Ann. Inst. Statist. Math.
39 (1987), Part A, 307-324

ON SUM OF 0-1 RANDOM VARIABLES
II. MULTIVARIATE CASE

KEI TAKEUCHI AND AKIMICHI TAKEMURA

(Received Nov. 28, 1985)

Summary

Distribution of sum of vectors of 0-1 random variables is discussed generalizing the univariate results obtained in our previous article Takeuchi and Takemura (1987, *Ann. Inst. Statist. Math.*, **39**, 85-102). As in our previous article no assumption is made on the independence of the 0-1 random variables.

1. Introduction

In our previous article (Takeuchi and Takemura [4]) we discussed sum of 0-1 random variables. Our discussion was restricted to the univariate case. Here we extend the results to the multivariate case. We consider sum of vectors of 0-1 random variables, which are indicator functions of several categories. First we consider the case, where the categories are not mutually exclusive (Section 2). Then we consider the multinomial case, where the categories are mutually exclusive (Section 3). Although the first case can be reduced to the second case by considering all the intersections of categories as separate categories we can obtain more meaningful results by preserving the relations among the categories. The results are especially simple when the categories are close to being independent.

In Sections 2 and 3 our development is carried out for the bivariate case ($k=2$) to avoid excessively cumbersome notation. Most of the results can be written down for more than 2 dimensions in an obvious way. However there are a few results in Section 2 which are particularly simple in the bivariate case. These results will be mentioned explicitly.

As in the univariate case the main point of this article is that we do not assume any condition on independence among 0-1 random variables.

Key words and phrases: Binomial distribution, central binomial moments, finite exchangeability, orthogonal polynomials.

Our main tool is again the central binomial moment appropriately generalized to the multivariate case. It will be seen that our previous development of the univariate case can be extended to the multivariate case with only slight modifications. Therefore we will concentrate on differences between the univariate and the multivariate cases and give a rather brief discussion when the developments are entirely parallel.

In the multivariate case we need more elaborate notations. For the most part we will use notational conventions which closely correspond to those of the univariate case. This makes generalizations of the univariate results transparent.

2. Sum of vectors of 0–1 random variables

In this section we consider a generalization of the results in Takeuchi and Takemura [4] to multivariate case. For notational simplicity we restrict our development to the bivariate case. Most of the results can be stated for more than 2 dimensions with obvious modifications. However there are a few results which are particularly simple for the bivariate case and do not generalize to more than 2 dimensions immediately. See Lemma 2.2 and Remark 2.2 below.

2.1 *Notations and definitions*

Let U_i, V_i, $i=1,\cdots,n$, be 0–1 random variables and let $\boldsymbol{X}_i=(U_i, V_i)'$, $i=1,\cdots,n$. Let $\boldsymbol{S}=(S_1, S_2)'=\sum_{i=1}^{n}\boldsymbol{X}_i$ be the sum. No assumption is made on the dependence among U_i's and V_i's. Let

$$p_{i_1,\dots,i_k;j_1,\dots,j_l}=\Pr(U_{i_1}=1,\cdots,U_{i_k}=1,\ V_{j_1}=1,\cdots,V_{j_l}=1) \tag{2.1}$$

and let

$$p_n(k,l)=\frac{1}{\binom{n}{k}\binom{n}{l}}\sum_{i_1<\cdots<i_k}\sum_{j_1<\cdots<j_l}p_{i_1,\dots,i_k;j_1,\dots,j_l}\,, \tag{2.2}$$

for $k\geqq 1$ or $l\geqq 1$, and $p_n(0,0)=1$. This corresponds to $p_n(k)$ (formula 2.2 of Takeuchi and Takemura [4]) of the univariate case. For any nonnegative integer k let $x^{(k)}=x(x-1)\cdots(x-k+1)$. The mixed (k,l) factorial moment of $\boldsymbol{S}$ is denoted as $\mu_{(k,l)}=\mathrm{E}(S_1^{(k)}S_2^{(l)})$. Then as in Lemma 2.1 of Takeuchi and Takemura [4] we have

$$\mu_{(k,l)}=n^{(k)}n^{(l)}p_n(k,l)\ . \tag{2.3}$$

This can be easily proved by considering the expected value of $(\sum U_{i_1}\cdots U_{i_k})(\sum V_{j_1}\cdots V_{j_l})$. See the proof of Lemma 1.4.1 of Galambos [1] for detail.

Remark 2.1. In the univariate case $p_n(k)$ was defined to be invariant with respect to permutations among X's. This corresponds to the finite exchangeability. In the present case the invariance in the definition of $p_n(k, l)$ is with respect to the product group of permutations among U's and permutations among V's. This contrasts with the usual exchangeability where the invariance is with respect to permutations among X's. However in the next section, where we discuss a generalization to the multinomial case, the finite exchangeability is again appropriate.

Now we define *bivariate central binomial moment* as follows. Let $p_{1\cdot}=p_n(1, 0)$, $p_{\cdot 1}=p_n(0, 1)$. Then

$$(2.4)\qquad \begin{aligned} &q_n(0, 0)=1\,, \\ &q_n(k, l)=\sum_{j=0}^{k}\sum_{h=0}^{l}(-1)^{j+h}\binom{k}{j}\binom{l}{h}p_{1\cdot}^{j}p_{\cdot 1}^{h}p_n(k-j, l-h)\,, \\ &\qquad\qquad\qquad\qquad\qquad k\geqq 1 \text{ or } l\geqq 1\,. \end{aligned}$$

Formally $q_n(k, l)$ is the mixed (k, l) moment about the mean when $p_n(k, l)$ is regarded as (k, l) moment about the origin. Lemma 2.4 below shows that in certain cases this interpretation is legitimate. In any case (2.4) can be easily inverted to yield

$$(2.5)\qquad p_n(k, l)=\sum_{j=0}^{k}\sum_{h=0}^{l}\binom{k}{j}\binom{l}{h}p_{1\cdot}^{j}p_{\cdot 1}^{h}q_n(k-j, l-h)\,.$$

Now let the probability generating function of S be

$$(2.6)\qquad G_n(\theta_1, \theta_2)=\sum_{k=0}^{n}\sum_{l=0}^{n}\theta_1^k\theta_2^l \Pr(S_1=k, S_2=l)\,.$$

Then $M_n(\theta_1, \theta_2)=G_n(1+\theta_1, 1+\theta_2)$ is the factorial moment generating function:

$$(2.7)\qquad M_n(\theta_1, \theta_2)=\sum_{k=0}^{n}\sum_{l=0}^{n}\frac{\theta_1^k\theta_2^l}{k!\,l!}\mu_{(k,l)}=\sum_{k=0}^{n}\sum_{l=0}^{n}\binom{n}{k}\binom{n}{l}\theta_1^k\theta_2^l p_n(k, l)\,.$$

The bivariate central binomial moment generating function $Q_n(\theta_1, \theta_2)$ is defined as

$$(2.8)\qquad Q_n(\theta_1, \theta_2)=\sum_{k=0}^{n}\sum_{l=0}^{n}\binom{n}{k}\binom{n}{l}\theta_1^k\theta_2^l q_n(k, l)\,.$$

Then using (2.5) and (2.7) we obtain after rearrangements

$$(2.9)\qquad M_n(\theta_1, \theta_2)=(1+p_{1\cdot}\theta_1)^n(1+p_{\cdot 1}\theta_2)^n Q_n\left(\frac{\theta_1}{1+p_{1\cdot}\theta_1}, \frac{\theta_2}{1+p_{\cdot 1}\theta_2}\right)\,.$$

See Lemma 2.3 of Takeuchi and Takemura [4] for detail. Letting $\tau_1=$

$\theta_1/(1+p_{1.}\theta_1)$, $\tau_2=\theta_2/(1+p_{.1}\theta_2)$ and solving (2.9) for Q_n we obtain

$$(2.10)\qquad Q_n(\tau_1, \tau_2)=(1-p_{1.}\tau_1)^n(1-p_{.1}\tau_2)^n M_n\left(\frac{\tau_1}{1-p_{1.}\tau_1}, \frac{\tau_2}{1-p_{.1}\tau_2}\right).$$

As in the univariate case $q_n(k, l)$ can be interpreted as representing deviation from independence. This point is illustrated by the followni g four lemmas.

LEMMA 2.1. *S_1 and S_2 are independent if and only if*

$$(2.11)\qquad q_n(k, l)=q_n(k, 0)q_n(0, l).$$

PROOF. If S_1 and S_2 are independent $\mu_{(k,l)}=\mathrm{E}(S_1^{(k)}S_2^{(l)})=\mu_{(k,0)}\mu_{(0,l)}$. From (2.3) we obtain $p_n(k, l)=p_n(k, 0)p_n(0, l)$. Substituting this into (2.4) we obtain (2.11). Conversely assume (2.11). Then $Q_n(\theta_1, \theta_2)=Q_n(\theta_1, 0)Q_n(0, \theta_2)$. Hence

$$M_n(\theta_1, \theta_2)=(1+p_{1.}\theta_1)^n Q_n(\theta_1/(1+p_{1.}\theta_1), 0)(1+p_{.1}\theta_2)^n Q_n(0, \theta_2/(1+p_{.1}\theta_2)).$$

This proves that S_1 and S_2 are independent.

Let $\tilde{X}_i=(\tilde{U}_i, \tilde{V}_i)$ be i.i.d. random vectors with $\Pr(\tilde{U}_i=1)=p_{1.}$, $\Pr(\tilde{V}_i=1)=p_{.1}$, and $\Pr(\tilde{U}_i=1, \tilde{V}_i=1)=p_{11}$, $i=1,\cdots,n$. Let $\tilde{S}=\sum \tilde{X}_i$. We call the distribution of $\tilde{S}$ bivariate binomial distribution with parameters n, $p_{1.}$, $p_{.1}$, p_{11}. Denote $p_{10}=p_{1.}-p_{11}$, $p_{01}=p_{.1}-p_{11}$, $p_{11}=1-p_{10}-p_{01}-p_{00}$. The following lemma characterizes this distribution.

LEMMA 2.2. *S has the bivariate binomial distribution if and only if*

$$(2.12)\qquad \begin{aligned} &q_n(k, l)=0, \qquad \textit{for} \quad k\neq l, \\ &q_n(k, k)=\Delta^k\Big/\binom{n}{k}, \end{aligned}$$

where $\Delta=p_{11}p_{00}-p_{10}p_{01}$.

PROOF. The factorial moment generating function of $\tilde{S}$ is given as

$$\tilde{M}_n(\theta_1, \theta_2)=(1+\theta_1 p_{1.}+\theta_2 p_{.1}+\theta_1\theta_2 p_{11})^n.$$

The lemma holds if and only if $M_n(\theta_1, \theta_2)=\tilde{M}_n(\theta_1, \theta_2)$. Using (2.10) this is equivalent to

$$\begin{aligned} Q_n(\tau_1, \tau_2)&=(1-p_{1.}\tau_1)^n(1-p_{.1}\tau_2)^n \\ &\quad\times\left(1+\frac{\tau_1 p_{1.}}{1-p_{1.}\tau_1}+\frac{\tau_2 p_{.1}}{1-p_{.1}\tau_2}+\frac{\tau_1\tau_2 p_{11}}{(1-p_{1.}\tau_1)(1-p_{.1}\tau_2)}\right)^n \\ &=(1+(p_{11}-p_{1.}p_{.1})\tau_1\tau_2)^n \\ &=(1+\Delta\tau_1\tau_2)^n. \end{aligned}$$

Expanding the right hand side and equating the coefficient of $\tau_1^k\tau_2^l$ we obtain the lemma.

Note that in the bivariate case only the diagonal terms remain in Q_n. In more than 2 dimensions we obtain various cross terms and the result is not as simple as in the two dimensional case. Combining the above two lemmas we obtain

LEMMA 2.3. *S_1 and S_2 are independent binomial random variables if and only if $q_n(k, l)=0$ for $k>0$ or $l>0$.*

Finally the following lemma gives a justification for calling q_n as a moment.

LEMMA 2.4. *Suppose that $S=(S_1, S_2)'$ is a mixture of independent binomial random variables with random success probabilities P_1 and P_2. Then*

$$p_n(k, l)=\mathrm{E}(P_1^k P_2^l)\ ,$$

$$q_n(k, l)=\mathrm{E}\{(P_1-\mathrm{E}(P_1))^k(P_2-\mathrm{E}(P_2))^l\}\ .$$

PROOF.

$$p_{i_1,\cdots,i_k;j_1,\cdots,j_l}=\mathrm{E}(\Pr(U_{i_1}=\cdots=V_{j_l}=1\,|\,P_1, P_2))=\mathrm{E}(P_1^k P_2^l)\ .$$

Hence $p_n(k, l)=\mathrm{E}(P_1^k P_2^l)$. The second equality is now obvious.

2.2 *Approximation by binomial distribution*

Lemma 2.3 suggests that if $q_n(k, l)$'s are small, then the distribution of $S=(S_1, S_2)'$ can be approximated by direct product of two binomial distributions. Actually using the orthogonal polynomials with respect to the binomial distribution this approximation can be explicitly written down. Let $p_{BN}(x; n, p)=\binom{n}{x}p^x(1-p)^{n-x}$ be the probability function of binomial distribution and let $L_j^n(x; p)=(d^j/dp^j)p_{BN}(x; n, p)/p_{BN}(x; n, p)$ be the j-th Krawtchouk polynomial. Then we have

THEOREM 2.1.

$$\Pr(S_1=x, S_2=y)=p_{BN}(x; n, p_{1\cdot})p_{BN}(y; n, p_{\cdot 1}) \times\left\{\sum_{k=0}^{n}\sum_{l=0}^{n}\frac{q_n(k, l)}{k!\,l!}L_k^n(x; p_{1\cdot})L_l^n(y; p_{\cdot 1})\right\}\ , \tag{2.13}$$

$$\Pr(S_1\leqq x, S_2\leqq y) \tag{2.14}$$

$$=F_1(x)F_2(y)-nF_1(x)p_{BN}(y; n-1, p_{\cdot 1})\sum_{l=2}^{n}\frac{q_n(0, l)}{l!}L_{l-1}^{n-1}(y; p_{\cdot 1})$$

$$-nF_2(y)p_{BN}(x; n-1, p_{1\cdot})\sum_{k=2}^{n}\frac{q_n(k, 0)}{k!}L_{k-1}^{n-1}(x; p_{1\cdot})$$
$$+n^2 p_{BN}(x; n-1, p_{1\cdot})p_{BN}(y; n-1, p_{\cdot 1})$$
$$\cdot\sum_{k=1}^{n}\sum_{l=1}^{n}\frac{q_n(k, l)}{k!\,l!}L_{k-1}^{n-1}(x; p_{1\cdot})L_{l-1}^{n-1}(y; p_{\cdot 1}) ,$$

where $F_1(x)=\sum_{u=0}^{x} p_{BN}(u; n, p_{1\cdot})$, $F_2(y)=\sum_{v=0}^{y} p_{BN}(v; n, p_{\cdot 1})$.

This theorem can be proved by checking that the right hand side of (2.13) has the same factorial moments as S. See the proof of Theorem 3.1 of Takeuchi and Takemura [4] for detail.

Remark 2.2. When S has the bivariate binomial distribution, then by Lemma 2.2 only diagonal terms remain in (2.13) and (2.14). This is a particular example of canonical representation of bivariate distributions. A general theory of the canonical representation of bivariate distributions has been given by Lancaster [2]. In more than 2 dimensions the result is more complicated.

2.3 *Convergence to Poisson distribution*

Here we consider the situation where $n\to\infty$, $np_n(1, 0)\to\lambda_1$, $np_n(0, 1)\to\lambda_2$ and the distribution of S_1 and S_2 converges to Poisson distribution. The simplest case is that S_1 and S_2 are asymptotically independent.

THEOREM 2.2. *S_1 and S_2 converge in distribution to independent Poisson variables with parameter λ_1 and λ_2, respectively, if $\lim_{n\to\infty} n^{k+l}p_n(k, l)=\lambda_1^k\lambda_2^l$ for each (k, l). Converse of this is true if $n^{k+l}p_n(k, l)$ is bounded in n for each (k, l).*

This follows from the fact that the k-th factorial moment of Poisson distribution with parameter λ is given by λ^k.

COROLLARY 2.1. *S_1 and S_2 converge in distribution to independent Poisson variables with parameters λ_1 and λ_2, respectively, if $\lim_{n\to\infty} np_n(1, 0)=\lambda_1$, $\lim_{n\to\infty} np_n(0, 1)=\lambda_2$, and $\lim_{n\to\infty} n^{k+l}q_n(k, l)=0$ for each (k, l) such that $k+l\geqq 2$. Converse of this is true if $np_n(1, 0)$, $np_n(0, 1)$ are bounded in n and $n^{k+l}q_n(k, l)$ is bounded in n for each (k, l) such that $k+l\geqq 2$.*

This convergence can be studied in more detail using asymptotic Charlier Type B expansion. For simplicity we only consider the term of order n^{-1}. Let $p(x; \lambda)=(\lambda^x/x!)e^{-\lambda}$ be the probability function of the Poisson distribution and let $L_k(x; \lambda)=(d^k/d\lambda^k)p(x; \lambda)/p(x; \lambda)$ be the k-th Charlier polynomial. Let $\lambda_1=np_n(1, 0)$ and $\lambda_2=np_n(0, 1)$. Taking the logarithm of (2.9) we obtain

$$\log M_n(\theta_1, \theta_2)=n\log(1+\lambda_1\theta_1/n)+n\log(1+\lambda_2\theta_2/n) + \log Q_n(\theta_1/(1+\lambda_1\theta_1/n), \theta_2/(1+\lambda_2\theta_2/n)) . \tag{2.15}$$

Now we assume that $\log Q_n(\tau_1, \tau_2)$ can be expanded as

$$\log Q_n(\tau_1, \tau_2)=\frac{1}{n}[b_{20}\tau_1^2+b_{11}\tau_1\tau_2+b_{02}\tau_2^2]+O(n^{-2}) , \tag{2.16}$$

where the remainder term is of order n^{-2} uniformly for $|\tau_1|\leqq 1+\varepsilon$, $|\tau_2|\leqq 1+\varepsilon$. Substituting (2.16) into (2.15) we obtain

$$\log M_n(\theta_1, \theta_2)=\lambda_1\theta_1+\lambda_2\theta_2+\frac{1}{n}[c_{20}\theta_1^2+c_{11}\theta_1\theta_2+c_{02}\theta_2^2]+O(n^{-2}) , \tag{2.17}$$

where $c_{20}=b_{20}-\lambda_1^2/2$, $c_{11}=b_{11}$, $c_{02}=b_{02}-\lambda_2^2/2$. From this we can obtain the following theorem corresponding to Theorem 4.2 of Takeuchi and Takemura [4].

THEOREM 2.3. *Let* $\lambda_1=np_n(1, 0)$ *and* $\lambda_2=np_n(0, 1)$. *Assume that* $\log Q_n$ *can be expanded as in* (2.16). *Then*

$$\Pr(S_1=x, S_2=y)=p(x;\lambda_1)p(y;\lambda_2)\left\{1+\frac{1}{n}[c_{20}L_2(x;\lambda_1) + c_{11}L_1(x;\lambda_1)L_1(y;\lambda_2)+c_{02}L_2(y;\lambda_2)]\right\}+O(n^{-2}) , \tag{2.18}$$

where c_{20}, c_{11}, c_{02} *are as in* (2.17).

Convergence to more general form of bivariate Poisson distribution is covered by the following theorem. It justifies taking a formal limit of (2.13) and (2.14) in view of the fact that as $n\to\infty$ and $np\to\lambda$, $n^{-j}L_j^n(x; n, p)\to L_j(x;\lambda)$ (see Appendix of Takeuchi and Takemura [4] for a proof).

THEOERM 2.4. *Suppose that as* $n\to\infty$, $np_n(1, 0)\to\lambda_1$, $np_n(0, 1)\to\lambda_2$, *and* $n^{k+l}q_n(k, l)\to q^*(k, l)$ *with* $\sum_{k,l} q^*(k, l)^2/(k!\,l!\,\lambda_1^k\lambda_2^l)<\infty$. *Then*

$$\lim_{n\to\infty}\Pr(S_1=x, S_2=y) =p(x;\lambda_1)p(y;\lambda_2)\left\{\sum_{k,l}^{\infty}\frac{q^*(k, l)}{k!\,l!}L_k(x;\lambda_1)L_l(y;\lambda_2)\right\} . \tag{2.19}$$

Remark 2.3. Under the assumption of the theorem, the marginal distributions of S_1 and S_2 approach Poisson if and only if $q^*(k, 0)=0$, $k\geqq 2$, and $q^*(0, l)=0$, $l\geqq 2$.

An essential step in proving Theorem 2.4 is that the distribution defined by the right hand side of (2.19) has a moment generating function and hence is uniquely determined by its moments. For a precise statement see Lemma A1 of Appendix. Rest of the proof is the same as that of Theorem 4.3 of Takeuchi and Takemura [4].

2.4 *Convergence to normal distribution*

Let

$$Z=(Z_1, Z_2)'=[(S_1-np_{1\cdot})/\sqrt{n}, (S_2-np_{\cdot 1})/\sqrt{n}]' . \tag{2.20}$$

We now consider the case where Z converges to a bivariate normal distribution. Let

$$C=\begin{pmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{pmatrix}$$

be a symmetric positive definite matrix and consider a bivariate normal distribution with mean 0 and covariance matrix C. Let

$$\mu_{k,l}(C)=\mu_{k,l}(c_{11}, c_{12}, c_{22}) \tag{2.21}$$

denote its mixed (k, l) moment. More precisely $\mu_{k,l}$ is the coefficient of $t_1^k t_2^l/(k!\,l!)$ in the expansion of $\exp\{(c_{11}t_1^2+2c_{12}t_1t_2+c_{22}t_2^2)/2\}$. Then $\mu_{k,l}$ is a polynomial in c_{11}, c_{12}, c_{22} and therefore it can be defined for all values of c_{11}, c_{12}, c_{22} (not only for positive definite C). Using this notation we can state the following theorem.

THEOREM 2.5. *Let $p_{1\cdot}=p_n(1, 0)$ and $p_{\cdot 1}=p_n(0, 1)$ be fixed and let Z be defined by* (2.20). *Let c_{11}, c_{12} and c_{22} be constants such that*

$$\Sigma=\begin{pmatrix} c_{11}+p_{1\cdot}(1-p_{1\cdot}) & c_{12} \\ c_{12} & c_{22}+p_{\cdot 1}(1-p_{\cdot 1}) \end{pmatrix}$$

is positive semidefinite. If

$$\lim_{n\to\infty} n^{(k+l)/2} q_n(k, l)=\mu_{k,l}(c_{11}, c_{12}, c_{22})$$

then Z converges to $N(0, \Sigma)$ in distribution. Converse of this is true if $n^{(k+l)/2}q_n(k, l)$ is bounded in n for each (k, l).

Proof of this is the same as that of Theorem 5.1 of Takeuchi and Takemura [4].

The following special case is worth mentioning because of its simplicity:

COROLLARY 2.2. *Let $p_{1\cdot}=p_n(1, 0)$, $p_{\cdot 1}=p_n(0, 1)$ be fixed and let Z be defined by* (2.20). *If*

$$\begin{aligned} &\lim_{n\to\infty} n^{(k+l)/2} q_n(k, l)=0 , \qquad \textit{for} \quad k\neq l , \\ &\lim_{n\to\infty} n^k q_n(k, k)=c_{12}^k k! , \end{aligned} \tag{2.22}$$

then $\boldsymbol{Z}$ *approaches* $N(0, \boldsymbol{\Sigma})$ *in distribution where*

$$\boldsymbol{\Sigma}=\begin{pmatrix} p_{1\cdot}(1-p_{1\cdot}) & c_{12} \\ c_{12} & p_{\cdot 1}(1-p_{\cdot 1}) \end{pmatrix}.$$

Note that (2.22) corresponds to (2.12) for finite n.

Concerning the limiting form of (2.14) we have the following theorem.

THEOREM 2.6. *Let* $p_{1\cdot}=p_n(1, 0)$, $p_{\cdot 1}=p_n(0, 1)$ *be fixed. Suppose that* $q^*(k, l)=\lim_{n\to\infty} n^{(k+l)/2}q_n(k, l)/[p_{1\cdot}(1-p_{1\cdot})]^{k/2}[p_{\cdot 1}(1-p_{\cdot 1})]^{l/2}$ *exists for each* (k, l) *with* $\sum q^*(k, l)^2/k!\,l!<\infty$. *Then*

$$\begin{aligned}(2.23)\quad &\lim_{n\to\infty} \Pr(S_1\leqq np_{1\cdot}+x\sqrt{np_{1\cdot}(1-p_{1\cdot})},\ S_2\leqq np_{\cdot 1}+y\sqrt{np_{\cdot 1}(1-p_{\cdot 1})}) \\ &=\Phi(x)\Phi(y)-\Phi(x)\phi(y)\sum_{l=2}^{\infty}\frac{q^*(0, l)}{l!}h_{l-1}(y)-\Phi(y)\phi(x) \\ &\quad\cdot\sum_{k=2}^{\infty}\frac{q^*(k, 0)}{k!}h_{k-1}(x)+\phi(x)\phi(y)\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\frac{q^*(k, l)}{k!\,l!}h_{k-1}(x)h_{l-1}(y),\end{aligned}$$

where $\phi=\Phi'$ *is the standard normal density and* h_k *is the k-th Hermite polynomial.*

As in Theorem 2.4 an essential step of the proof is to show that the right hand side of (2.23) has a moment generating function. For a precise statement see Lemma A2 of Appendix.

Although the other results of Section 5 of Takeuchi and Takemura [4] can be generalized to the bivariate case, we omit them here.

3. Sum of multinomial random vectors

In this section we discuss a generalization of the univariate results to the multinomial case. The difference from Section 2 is that the 0–1 random variables are now mutually exclusive. Again for notational simplicity we discuss the bivariate case. All results of this section can be stated for more than 2 dimensions with obvious modifications.

As in Section 2 we keep our notation consistent with the univariate case. Since the generalizations in Section 2 and in this section are different, this results in the same symbols used differently in Section 2 and in this section.

3.1 *Notations and definitions*

Consider $k+1$ exhaustive and mutually exclusive categories $C_1,\cdots,$ C_k, C_{k+1}. Let $k=2$ for notational simplicity. Let U, V be the indicator functions of C_1 and C_2, respectively. Note that U and V are mutually exclusive, i.e. $\Pr(U=1, V=1)=0$. Let $\boldsymbol{X}_i=(U_i, V_i)'$, $i=1,\cdots,n$, and let $\boldsymbol{S}=(S_1, S_2)'=\sum \boldsymbol{X}_i$ be the sum. We do not assume any condition on dependence among U's and V's. Let $p_{i;j}=p_{i_1,\cdots,i_k;j_1,\cdots,j_l}$ be defined as in (2.1). Note that if there is a common element among i's and j's then $p_{i;j}=0$ by definition. This implies that if $k+l>n$, then $p_{i;j}=0$. Now let

$$p_n(k, l)=\frac{1}{\binom{n}{k,l}} \sum_{i_1<\cdots<i_k} \sum_{j_1<\cdots<j_l} p_{i_1,\cdots,i_k;j_1,\cdots,j_l}, \tag{3.1}$$

where $\binom{n}{k,l}$ is the multinomial coefficient:

$$\binom{n}{k,l}=\frac{n!}{k!\,l!\,(n-k-l)!}\,. \tag{3.2}$$

Note that $p_n(k, l)$ is defined differently from Section 2. In particular if $k+l>n$, then $p_n(k, l)=0$. If $\boldsymbol{X}_i$'s are exchangeable, then simply $p_n(k, l)=\Pr(U_1=\cdots=U_k=V_{k+1}=\cdots=V_{k+l}=1)$. Let $\mu_{(k,l)}$ again denote the (k, l) factorial moment of $\boldsymbol{S}$. Then in the present case we have

$$\mu_{(k,l)}=n^{(k+l)}p_n(k, l)\,. \tag{3.3}$$

Now let $p_1=p_n(1, 0)$ and $p_2=p_n(0, 1)$. Now we define *central multinomial moment* $q_n(k, l)$ for (k, l) with $k+l\leqq n$ by

$$\begin{aligned} &q_n(0, 0)=1\,, \\ &q_n(k, l)=\sum_{j=0}^{k}\sum_{h=0}^{l}(-1)^{j+h}\binom{k}{j}\binom{l}{h}p_1^j p_2^h p_n(k-j, l-h)\,, \\ &\qquad\qquad\qquad\qquad k\geqq 1 \text{ or } l\geqq 1\,. \end{aligned} \tag{3.4}$$

Inverting this relation for $p_n(k, l)$ with $k+l\leqq n$ we have

$$p_n(k, l)=\sum_{j=0}^{k}\sum_{h=0}^{l}\binom{k}{j}\binom{l}{h}p_1^j p_2^h q_n(k-j, l-h)\,. \tag{3.5}$$

The probability generating function of $\boldsymbol{S}$ is defined as (2.6). The factorial moment generating function $M_n(\theta_1, \theta_2)=G_n(1+\theta_1, 1+\theta_2)$ is now given as

$$M_n(\theta_1, \theta_2)=\sum_{k+l\leqq n}\frac{\theta_1^k\theta_2^l}{k!\,l!}\mu_{(k,l)}=\sum_{k+l\leqq n}\binom{n}{k,l}\theta_1^k\theta_2^l p_n(k, l)\,. \tag{3.6}$$

The central multinomial moment generating function is defined as

$$Q_n(\theta_1, \theta_2)=\sum_{k+l\leqq n}\binom{n}{k,l}\theta_1^k\theta_2^l q_n(k, l)\,. \tag{3.7}$$

Substituting (3.5) into (3.6) we easily obtain

$$(3.8)\quad M_n(\theta_1, \theta_2)=(1+p_1\theta_1+p_2\theta_2)^n Q_n\left[\frac{\theta_1}{1+p_1\theta_1+p_2\theta_2}, \frac{\theta_2}{1+p_1\theta_1+p_2\theta_2}\right].$$

Note the difference of this expression compared to (2.9).

As in Lemma 2.4, $p_n(k, l)$ and $q_n(k, l)$ can be interpreted as moments when S has a certain mixture distribution. The following lemma seems to be a more natural generalization of Lemma 2.2 of Takeuchi and Takemura [4] than Lemma 2.4.

LEMMA 3.1. *Suppose that S is a mixture of multinomial distributions with random success probabilities P_1, P_2 where $(P_1, P_2, P_3)=(P_1, P_2, 1-P_1-P_2)$ is distributed over the simplex: $P_1+P_2+P_3=1$. Then for (k, l) such that $k+l\leqq n$*

$$(3.9)\quad \begin{aligned} p_n(k, l)&=\mathrm{E}(P_1^k P_2^l)\ ,\\ q_n(k, l)&=\mathrm{E}\{(P_1-\mathrm{E}(P_1))^k(P_2-\mathrm{E}(P_2))^l\}\ . \end{aligned}$$

Proof of this is similar to that of Lemma 2.4 and omitted. We also note

LEMMA 3.2. *$q_n(k, l)=0$ for $k>0$ or $l>0$ if and only if S has a multinomial distribution.*

PROOF. Clear from (3.8) since for a multinomial distribution the factorial moment generating function is given by $(1+p_1\theta_1+p_2\theta_2)^n$.

3.2 *Generalized Krawtchouk polynomials and approximation by multinomial distribution*

In the univariate case the distribution of S was expanded around binomial distribution using the central binomial moment and the Krawtchouk polynomials. In Section 2 of the present article the expansion was carried out around the product of binomial distributions using products of the Krawtchouk polynomials. Here we obtain analogous results.

In the present case there are some differences. In the univariate case as well as in Section 2, the polynomials employed were orthogonal polynomials. In this section we first define *generalized Krawtchouk polynomials* which are no longer orthogonal polynomials. As a byproduct we also define another system of polynomials, which are dual to the generalized Krawtchouk polynomials.

Let the probability function of the multinomial distribution be denoted as

$$(3.10)\quad p_{MN}(x_1,\cdots,x_k; n, p_1,\cdots,p_k)=\binom{n}{x_1,\cdots,x_k}p_1^{x_1}\cdots p_k^{x_k}p_{k+1}^{x_{k+1}}\ ,$$

where $p_{k+1}=1-p_1-\cdots-p_k$ and $x_{k+1}=n-x_1-\cdots-x_k$. For notational simplicity we again consider the case $k=2$. Now we define *a generalized Krawtchouk polynomial* by

$$(3.11)\quad L^n_{k,l}(x, y; p_1, p_2)=\frac{\partial^{k+l}}{\partial p_1^k \partial p_2^l} p_{MN}(x, y; n, p_1, p_2)/p_{MN}(x, y; n, p_1, p_2) .$$

It can be easily seen that $L^n_{k,l}$ is a polynomial in (x, y) of degree $k+l$ and $L^n_{k+l}=0$ if $k+l>n$.

For a multinomial distribution $p_n(k, l)=p_1^k p_2^l$. Hence by (3.3)

$$(3.12)\qquad E(S_1^{(k)}S_2^{(l)})=\mu_{(k,l)}=n^{(k+l)}p_1^k p_2^l .$$

Differentiating this relation j times with respect to p_1 and h times with respect to p_2 we obtain

LEMMA 3.3.

$$(3.13)\qquad \sum_{x+y\leq n} x^{(k)}y^{(l)}L^n_{j,h}(x, y; p_1, p_2)p_{MN}(x, y; n, p_1, p_2) = \begin{cases} n^{(k+l)}k^{(j)}l^{(h)}p_1^{k-j}p_2^{l-h} , & \text{if } j\leq k \text{ and } h\leq l , \\ 0 , & \text{otherwise} . \end{cases}$$

This relation can also be used to show

LEMMA 3.4.

$$(3.14)\qquad \begin{aligned} &L^n_{j,h}(x, y; p_1, p_2)p_{MN}(x, y; n, p_1, p_2) \\ &\quad =(-\Delta_x)^j(-\Delta_y)^h n^{(j+h)}p_{MN}(x, y; n-j-h, p_1, p_2) , \end{aligned}$$

where $\Delta_x f(x, y)=f(x, y)-f(x-1, y)$ *and* $\Delta_y f(x, y)=f(x, y)-f(x, y-1)$.

PROOF. It suffices to show that for all k and l

$$\begin{aligned} &\sum x^{(k)}y^{(l)}(-\Delta_x)^j(-\Delta_y)^h n^{(j+h)}p_{MN}(x, y; n-j-h, p_1, p_2) \\ &\quad =\sum x^{(k)}y^{(l)}L^n_{j,h}(x, y; p_1, p_2)p_{MN}(x, y; n, p_1, p_2) . \end{aligned}$$

Summing by parts, the left hand side is equal to

$$n^{(j+h)}\sum \Delta_x^j(x+j)^{(k)}\Delta_y^h(y+h)^{(l)}p_{MN}(x, y; n-j-h, p_1, p_2) .$$

But $\Delta_x^j(x+j)^{(k)}=k^{(j)}x^{(k-j)}$, $\Delta_y^h(y+h)^{(l)}=l^{(h)}y^{(l-h)}$. Therefore this is equal to

$$\begin{aligned} &n^{(j+h)}k^{(j)}l^{(h)}\sum x^{(k-j)}y^{(l-h)}p_{MN}(x, y; n-j-h, p_1, p_2) \\ &\quad =n^{(j+h)}k^{(j)}l^{(h)}(n-j-h)^{(k-j+l-h)}p_1^{k-j}p_2^{l-h} \qquad \text{(by 3.12)} \\ &\quad =n^{(k+l)}k^{(j)}l^{(h)}p_1^{k-j}p_2^{l-h} . \end{aligned}$$

This proves the lemma.

Now we are ready to prove the following theorem.

THEOREM 3.1.

$$\Pr(S_1=x,\ S_2=y)=p_{MN}(x,\ y\ ;\ n,\ p_1,\ p_2)\times\left\{1+\sum_{2\leq k+l\leq n}\frac{q_n(k,\ l)}{k!\,l!}L_{k,l}^n(x,\ y\ ;\ p_1,\ p_2)\right\}. \tag{3.15}$$

$$\begin{aligned}\Pr(S_1\leq x,\ S_2\leq y)&=\sum_{u=2}^{x}\sum_{v=0}^{y}p_{MN}(u,\ v\ ;\ n,\ p_1,\ p_2)\\&-n\sum_{l=2}^{n}\sum_{u=0}^{x}\frac{q_n(0,\ l)}{l!}L_{0,l-1}^{n-1}(u,\ y\ ;\ p_1,\ p_2)p_{MN}(u,\ y\ ;\ n-1,\ p_1,\ p_2)\\&-n\sum_{k=2}^{n}\sum_{v=0}^{y}\frac{q_n(k,\ 0)}{k!}L_{k-1,0}^{n-1}(x,\ v\ ;\ p_1,\ p_2)p_{MN}(x,\ v\ ;\ n-1,\ p_1,\ p_2)\\&+n(n-1)\sum_{\substack{k>0,\ l>0\\k+l\leq n}}\frac{q_n(k,\ l)}{k!\,l!}L_{k-1,l-1}^{n-2}(x,\ y\ ;\ p_1,\ p_2)\\&\cdot p_{MN}(x,\ y\ ;\ n-2,\ p_1,\ p_2)\,.\end{aligned} \tag{3.16}$$

PROOF. (3.16) is a straightforward consequence of (3.15) in view of (3.14). To prove (3.15) it suffices to show that the right hand side of (3.15) has the same factorial moments as S since the factorial moments uniquely determine the distribution. Now by Lemma 3.3

$$\begin{aligned}&\sum_{x,y}\sum_{j+h\leq n}x^{(k)}y^{(l)}p_{MN}(x,\ y\ ;\ n,\ p_1,\ p_2)\frac{q_n(j,\ h)}{j!\,h!}L_{j,h}^n(x,\ y\ ;\ p_1,\ p_2)\\&=\sum_{j=0}^{k}\sum_{h=0}^{l}n^{(k+l)}\frac{k^{(j)}}{j!}\frac{l^{(h)}}{h!}p_1^{k-j}p_2^{l-h}q_n(j,\ h)\\&=n^{(k+l)}p_n(k,\ l)=\mu_{(k,l)}\,.\end{aligned}$$

This proves the theorem.

In the univariate case the Krawtchouk polynomials form a system of orthogonal polynomials. Hence by Perseval's identity we had

$$q_n(k)=\frac{p^k(1-p)^k}{k!\,n^{(k)}}\mathrm{E}(L_k^n)\,.$$

See formula (3.8) of Takeuchi and Takemura [4]. In the present case $\{L_{k,l}^n\}$ do not form a system of orthogonal polynomials. However from Theorem 3.1 we can obtain the following analogous result. Define

$$\tilde{L}_{k,l}^n(x,\ y\ ;\ p_1,\ p_2)=\sum_{j=0}^{k}\sum_{h=0}^{l}\binom{k}{j}\binom{l}{h}\frac{1}{n^{(k-j+l-h)}}(-1)^{j+h}p_1^jp_2^hx^{(k-j)}y^{(l-h)}\,, \tag{3.17}$$

then from (3.3) we easily obtain

$$\mathrm{E}(\tilde{L}^n_{k,l}(S_1, S_2; p_1, p_2)) = q_n(k, l) \,. \tag{3.18}$$

Now write out the left hand side expectation term by term using Theorem 3.1 and equate each term. Then we have

PROPOSITION 3.1.

$$\sum_{x,y} \tilde{L}^n_{k,l}(x, y; p_1, p_2) L^n_{j,h}(x, y; p_1, p_2) p_{MN}(x, y; n, p_1, p_2) = \begin{cases} k!\,l!\,, & \textit{if } k=j \textit{ and } l=h\,, \\ 0\,, & \textit{otherwise}\,. \end{cases} \tag{3.19}$$

We see that $\{L^n_{k,l}\}$ and $\{\tilde{L}^n_{k,l}\}$ form dual systems of polynomials with respect to the multinomial distribution. As seen above, these polynomials are almost as convenient as orthogonal polynomials.

Finally we derive generating function $G_n(t_1, t_2)$ of generalized Krawtchouk polynomials useful for later developments. By definition

$$p_{MN}(x, y; n, p_1+t_1, p_2+t_2) = p_{MN}(x, y; n, p_1, p_2)\left\{1 + \sum_{k+l\le n} \frac{t_1^k t_2^l}{k!\,l!} L^n_{k,l}(x, y; p_1, p_2)\right\} \,.$$

Hence

$$\begin{aligned} G_n(t_1, t_2) &= p_{MN}(x, y; n, p_1+t_1, p_2+t_2)/p_{MN}(x, y; n, p_1, p_2) \\ &= [1+t_1/p_1]^x [1+t_2/p_2]^y [1-(t_1+t_2)/(1-p_1-p_2)]^{n-x-y} \,. \end{aligned} \tag{3.20}$$

3.3 *Convergence to Poisson distribution*

As far as convergence to Poisson distribution is concerned there is not much to discuss here. Actually, Theorem 2.2, Corollary 2.1 and Theorem 2.4 hold word by word in the present setting. Concerning Theorem 2.2 and Corollary 2.1 this is clear because in (2.3) and (3.3) we have $n^{(k)}n^{(l)}/n^{k+l} \to 1$, $n^{(k+l)}/n^{k+l} \to 1$, respectively, as $n \to \infty$. Concerning Theorem 2.4 we only have to check that as $n \to \infty$, $n^{-(k+l)} L^n_{k,l}(x, y; p_1, p_2) \to L_k(x; \lambda_1) L_l(y; \lambda_2)$. We state this as a lemma.

LEMMA 3.5. *As $n \to \infty$, $np_1 \to \lambda_1$, and $np_2 \to \lambda_2$*

$$n^{-(k+l)} L^n_{k,l}(x, y; p_1, p_2) \to L_k(x; \lambda_1) L_l(y; \lambda_2) \,, \tag{3.21}$$

where $L_k(x; \lambda_1)$ and $L_l(y; \lambda_2)$ are the Charlier polynomials.

PROOF. From (3.20) the generating function of $n^{-k-l} L^n_{k,l}$ is given by

$$\tilde{G}_n(t_1, t_2) = [1+t_1/np_1]^x [1+t_2/np_2]^y [1-(t_1+t_2)/n(1-p_1-p_2)]^{n-x-y} \,.$$

As $n \to \infty$, $np_1 \to \lambda_1$, and $np_2 \to \lambda_2$, $\tilde{G}_n(t_1, t_2)$ converges to

$$\tilde{G}(t_1, t_2)=[1+t_1/\lambda_1]^x[1+t_2/\lambda_2]^y \exp(-t_1-t_2) .$$

This is the generating function of $L_k(x; \lambda_1)L_l(y; \lambda_2)$. Since $\tilde{G}_n$ and $\tilde{G}$ are analytic (around the origin), the lemma follows.

Now consider Theorem 2.3. In the present situation a difference appears in the term of order n^{-1}. Let $p_1=\lambda_1/n$, $p_2=\lambda_2/n$, and take the logarithm of (3.8). Then

$$\log M_n(\theta_1, \theta_2)=n \log(1+\lambda_1\theta_1/n+\lambda_2\theta_2/n) + \log Q_n\left[\frac{\theta_1}{1+\lambda_1\theta_1/n+\lambda_2\theta_2/n}, \frac{\theta_1}{1+\lambda_1\theta_1/n+\lambda_2\theta_2/n}\right] . \tag{3.22}$$

Now assume that $\log Q_n(\tau_1, \tau_2)$ can be expanded as in (2.16). Substitute (2.16) into (3.22). Then $\log M_n(\theta_1, \theta_2)$ is expressed as in (2.17) with c_{11} now defined as $c_{11}=b_{11}-\lambda_1\lambda_2$. Therefore

THEOREM 3.2. *Let $\lambda_1=np_1$ and $\lambda_2=np_2$. Assume that $\log Q_n$ can be expanded as in* (2.16). *Then*

$$\Pr(S_1=x, S_2=y)=p(x; \lambda_1)p(y; \lambda_2)\left\{1+\frac{1}{n}[c_{20}L_2(x; \lambda_1) + c_{11}L_1(x; \lambda_1)L_1(y; \lambda_2)+c_{02}L_2(y; \lambda_2)]\right\}+O(n^{-2}) , \tag{3.23}$$

where $c_{20}=b_{20}-\lambda_1^2/2$, $c_{11}=b_{11}-\lambda_1\lambda_2$, $c_{02}=b_{02}-\lambda_2^2/2$.

3.4 *Convergence to normal distribution*

Convergence to normal distribution can be stated in a similar way as Theorem 2.5. The difference in the present case is that the asymptotic covariance matrix is based on the multinomial distribution.

THEOREM 3.3. *Let $p_1=p_n(1, 0)$ and $p_2=p_n(0, 1)$ be fixed and let $\boldsymbol{Z}$ be defined by* (2.20). *Let c_{11}, c_{12} and c_{22} be constants such that*

$$\boldsymbol{\Sigma}=\begin{pmatrix} c_{11}+p_1(1-p_1) & c_{12}-p_1p_2 \\ c_{12}-p_1p_2 & c_{22}+p_1(1-p_1) \end{pmatrix}$$

is positive semidefinite. If

$$\lim_{n\to\infty} n^{(k+l)/2}q_n(k, l)=\mu_{k,l}(c_{11}, c_{12}, c_{22}) ,$$

then $\boldsymbol{Z}$ converges to $N(0, \boldsymbol{\Sigma})$ in distribution. Converse of this is true if $n^{(k+l)/2}q_n(k, l)$ is bounded in n for each (k, l).

A substantial modification is needed to adapt Theorem 2.6 to the present case. First we need bivariate Hermite polynomials. Let $\phi(x, y; \rho)$ denote the density function of standard bivariate normal distribution with correlation coefficient ρ. Define Hermite polynomial of degree (k, l) by

$$h_{k,l}(x, y; \rho)=\frac{\partial^{k+l}}{\partial x^k \partial y^l}\phi(x, y; \rho)/\phi(x, y; \rho) . \tag{3.24}$$

Then we have the following lemma.

LEMMA 3.6. *Let* p_1, p_2 *be fixed. As* $n\rightarrow\infty$

$$\begin{aligned}[p_1(1-p_1)]^{k/2}[p_2(1-p_2)]^{l/2}n^{-(l+k)/2}\\ \cdot L^n_{k,l}(np_1+x\sqrt{np_1(1-p_1)}, np_2+y\sqrt{np_2(1-p_2)}; p_1, p_2)\\ \rightarrow h_{k,l}(x, y; \rho) ,\end{aligned} \tag{3.25}$$

where

$$\rho=-\left[\frac{p_1p_2}{(1-p_1)(1-p_2)}\right]^{1/2} . \tag{3.26}$$

Proof of this is given in Appendix. (3.25) allows us to write down the formal limit of (3.16). This results in Bivariate Edgeworth expansion (see Chapter 3 of Mardia [3] for example). However since the bivariate Hermite polynomials are not orthogonal to each other the regularity condition for convergence is more cumbersome. Define

$$k!\,l!\,k'!\,l'!\,a_{k,l;k',l'}=\int h_{k,l}(x, y; \rho)h_{k',l'}(x, y; \rho)\phi(x, y; \rho)dxdy . \tag{3.27}$$

A more concrete expression of $a_{k,l;k',l'}$ is discussed in Appendix. Now we can state the following theorem.

THEOREM 3.4. *Let* $p_1=p_n(1, 0)$, $p_2=p_n(0, 1)$ *be fixed. Suppose that* $q^*(k, l)=\lim_{n\rightarrow\infty} n^{(k+l)/2}q_n(k, l)/[p_1(1-p_1)]^{k/2}[p_2(1-p_2)]^{l/2}$ *exists for each* (k, l) *such that*

$$\sum_{k,l,k',l'} |q^*(k, l)q^*(k', l')a_{k,l;k',l'}|<\infty . \tag{3.28}$$

Then

$$\begin{aligned}&\lim_{n\rightarrow\infty} \Pr(S_1\leqq np_1+x\sqrt{np_1(1-p_1)}, S_2\leqq np_2+y\sqrt{np_2(1-p_2)})\\ &=\Phi(x, y; \rho)-\sum_{l=2}^{\infty}\int_{-\infty}^{x}\frac{q^*(0, l)}{l!}h_{0,l-1}(u, y; \rho)\phi(u, y; \rho)du\\ &\quad-\sum_{k=2}^{\infty}\int_{-\infty}^{y}\frac{q^*(k, 0)}{k!}h_{k-1,0}(x, v; \rho)\phi(x, v; \rho)dv\\ &\quad+\sum_{k>0}\sum_{l>0}\frac{q^*(k, l)}{k!\,l!}h_{k-1,l-1}(x, y; \rho)\phi(x, y; \rho) ,\end{aligned} \tag{3.29}$$

where

$$\rho=-\left[\frac{p_1p_2}{(1-p_1)(1-p_2)}\right]^{1/2}.$$

Note that the regularity condition guarantees that the density of the distribution on the right hand side converges in L^2. Using Lemma A2, Theorem 3.4 can be proved as Theorem 2.4 or Theorem 2.6.

UNIVERSITY OF TOKYO

REFERENCES

[1] Galambos, J. (1978). *The Asymptotic Theory of Extreme Order Statistics*, Wiley, New York.

[2] Lancaster, H. O. (1958). The structure of bivariate distributions, *Ann. Math. Statist.*, **29**, 719-736.

[3] Mardia, K. V. (1970). *Families of Bivariate Distributions*, Griffin, London.

[4] Takeuchi, K. and Takemura, A. (1987). On sum of 0-1 random variables I. Univariate case, *Ann. Inst. Statist. Math.*, **39**, 85-102.

Appendix

The following lemma is needed for proving Theorem 2.4.

LEMMA A1. *Let P^* be a distribution over pairs of nonnegative integers such that* $\sum_{x,y} P^*(x, y)^2/[p(x;\lambda_1)p(x;\lambda_2)]=M<\infty$. *Then P^* has a moment generating function.*

PROOF. By Schwarz, for any $a>0$, $b>0$

$$\begin{aligned}\sum a^x b^y P^*(x, y) &\leqq \left[\sum a^{2x}b^{2y}\lambda_1^x\lambda_2^y/(x!\,y!)\sum P^*(x, y)^2 x!\,y!/(\lambda_1^x\lambda_2^y)\right]^{1/2}\\ &= e^{[(a^2-1)\lambda_1+(b^2-1)\lambda_2]/2}M^{1/2}<\infty\,.\end{aligned}$$

The following lemma is needed for proving Theorems 2.6 and 3.4.

LEMMA A2. *Let $f(x, y)$ be a density function such that* $\int f(x, y)^2/\phi(x, y;\rho)dxdy<\infty$, *where $\phi(x, y;\rho)$ is the density function of the standard bivariate normal distribution with correlation coefficient ρ. Then f has a moment generating function.*

Proof is similar to the proof of Lemma A1 and omitted.

PROOF OF LEMMA 3.6. The generating function of $L_{k,l}^n(x, y; p_1, p_2)$ is given by (3.20). Similarly the generating function of the bivariate Hermite polynomials is given by

(A1)
$$\phi(x+t_1, y+t_2; \rho)/\phi(x, y; \rho)$$
$$=\exp\left[-\frac{t_1^2-2xt_1+t_2^2-2yt_2}{2(1-\rho^2)}+\frac{\rho(t_1t_2-xt_2-yt_1)}{(1-\rho^2)}\right].$$

Now transforming (3.20) to the generating function of $[p_1(1-p_1)]^{k/2}\cdot[p_2(1-p_2)]^{l/2}n^{-(l+k)/2}L_{k,l}^n$ and letting $n\to\infty$ we easily see that the limit coincides with (A1). Therefore by analyticity of generating functions Lemma 3.6 follows.

We finally discuss how (3.27) can be calculated. Let

$$\Sigma^{-1}=(\sigma^{ij})=\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}^{-1}.$$

Then (A1) can be written as

$$g(t_1, t_2)=\exp\left[-\frac{1}{2}(t_1, t_2)\Sigma^{-1}(t_1, t_2)'+(t_1, t_2)\Sigma^{-1}(x, y)'\right].$$

Therefore

$$\sum_{k,l,k',l'} t_1^k t_2^l s_1^{k'} s_2^{l'} a_{k,l;k',l'}=\int g(t_1, t_2)g(s_1, s_2)\phi(x, y; \rho)dxdy$$
$$=\exp[(t_1, t_2)\Sigma^{-1}(s_1, s_2)']$$
$$=\sum_n \frac{1}{n!(1-\rho^2)^n}[t_1s_1+t_2s_2-\rho(t_1s_2+t_2s_1)]^n.$$

From this $a_{k,l;k',l'}$ can be easily obtained. In particular $a_{k,l;k',l'}=0$ if $k+l\neq k'+l'$.