right\}(1-p_1)\right]^m$$
$$\times\{p_1t_1+(1-p_1)\}^{(n_1-n_2)\delta(n_1-m)}\{p_2t_2+(1-p_2)\}^{(n_2-n_1)\delta(n_2-m)}. \tag{2.2}$$

Under independence of Y_1 and Y_2, we have $\alpha = 0$, and consequently (2.2) reduces to

$$\{p_1t_1 + (1-p_1)\}^{n_1}\{p_2t_2 + (1-p_2)\}^{n_2}.$$

The coefficient of $t_1^{y_1}t_2^{y_2}$ in the expansion of (2.2) will provide $\Pr(Y_1 = y_1, Y_2 = y_2)$, which is given by (1.1) and (1.2). For example, when $n_1 \leqslant n_2$, we have

$$\Pr(Y_1 = 0, Y_2 = 0) = \left\{ \left(\frac{1 - p_2 - \alpha(p_2 - p_1) + \alpha}{1+\alpha} \right) (1-p_1) \right\}^{n_1} (1-p_2)^{n_2 - n_1},$$

which reduces to $(1-p_1)^{n_1}(1-p_2)^{n_2}$ for $\alpha = 0$. Marginal moments of Y_1 and Y_2 are those of the corresponding binomial distributions. To get the joint moment generating function we replace t_1 and t_2 in (2.2) by $\exp(t_1)$ and $\exp(t_2)$, respectively. Note that the conditional distribution of any Y_i given the other is not binomial. In fact,

$$\Pr(Y_2 = y_2 | Y_1 = y_1) = f(y_2|y_1),$$

where $f(y_2|y_1)$ is given in (1.2).

3. Estimation of the parameters

If no covariate is considered we do not need any further modeling of the parameters. In that case there are only 3 parameters in model (1.1) and (1.2), namely p_1, p_2 and α. Suppose there are $k(\geqslant 2)$ independent experiments with $n_{1i}, n_{2i}, y_{1i}, y_{2i}$ being the n_1, n_2, y_1, y_2 values in the ith experiment. Then the log-likelihood can be written as

$$\begin{aligned} l = \text{Constant} &+ \left(\sum_{i=1}^{k} y_{1i}\right) \log p_1 + \left(\sum_{i=1}^{k} (n_{1i} - y_{1i})\right) \log(1-p_1) \\ &- \left(\sum_{i=1}^{k} m_i\right) \log(1+\alpha) + \sum_{i=1}^{k} \log U_i, \end{aligned}$$

where $U_i = \sum_{S_i} W_i$ with

$$\begin{aligned} W_i = &\binom{y_{1i}}{j_1}\binom{n_{1i} - y_{1i}}{j_2}\binom{n_{2i} - y_{1i}}{j_3} \\ &\times \{p_2 + \alpha(p_2 - p_1) + \alpha\}^{j_1}\{1 - p_2 - \alpha(p_2 - p_1)\}^{y_{1i} - j_1} \\ &\times \{p_2 + \alpha(p_2 - p_1)\}^{j_2}\{1 - p_2 - \alpha(p_2 - p_1) + \alpha\}^{n_{1i} - y_{1i} - j_2} \\ &\times p_2^{j_3}(1-p_2)^{n_{2i} - n_{1i} - j_3}. \end{aligned}$$

The estimates of the parameters then can be obtained by employing grid search method for p_1 and p_2 over their possible ranges, and, for every possible (p_1, p_2), by solving $\partial l/\partial\alpha = 0$, where

$$\frac{\partial l}{\partial \alpha} = -\frac{\sum m_i}{1+\alpha} + \sum \frac{1}{U_i} \sum_{S_i} W_i \frac{\partial \log W_i}{\partial \alpha}.$$

To illustrate the applicability of the procedure we carry out some computations. Using random numbers in S-Plus we generate two independent 2×2 tables (i.e., $k = 2$) with $n_{11} = 8$, $n_{21} = 11$, $n_{12} = 10$, $n_{22} = 12$ and $p_1 = 0.2$, $p_2 = 0.3$, $\alpha = 1$. The generated $y_{11} = 2$, $y_{12} = 1$, $y_{21} = 0$ and $y_{22} = 1$. The maximum likelihood estimates of p_1, p_2 and α in this case comes out to be 0.22, 0.28 and 0.9749, which are quite close to the actual values. Thus the method works well for this particular observed values. We have also carried out an extensive simulation based computation. In each simulation, taking some fixed values of (p_1, p_2, α), K 2×2 tables with some (n_{1i}, n_{2i}) for the ith table, $i = 1, \ldots, K$, are generated. Then the (p_1, p_2, α)-values are estimated from each simulation, and all the simulated estimates are then averaged. It is observed that simulation results provide good estimates of p_1 and p_2, but the estimates of α are not quite good for small value of K. The estimate of α has an upward bias. The bias reduces with the increase of K. However, if K, the number of tables, is large, say 20 or more, the estimate of α is quite close to the actual value. Note that some other estimating procedures can also be used. After getting the estimates of p_1 and p_2, and noting that

$$\operatorname{cov}\left(\frac{Y_1}{\sqrt{m}}, \frac{Y_2}{\sqrt{m}}\right) = \left(\frac{\alpha}{1+\alpha}\right) p_1(1 - p_1),$$

α can be estimated from the sample covariance of the K available (y_{1i}, y_{2i}) values. But, this approach is also not very efficient for estimating α. Note that, the proposed method is flexible in the real life situations where p_1 and p_2 are functions of covariates.

4. Concluding remarks

There is a basic structural difference of our model with the earlier works. For example, consider the two groups namely smokers (S) and non-smokers (NS). Each group is divided into two parts, occurrence of cancer (C) or no cancer (NC). Thus there are four cells corresponding to S/C, S/NC, NS/C and NS/NC in the 2×2 table. Then the model by Aitken and Gonin (1935) and the following works in that direction, assumes the total number of individuals under study to be fixed, but the numbers under S and NS are allowed to be random. In contrast, our model considers fixed values of the numbers under S and NS. And, quite often, the data may be collected in that fashion. Note that the model of Crowder and Sweeting (1989) is of form (1.1) in some sense, where Y_1 is binomial $(n_1 + n_2, p)$ and $f(y_2|y_1)$ is also a binomial probability function.

From (1.3), we observe that, even when $p_1 = p_2$, the correlation coefficient between Y_1 and Y_2 will never be 1 for large α, when $n_1 \neq n_2$. This has intuitive justification in the sense that, when $n_1 \neq n_2$, we can never express $Y_1 = \beta_0 + \beta_1 Y_2$ with probability 1. The proposed methodology can be extended to obtain bivariate multinomial distribution to explain $2 \times k$ tables. The details are under study.

The present paper is an attempt to provide a solution of the distributional problem. The present authors believe that a lot more theoretical and practical issues are to be taken care of, which may be a subject of future research. One can as well redefine Y_1 and Y_2, and generate any one stating from the other. The expression of the correlation provides the insight of the stating Y_i. But, one can make the expression invariant by choosing α appropriately. Odds ratio is an important feature in such studies, which has widely been used in several experimental and observational studies. The estimate of the odds ratio in the present situation will be same as the case of independence. But its distribution will differ from the usual one for the possible correlation.

Acknowledgements

The authors wish to thank the referee for some valuable suggestions which led some improvement over an earlier version of the paper. Part of the research was done when the first author was visiting the Institute of Statistical Science, Academia Sinica at Taiwan. The first author thanks the Institute for its hospitality.

References

Aitken, A.C., Gonin, H.T., 1935. Fourfold sampling with and without replacement. Proc. Roy. Soc. Edinburgh 55, 114–125.

Consul, P.C., 1974. A simple urn-model dependent on pre-determined strategy. Sankhya Ser. B 36, 391–399.

Crowder, M., Sweeting, T., 1989. Bayesian inference for a bivariate binomial distribution. Biometrika 76, 599–603.

Fisher, R.A., 1950. Statistical Methods for Research Workers, 11th Edition. Hafner Publishing Co., New York.

Hamdan, M.A., 1972. Canonical expansion of the bivariate binomial distribution with unequal marginal indices. Internat. Statist. Rev. 40, 277–288.

Hamdan, M.A., 1975. A note on the trinomial distribution. Internat. Statist. Rev. 43, 219–220.

Hamdan, M.A., Jensen, D.R., 1976. A bivariate binomial distribution and some applications. Austral. J. Statist. 18, 163–169.

Hamdan, M.A., Martinson, E.O., 1971. Maximum likelihood estimation in the bivariate binomial (0,1) distribution: application to 2×2 tables. Austral. J. Statist. 13, 154–158.

Hamdan, M.A., Nasro, M.O., 1986. Maximum likelihood estimation of the parameters of the bivariate binomial distribution. Comm. Statist. A—Theory Methods 15, 747–754.

Hamdan, M.A., Tsokos, C.P., 1971. A model for physical and biological problems: the bivariate-compounded Poisson distribution. Internat. Statist. Rev. 39, 59–63.

Johnson, N.L., Kotz, S., Balakrishnan, N., 1997. Discrete Multivariate Distributions. Wiley, New York.

Kocherlakota, S., 1989. A note on the bivariate binomial distribution. Statist. Probab. Lett. 8, 21–24.

Kocherlakota, S., Kocherlakota, K., 1992. Bivariate Discrete Distributions. Marcel Dekker, New York.

Le, C.T., 1984. A symmetric bivariate binomial distribution and its application to the analysis of clustered samples in medical research. Biometrical J. 26, 289–294.

Ling, K.D., Tai, T.H., 1990. On bivariate binomial distributions of order k. Soochow J. Math. 16, 211–220.

Loukas, S., Kemp, C.D., 1986. The computer generation of bivariate binomial and negative binomial random variables. Comm. Statist. B—Simulation Comput. 15, 15–25.

Mardia, K.V., 1970. Families of Bivariate Distributions. Griffin, London.

Marshall, A.W., Olkin, I., 1985. A family of bivariate distributions generated by the bivariate Bernoulli distribution. J. Amer. Statist. Assoc. 80, 332–338.

Mishra, A., 1996. A generalised bivariate binomial distribution applicable in four-fold sampling. Comm. Statist. A—Theory Methods 25, 1943–1956.

Oluyede, B.O., 1994. A family of bivariate binomial distributions generated by extreme Bernoulli distributions. Comm. Statist. A—Theory Methods 23, 1531–1547.

Ong, S.H., 1992. The computer generation of bivariate binomial variables with given marginals and correlation. Comm. Statist. B—Simulation Comput. 21, 285–299.

Papageorgiou, H., David, K.M., 1994. On countable mixtures of bivariate binomial distributions. Biometrical J. 36, 581–601.

Patil, G.P., Joshi, S.W., 1968. A Dictionary and Bibliography of Discrete Distribution. Oliver and Boyd, Edinburgh.

Pearson, K., 1900. Mathematical contribution to the theory of evolution VII. On the correlation of characters not quantitatively measurable. Philos. Trans. A195, 1–47.

Polson, N., Wasserman, L., 1990. Prior distributions for the bivariate binomial. Biometrika 77, 901–904.

Szego, G., 1959. Orthogonal Polynomials. Colloquium Publication No. 23, American Mathematical Society, Providence, RI.

Takeuchi, K., Takemura, A., 1987. On sum of 0–1 random variables (multivariate). Ann. Inst. Statist. Math. 39, 307–324.