PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 102, Number 3, March 1988

MULTINOMIAL PROBABILITIES, PERMANENTS AND A CONJECTURE OF KARLIN AND RINOTT

R. B. BAPAT

(Communicated by Bert E. Fristedt)

ABSTRACT. The probability density function of a multiparameter multinomial distribution can be expressed in terms of the permanent of a suitable matrix. This fact and certain results on conditionally negative definite matrices are used to prove a conjecture due to Karlin and Rinott.

1. Introduction. If $A = ((a_{ij}))$ is an $n \times n$ matrix, the permanent of A, denoted by $\operatorname{per} A$, is defined as

$$\operatorname{per} A = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i\sigma(i)},$$

where S_n is the group of permutations of $1, 2, \ldots, n$.

The probability density function of a multinomial distribution can be expressed in terms of the permanent of a suitable matrix. To begin with a simple example, suppose X denotes the number of heads resulting from n tosses of a coin, with probability of heads equal to p, $0 \leq p \leq 1$, on a single toss. Then X has the Binomial distribution and its probability density function is given by

$$P(X = x) = \frac{n!}{x!(n-x)!} p^x q^{n-x}, \qquad x = 0, 1, \ldots, n; \tag{1}$$

where $P(X = x)$ denotes the probability that X equals x, $q = 1 - p$, and, as usual, $P(X = x)$ is understood to be zero for all values of x not specified in (1).

The probability in (1) can be expressed in terms of a permanent as follows:

$$P(X = x) = \frac{1}{x!(n-x)!} \operatorname{per} \begin{bmatrix} p & \cdots & p \\ \vdots & & \vdots \\ p & \cdots & p \\ q & \cdots & q \\ \vdots & & \vdots \\ q & \cdots & q \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} p \\ \vdots \\ p \end{matrix}}\right\} x \text{ times,} \\ \left.\vphantom{\begin{matrix} q \\ \vdots \\ q \end{matrix}}\right\} (n-x) \text{ times,} \end{matrix} \qquad x = 0, 1, \ldots, n. \tag{2}$$

Received by the editors July 11, 1986 and, in revised form, December 15, 1986. Presented at the Second Utah State Matrix Theory Conference held in Logan, Utah, January 29–31, 1987, sponsored by the Department of Mathematics, Utah State University.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 15A15, 60E15; Secondary 15A48, 94A17.

Key words and phrases. Permanents, multinomial distribution, conditionally negative definite matrices.

0002-9939/88 \$1.00 + \$.25 per page

The expression in (2) admits generalizations. For example, suppose n different coins are tossed once and let X be the number of heads obtained. If p_i is the probability of heads on a single toss of the ith coin and if $q_i = 1-p_i$, $i = 1, 2, \ldots, n$, then it can be verified that

$$(3)\quad P(X = x) = \frac{1}{x!(n-x)!}\operatorname{per}\begin{bmatrix} p_1 & \cdots & p_n \\ \vdots & & \vdots \\ p_1 & \cdots & p_n \\ q_1 & \cdots & q_n \\ \vdots & & \vdots \\ q_1 & \cdots & q_n \end{bmatrix}\begin{matrix} \left.\vphantom{\begin{matrix} a\\ a\\ a\end{matrix}}\right\} x \text{ times,} \\ \left.\vphantom{\begin{matrix} a\\ a\\ a\end{matrix}}\right\} (n-x) \text{ times,} \end{matrix} \qquad x = 0, 1, \ldots, n.$$

More generally, instead of tossing n coins, it may be an experiment of rolling n dice, differently loaded, and if X_i denotes the number of times i spots are obtained, $i = 1, 2, \ldots, 6$, the density function of $(X_1, \ldots, X_6)$ can be written in terms of a permanent.

To make the concepts precise, consider an experiment which can result in any of r possible outcomes and suppose n trials of the experiment are performed. Let π_{ij} be the probability that the experiment results in the ith outcome at the jth trial, $i = 1, 2, \ldots, r$; $j = 1, 2, \ldots, n$. Let Π denote the $r \times n$ matrix $((\pi_{ij}))$, which, of course, is column stochastic. Let X_i denote the number of times the ith outcome is obtained in the n trials, $i = 1, 2, \ldots, r$, and let $X = (X_1, \ldots, X_r)$. In this setup we will say that X has the multiparameter multinomial distribution with the $r \times n$ parameter matrix Π. Let π^i denote the ith row of Π, $i = 1, 2, \ldots, r$. The density function of X can again be expressed in terms of a permanent as

$$(4)\qquad P(X = t) = \frac{1}{t_1! \cdots t_r!}\operatorname{per}\begin{bmatrix} \pi^1 \\ \vdots \\ \pi^1 \\ \vdots \\ \pi^r \\ \vdots \\ \pi^r \end{bmatrix}\begin{matrix} \left.\vphantom{\begin{matrix} a\\ a\\ a\end{matrix}}\right\} t_1 \text{ times,} \\ \\ \left.\vphantom{\begin{matrix} a\\ a\\ a\end{matrix}}\right\} t_r \text{ times,} \end{matrix} \qquad t \in \mathcal{K}_n,$$

where

$$\mathcal{K}_n = \left\{ k = (k_1, \ldots, k_r)\colon\ k_i \text{ nonnegative integers}, \sum_{i=1}^{r} k_i = n \right\}.$$

It seems that the representation (4) is important in understanding certain properties of the multinomial distribution. The purpose of this paper is to exploit (4) to settle a conjecture due to Karlin and Rinott [**4**].

2. The problem. Before stating the main problem, we need some notation. Let

$$H^r = \left\{ x \in R^r\colon\ \sum_{i=1}^{r} x_i = 0 \right\}.$$

DEFINITION 1. A real, symmetric $r \times r$ matrix A is said to be conditionally negative definite (c.n.d.) if for any $x \in H^r$,

$$\sum_{i=1}^{r} \sum_{j=1}^{r} a_{ij} x_i x_j \leq 0.$$

Now suppose $X = (X_1, \ldots, X_r)$ follows the multiparameter multinomial distribution with the $r \times n$ parameter matrix $\Pi = ((\pi_{ij}))$. We assume throughout that Π is positive and that $n \geq 2$, $r \geq 2$.

Fix $k \in \mathcal{K}_{n-2}$ and let

$$k_{ij} = (k_1, \ldots, k_{i-1},\ k_i + 1,\ k_{i+1}, \ldots, k_{j-1},\ k_j + 1,\ k_{j+1}, \ldots, k_r), \qquad 1 \leq i \neq j \leq r, \tag{5}$$
$$k_{ii} = (k_1, \ldots, k_{i-1},\ k_i + 2,\ k_{i+1}, \ldots, k_r), \qquad 1 \leq i \leq r.$$

Clearly, $k_{ij} \in \mathcal{K}_n$, $i, j = 1, 2, \ldots, r$. Our main result is that $((\log P(X = k_{ij})))$ is c.n.d. This appears as Conjecture 2.1 in Karlin and Rinott [**4**], where it has been confirmed for $r = 2, 3$ and for any n. We refer to [**4**] for a discussion concerning the relevance of the problem to multivariate majorization and inequalities as well as for certain consequences of proving the conjecture. In particular, Conjecture 2.2 of [**4**] is also verified once Conjecture 2.1 is established. It must be remarked that we have used a notation different than that of [**4**] at various places in the paper. Also, we find it more convenient to work with conditionally negative definite matrices rather than conditionally positive definite matrices.

3. Permanents of positive matrices. In this section we give certain results on permanents that will be used. Let A be a real matrix of order $(n-2) \times n$, $n \geq 2$. Let e^j denote the unit row vector $(0, \ldots, 0, 1, 0, \ldots, 0)$, where the 1 occurs at the jth place, $j = 1, 2, \ldots, n$. Define an $n \times n$ matrix $\hat{A}$ as follows. The (i, j)th entry of $\hat{A}$ is the permanent of the augmented matrix

$$\begin{bmatrix} A \\ e^i \\ e^j \end{bmatrix}, \qquad i, j = 1, 2, \ldots, n.$$

Note that the diagonal entries of $\hat{A}$ are all zero and $\hat{A}$ is symmetric. Of course if $n = 2$, then A is absent and

$$\hat{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The following theorem is a deep result in the theory of permanents. It was originally proved by Alexandroff [**1**] in a more general form. It serves as a crucial step in the proof of the well-known van der Waerden conjecture due to Egorychev [**2, 10**] and Falikman [**3**]. For a proof of the result, see Theorem 2.8 of [**10**]. Again, beware of differences in notation.

THEOREM 2. *If A is a positive $(n-2) \times n$ matrix, $n \geq 2$, then the matrix $\hat{A}$ is nonsingular and has exactly one, simple, positive eigenvalue.*

The following result will be deduced from Theorem 2.

THEOREM 3. *Let A be a positive $(n-2)\times n$ matrix, $n \geq 2$, and let $x^1,\dots,x^r$ be row vectors in R^n. Define the $r\times r$ matrix $B = ((b_{ij}))$ as*

$$b_{ij} = \operatorname{per}\begin{bmatrix} A \\ x^i \\ x^j \end{bmatrix}, \qquad i,j = 1,2,\dots,r.$$

Then B has at most one, simple, positive eigenvalue.

PROOF. The permanent admits a Laplace expansion in terms of a set of rows [**7**, p. 16].

Expand

$$\operatorname{per}\begin{bmatrix} A \\ x^i \\ x^j \end{bmatrix}$$

in terms of the last two rows and observe that

$$b_{ij} = \langle x^i\hat{A}, x^j\rangle, \qquad i,j = 1,2,\dots,r. \tag{6}$$

Let X be the $r\times n$ matrix whose ith row is x^i, $i = 1,2,\dots,r$. Then it follows from (6) that $B = X\hat{A}X'$. By Theorem 2, $\hat{A}$ has exactly one positive eigenvalue and hence (essentially) by Sylvester's law, B has at most one positive eigenvalue. This completes the proof.

4. Conditionally negative definite matrices. The definition of a c.n.d. matrix was given in §2. The following result will be useful (see, for example, Parthasarathy and Schmidt [**8**, p. 3]).

LEMMA 4. *A real, symmetric matrix A is c.n.d. if and only if for each $\alpha > 0$, the matrix $((e^{-\alpha a_{ij}}))$ is positive semidefinite.*

Recall that a function F: $(0,\infty) \to R$ is said to be completely monotonic if it is in $C^\infty(0,\infty)$ and $(-1)^k F^{(k)}(x) \geq 0$, $x \in (0,\infty)$, $k = 0,1,2,\dots,$ where, by definition, $F^{(0)} = F$.

The following result appears in a recent paper by Micchelli [**6**]. A proof is included for completeness.

LEMMA 5. *Let A be a c.n.d. $n\times n$ matrix with positive entries and let F: $(0,\infty) \to R$ be completely monotonic. Then the matrix $((F(a_{ij})))$ is positive semidefinite.*

PROOF. By a well-known theorem of Bernstein (see, for example, [**11**, p. 160]), F admits the representation

$$F(t) = \int_0^\infty e^{-t\sigma}\, d\mu(\sigma), \qquad t > 0,$$

where $d\mu(\sigma)$ is a Borel measure on $(0,\infty)$. The result follows by Lemma 4.

Now we have the following.

LEMMA 6. *Let A be a symmetric, positive $r\times r$ matrix with exactly one, simple, positive eigenvalue. Then, for any $x \in H^r$,*

$$\prod_{i,j=1}^{r} a_{ij}^{x_i x_j} \leq 1.$$

PROOF. It is known that if A is a symmetric, positive $r \times r$ matrix, then there exists a positive vector z such that the matrix $B = ((b_{ij})) = ((a_{ij}z_iz_j))$ is doubly stochastic (see, for example, [**5**, **9**]). Clearly, if A satisfies the hypothesis of the lemma, then B also has only one positive eigenvalue by Sylvester's law, and, furthermore,

$$\prod_{i,j=1}^{r} b_{ij}^{x_ix_j} = \prod_{i,j=1}^{r} (a_{ij}z_iz_j)^{x_ix_j} = \prod_{i,j=1}^{r} a_{ij}^{x_ix_j} \quad \text{for any } x \in H^r.$$

Thus, we may assume, without loss of generality, that A is doubly stochastic, so that the vector $(1, \ldots, 1)$ is an eigenvector of A corresponding to the eigenvalue 1. Since A has only one positive eigenvalue, any vector $x \in H^r$ lies in the span of eigenvectors of A corresponding to only nonpositive eigenvalues and, hence,

$$\sum_{i=1}^{r}\sum_{j=1}^{r} a_{ij}x_ix_j \leq 0.$$

Thus, A is c.n.d. The function $F(t) = t^{-\alpha}$, $t > 0$, is completely monotonic for any $\alpha > 0$ and by Lemma 5, $((a_{ij}^{-\alpha}))$ is positive semidefinite.

Hence, for any $\alpha > 0$, $((e^{-\alpha \log a_{ij}}))$ is positive semidefinite and by Lemma 4, $((\log a_{ij}))$ is c.n.d. This completes the proof.

5. The main result. We are now in a position to prove the following statement, conjectured by Karlin and Rinott [**4**].

THEOREM 7. *Let $X = (X_1, \ldots, X_r)$ have the multiparameter multinomial distribution with the $r \times n$ parameter matrix Π. Let $k \in \mathcal{K}_{n-2}$, let k_{ij} be defined as in* (5), *and let $m_{ij} = P(X = k_{ij})$, $i, j = 1, 2, \ldots, r$. Then the matrix $((\log m_{ij}))$ is c.n.d.*

PROOF. We have to prove that for any $x \in H^r$,

$$\prod_{i,j=1}^{r} m_{ij}^{x_ix_j} \leq 1. \tag{7}$$

Let A be the $(n-2) \times n$ matrix formed by taking k_i copies of π^i, the ith row of Π, $i = 1, 2, \ldots, n$. Define the $r \times r$ matrix $B = ((b_{ij}))$ as

$$b_{ij} = \operatorname{per}\begin{bmatrix} A \\ \pi^i \\ \pi^j \end{bmatrix}, \qquad i, j = 1, 2, \ldots, r.$$

By Theorem 3, B has at most one positive eigenvalue, and since B is a positive matrix, it must have exactly one positive eigenvalue. So by Lemma 6,

$$\prod_{i,j=1}^{r} b_{ij}^{x_ix_j} \leq 1 \quad \text{for any } x \in H^r. \tag{8}$$

By the relationship (4),

$$(9)\quad m_{ij} = P(X = k_{ij}) = \begin{cases} \dfrac{b_{ij}}{(k_1!, \ldots, k_r!)(k_i+1)(k_j+1)}, & 1 \le i \ne j \le r, \\ \dfrac{b_{ij}}{(k_1! \cdots k_r!)(k_i+1)(k_i+2)}, & 1 \le i \le r. \end{cases}$$

Let D be the $r \times r$ diagonal matrix with its ith diagonal entry equal to $(k_i+1)^{-1}$, $i = 1, 2, \ldots, r$, and let

$$(10)\qquad C = \frac{1}{k_1! \cdots k_r!} DBD.$$

It follows from (8) and (10) that for any $x \in H^r$,

$$(11)\qquad \prod_{i,j=1}^{r} c_{ij}^{x_i x_j} \le 1.$$

From (9) and (10),

$$(12)\qquad m_{ij} = \begin{cases} c_{ij}, & 1 \le i \ne j \le r, \\ ((k_i+1)/(k_i+2))c_{ii}, & 1 \le i \le r, \end{cases}$$

Since $(k_i+1)/(k_i+2) < 1$, $i = 1, 2, \ldots, r$, it follows from (12) and (11) that

$$\prod_{i,j=1}^{r} m_{ij}^{x_i x_j} \le \prod_{i,j=1}^{r} c_{ij}^{x_i x_j} \le 1 \quad \text{for any } x \in H^r,$$

and the proof is complete.

References

1. A. D. Alexandroff, *Zur theorie der gemischten volumina von konvexen körpern.* IV, Mat. Sb. **45** (1938), no. 3, 227–251. (Russian; German summary).
2. G. P. Egorychev, *The solution of van der Waerden's problem for permanents*, Adv. in Math. **42** (1981), 299–305.
3. D. I. Falikman, *A proof of van der Waerden's conjecture on the permanent of a doubly stochastic matrix*, Mat. Zametki **29** (1981), 931–938. (Russian)
4. S. Karlin and Y. Rinott, *Entropy inequalities for classes of probability distributions.* II, *The multivariate case*, Adv. in Appl. Probab. **13** (1981), 325–351.
5. M. Marcus and M. Newman, *The permanent of a symmetric matrix*, Notices Amer. Math. Soc. **8** (1981), 595.
6. C. A. Micchelli, *Interpolation of scattered data: Distance matrices and conditionally positive definite matrices*, Constr. Approx. **2** (1986), 11–22.
7. H. Minc, *Permanents*, Encyclopedia of Mathematics and its Applications, Vol. 6, Addison-Wesley, Reading, Mass., 1978.
8. K. R. Parthasarathy and K. Schmidt, *Positive definite kernels, continuous tensor products and central limit theorems of probability theory*, Springer-Verlag, Berlin, 1972.
9. R. Sinkhorn, *A relationship between arbitrary positive matrices and doubly stochastic matrices*, Ann. Math. Statist. **35** (1964), 876–879.
10. J. H. van Lint, *Notes on Egoritsjev's proof of the van der Waerden conjecture*, Linear Algebra Appl. **39** (1981), 1–8.
11. D. V. Widder, *The Laplace transform*, Princeton Univ. Press, Princeton, N.J., 1946.

INDIAN STATISTICAL INSTITUTE, DELHI CENTRE 7, S. J. S. SANSANWAL MARG, NEW DELHI-110016, INDIA