Permanents in Probability and Statistics

R. B. Bapat
Indian Statistical Institute
New Delhi, 110016, India

Submitted by M. D. Perlman

ABSTRACT

This is a survey of several topics in probability and statistics in which permanents seem to have a role. The topics covered include multiparameter versions of the multinomial and the negative multinomial distributions, arrangement-decreasing functions, order statistics for nonidentically distributed random variables, sequential experiments with feedback, and sampling. It is shown by means of examples how Alexandroff's inequality can be applied to demonstrate log-concavity of certain sequences in some of these areas.

1. INTRODUCTION

If A is an $n \times n$ matrix, then 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 set of permutations of $1, 2, \ldots, n$. Thus the definition of the permanent is similar to that of the determinant except that all terms in the expansion get a positive sign.

The permanent function was introduced by Binet and independently by Cauchy as early as in 1812, more or less simultaneously with the determinant. However, the major developments in the theory of permanents have taken place only in the last twenty years or so. Much of this development was inspired by a famous conjecture posed by van der Waerden in 1926 concern-

LINEAR ALGEBRA AND ITS APPLICATIONS 127:3–25 (1990)

655 Avenue of the Americas, New York, NY 10010 0024-3795/90/$3.50

ing the minimum permanent over the set of doubly stochastic matrices. The conjecture was solved by Egorychev, and independently by Falikman, around 1980, and that led to increased activity in the area. The book *Permanents* by Minc [20] and the survey papers by Minc [21, 22] provide an excellent source of references on permanents.

The purpose of this paper is to describe various topics in probability and statistics where permanents seem to have a role. It appears that there are two main advantages of employing permanents in these areas. Firstly the permanent serves as a convenient notational device which facilitates manipulation of complicated expressions. The second advantage is more important. In all the areas described in this paper the matrices that appear are nonnegative. Thus it is possible to apply results from the theory of permanents of nonnegative matrices. For instance, several applications of Alexandroff's inequality are illustrated.

This is intended as a survey paper, and the results are stated mostly without proof, giving references to other sources for more details.

2. LOG-CONCAVE SEQUENCES

A sequence of nonnegative numbers $\alpha_1, \alpha_2, \dots, \alpha_n$ is said to be log-concave if $\alpha_i^2 \geqslant \alpha_{i-1}\alpha_{i+1}$, $i = 2,3,\dots, n-1$. Log-concave sequences arise frequently in statistics and in combinatorics. In the next result we summarize a number of elementary properties of such sequences.

LEMMA 2.1. *Let $\alpha_1, \dots, \alpha_n; \beta_1, \dots, \beta_n$ be log-concave sequences. Then the following assertions are true.*

(i) *If $\alpha_i > 0$, $i = 1,2,\dots, n$, then*

$$\frac{\alpha_i}{\alpha_{i-1}} \geqslant \frac{\alpha_{i+1}}{\alpha_i}, \qquad i = 2,\dots, n-1,$$

i.e., α_i/α_{i-1} is nonincreasing in i.

(ii) *If $\alpha_i > 0$, $i = 1,2,\dots, n$, then $\alpha_1, \dots, \alpha_n$ is unimodal, i.e., for some k, $1 \leqslant k \leqslant n$,*

$$\alpha_1 \leqslant \alpha_2 \leqslant \cdots \leqslant \alpha_k \geqslant \alpha_{k+1} \geqslant \cdots \geqslant \alpha_n.$$

(iii) *$\alpha_1\beta_1, \dots, \alpha_n\beta_n$ is log-concave.*

(iv) $\gamma_1, \dots, \gamma_n$ *is log-concave, where*

$$\gamma_k = \sum_{i=1}^{k} \alpha_i \beta_{k+1-i}, \qquad k = 1, 2, \dots, n.$$

(v) $\alpha_1, \alpha_1 + \alpha_2, \dots, \alpha_1 + \cdots + \alpha_n$ *and* $\alpha_n, \alpha_n + \alpha_{n-1}, \dots, \alpha_1 + \cdots + \alpha_n$ *are both log-concave.*

(vi) *The sequence*

$$\binom{n}{i}, \qquad i = 0, 1, \dots, n,$$

is log-concave.

We offer some remarks concerning the proof of Lemma 2.1. Assertions (i), (iii), (vi) are elementary. To prove (ii), note that by (i),

$$\frac{\alpha_2}{\alpha_1} \geqslant \frac{\alpha_3}{\alpha_2} \geqslant \cdots \geqslant \frac{\alpha_n}{\alpha_{n-1}},$$

and there must exist k, $1 \leqslant k \leqslant n$, such that

$$\frac{\alpha_2}{\alpha_1} \geqslant \cdots \geqslant \frac{\alpha_k}{\alpha_{k-1}} \geqslant 1 \geqslant \frac{\alpha_{k+1}}{\alpha_k} \geqslant \cdots \geqslant \frac{\alpha_n}{\alpha_{n-1}}.$$

Assertion (iv) can be proved by showing $\gamma_k^2 \geqslant \gamma_{k-1}\gamma_{k+1}$, and this involves making a careful pairing of terms on the two sides of the inequality. Then by taking $\beta_i = 1$, $i = 1, 2, \dots, n$, we get the first part of (v). Since $\alpha_n, \dots, \alpha_1$ is also log-concave, we have the second part of (v).

An important result in the theory of permanents of nonnegative matrices is the Alexandroff inequality, which we state next. We refer to van Lint [25] for a proof.

THEOREM 2.2. *Let* $A = (a_1, \dots, a_n)$ *be a nonnegative* $n \times n$ *matrix. Then*

$$(\operatorname{per} A)^2 \geqslant \operatorname{per}(a_1, \dots, a_{n-2}, a_{n-1}, a_{n-1}) \operatorname{per}(a_1, \dots, a_{n-2}, a_n, a_n). \tag{2.1}$$

Alexandroff's inequality has an interesting history. It was proved by A. D. Alexandroff [1] in 1938 for a more general function called the "mixed discriminant." In fact there is no mention of permanents in [1]. Then, after almost forty years, Egorychev realized that the result specializes to a permanental inequality and it provided exactly what was needed for him to complete the proof of the famous van der Waerden conjecture.

Alexandroff's inequality provides a powerful tool for producing log-concave sequences, and this will be illustrated in the present paper by means of several examples. A typical way in which one obtains log-concave sequences using (2.1) is as follows.

Suppose B is a nonnegative $n \times m$ matrix, $m < n$, and let $x \geqslant 0$, $y \geqslant 0$ be vectors in R^n. Let

$$\alpha_r = \operatorname{per}\Big(B, \underbrace{x, \dots, x}_{r}, \underbrace{y, \dots, y}_{n-m-r}\Big), \qquad r = 0, 1, \dots, n-m.$$

Then by (2.1), $\alpha_0, \alpha_1, \dots, \alpha_{n-m}$ is log-concave.

3. THE MULTIPARAMETER MULTINOMIAL DISTRIBUTION

Suppose a coin turns up heads with probability p on any single toss. Let the coin be tossed n times, and let X denote the number of heads obtained in the n tosses. Then X has the binomial distribution with parameters n, p, and the density function of X is given by

$$\Pr(X = x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}, \qquad x = 0, 1, \dots, n.$$

This density can be expressed in terms of permanents as follows:

$$\Pr(X = x) = \frac{1}{x!(n-x)!} \operatorname{per} \left[\underbrace{\begin{matrix} p & \cdots & p \\ \vdots & & \vdots \\ p & \cdots & p \end{matrix}}_{x} \quad \underbrace{\begin{matrix} 1-p & \cdots & 1-p \\ \vdots & & \vdots \\ 1-p & \cdots & 1-p \end{matrix}}_{n-x} \right]. \tag{3.1}$$

The expression (3.1) admits generalizations. For example, suppose n coins, not necessarily identical, 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, $i=1,2,\dots,n$, then it can be verified that

$$\Pr(X=x)=\frac{1}{x!(n-x)!}\operatorname{per}\begin{bmatrix}\underbrace{\begin{matrix}p_1 & \cdots & p_1\\ \vdots & & \vdots\\ p_n & \cdots & p_n\end{matrix}}_{x} & \underbrace{\begin{matrix}1-p_1 & \cdots & 1-p_1\\ \vdots & & \vdots\\ 1-p_n & \cdots & 1-p_n\end{matrix}}_{n-x}\end{bmatrix}. \tag{3.2}$$

We now consider a further generalization. Thus, suppose an experiment can result in any one of r possible outcomes, and suppose n trials of the experiment are performed. Let p_{ij} be the probability that the experiment results in the jth outcome at the ith trial, $i=1,2,\dots,n$, $j=1,2,\dots,r$. Let P denote the $n\times r$ matrix (p_{ij}), which, of course, is row-stochastic. Let X_j denote the number of times the jth outcome is obtained in the n trials, $j=1,2,\dots,r$; and let $X=(X_1,\dots,X_r)$. In this setup X is said to have the multiparameter multinomial distribution with the parameter matrix P. If the rows of P are all identical, then X has the usual multinomial distribution. Let

$$\mathcal{K}_{n,r}=\left\{k=(k_1,\dots,k_r): k_i\geqslant 0,\text{ integers, }\sum_{i=1}^{r}k_i=n\right\}$$

If A is an $m\times r$ matrix and if $k\in\mathcal{K}_{n,r}$, then $A(k)$ will denote the $m\times n$ matrix obtained by taking k_j copies of the jth column of A, $j=1,2,\dots,r$. Also, for $k\in\mathcal{K}_{n,r}$, we define $k!=k_1!\cdots k_r!$.

If X has the multiparameter multinomial distribution with the $n\times r$ parameter matrix P, then as a simple generalization of (3.2) we have

$$\Pr(X=k)=\frac{1}{k!}\operatorname{per}P(k),\qquad k\in\mathcal{K}_{n,r} \tag{3.3}$$

An alternative way of arriving at (3.3) is the following. The probability generating function of X is given by

$$\prod_{i=1}^{n}\left(\sum_{j=1}^{r}p_{ij}s_j\right). \tag{3.4}$$

Equation (3.3) thus says that for any $k\in\mathcal{K}_{n,r}$ the coefficient of $s_1^{k_1}\cdots s_r^{k_r}$ in

(3.4) is $(k!)^{-1}\operatorname{per}P(k)$. This statement is well known and is related to MacMahon's master theorem (see, for example, [6; 24, p. 15]).

We now introduce the concept of matrix majorization. If P, Q are $n \times r$ matrices, we say that Q is majorized by P $(Q \prec P)$ if Q is obtained from P by a repeated averaging of rows. To say this more formally, let $\mathscr{D}_n$ denote the class of $n \times n$ doubly stochastic matrices which can be written as a product of matrices of the form $tI+(1-t)T$, $0 \leqslant t \leqslant 1$, where T is an $n \times n$ permutation matrix that interchanges only two coordinates. Then we have the following.

DEFINITION. Let P, Q be $n \times r$ matrices. Then $Q \prec P$ if and only if $Q = DP$ for some $D \in \mathscr{D}_n$. The same concept has been called chain majorization by Marshall and Olkin [19, p. 430].

DEFINITION. A function g defined on a set of $n \times r$ matrices is said to be (multivariate) Schur-concave if $Q \prec P$ implies $g(Q) \geqslant g(P)$. Similarly g is said to be Schur-convex if $Q \prec P$ implies $g(Q) \leqslant g(P)$.

DEFINITION. Let A be a real, symmetric $n \times n$ matrix. Then A is said to be conditionally positive definite if $x'Ax \geqslant 0$ for any x satisfying $\sum_{i=1}^{n} x_i = 0$.

If $k \in \mathscr{K}_{n,r}$, then we set $k_{ij} = k + e_i + e_j$, where e_i denotes the ith row of the $r \times r$ identity matrix for any i. It is important to keep in mind that for each i, j $(i = 1, 2, \dots, r,\ j = 1, 2, \dots, r)$, k_{ij} has been defined as a vector, although the notation is perhaps a bit misleading.

DEFINITION. Let $\phi: \mathscr{K}_{n,r} \to (-\infty, \infty)$. Then ϕ is said to be positive semidefinite (conditionally positive definite) if for any $k \in \mathscr{K}_{n-2,r}$, the matrix $(\phi(k_{ij}))$ is positive semidefinite (conditionally positive definite).

Suppose $\phi(k) = \Pr(X = k)$, given in (3.3). The representation (3.3) along with some results on permanents has been used in [3] to show that $\log \phi(k)$ is conditionally positive definite. This leads to a proof of a conjecture made by Karlin and Rinott [16] concerning the Schur-concavity of the entropy of a multiparameter multinomial. We only state the result here and refer to [3, 16] for further details.

THEOREM 3.1. *If P is a row-stochastic $n \times r$ matrix, let $g(P)$ denote the entropy function*

$$g(P) = -\sum_{k \in \mathscr{K}_{n,r}} \Pr(X = k) \log \Pr(X = k),$$

where X *has the multiparameter multinomial distribution with the parameter matrix* P. *Then*

(a) g *is Schur-concave;*
(b) *if* Q *is the* $n \times r$ *matrix with* $nq_{ij} = \sum_{l=1}^{n} p_{lj}$ *for all* i, j, *then*

$$g(Q) \geqslant g(P);$$

(c) *the maximum of* $g(P)$ *over all* $n \times r$ *row-stochastic matrices* P *is attained when* $p_{ij} = 1/r$ *for all* i, j.

It is easy to see that the binomial density is log-concave. In the next result we give an analogous statement for the multiparameter multinomial.

Lemma 3.2. *Let* $X = (X_1, \dots, X_r)$ *have the multiparameter multinomial distribution with the* $n \times r$ *parameter matrix* P. *Let* $x_3, \dots, x_r$ *be nonnegative integers such that* $y = x_3 + \cdots + x_r \leqslant n$, *and let*

$$f(x) = \Pr(X_1 = x,\ X_2 = n - y - x \mid X_3 = x_3, \dots,\ X_r = x_r),$$
$$x = 0, 1, \dots, n - y.$$

Then $f(0), \dots, f(n-y)$ *is log-concave.*

Proof. Let

$$g(x) = \Pr(X_1 = x,\ X_2 = n - y - x,\ X_3 = x_3, \dots,\ X_r = x_r).$$

Then by (3.3)

$$g(x) = \frac{1}{x!(n-y-x)!x_3! \dots x_r!} \operatorname{per} P(x, n-y-x, x_3, \dots, x_r).$$

By Alexandroff's inequality, $\operatorname{per} P(x, n-y-x, x_3, \dots, x_r)$, $x = 0, 1, \dots, n-y$, is a log-concave sequence. The sequence $\{x!(n-y-x)!\}^{-1}$, $x = 0, 1, \dots, n-y$, is log-concave by Lemma 2.1(vi), and then, by Lemma 2.1(iii), $g(x)$ $(x = 0, 1, \dots, n-y)$ is log-concave.

Since

$$f(x) = \frac{1}{\Pr(X_3 = x_3, \dots,\ X_r = x_r)} g(x),$$

it follows that $f(x)$, $x = 0, 1, \dots, n-y$, is log-concave. ∎

The next result illustrates yet another application of Alexandroff's inequality.

LEMMA 3.3. *Suppose there are n coins, of which m are identical with the same probability of heads equal to p, whereas the remaining $n - m$ have the same probability of heads equal to p'. Let x be fixed, $0 \leqslant x \leqslant n$, and let $f(m)$ denote the probability of getting x heads when the n coins are tossed. Then $f(0), \dots, f(m)$ is log-concave.*

Proof. By (3.3) we have

$$f(m) = \frac{1}{x!(n-x)!} \operatorname{per} \begin{bmatrix} \underbrace{\begin{matrix} p & \cdots & p \\ \vdots & & \vdots \\ p & \cdots & p \\ p' & \cdots & p' \\ \vdots & & \vdots \\ p' & \cdots & p' \end{matrix}}_{x} & \underbrace{\begin{matrix} 1-p & \cdots & 1-p \\ \vdots & & \vdots \\ 1-p & \cdots & 1-p \\ 1-p' & \cdots & 1-p' \\ \vdots & & \vdots \\ 1-p' & \cdots & 1-p' \end{matrix}}_{n-x} \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} a \\ \vdots \\ a \end{matrix}}\right\} m \\ \left.\vphantom{\begin{matrix} a \\ \vdots \\ a \end{matrix}}\right\} n-m \end{matrix} \tag{3.5}$$

Since the permanent of a matrix is equal to that of its transpose, we can write down a version of Alexandroff's inequality in which rows are repeated instead of columns. The log-concavity of $f(0), \dots, f(m)$ then easily follows from that version in view of (3.5). ∎

Lemma 3.2 and 3.3 appear in [2], and it may be remarked that the proof of Lemma 3.3 contained there is incorrect. Part of the material covered in this and in the next section has appeared in a rather condensed form in [5, p. 201].

The relationship between permanents and the multiparameter multinomial has been noted by Gleason [10]. The main emphasis in [10] is on providing a probabilistic interpretation of the van der Waerden conjecture. We now give an outline of the interpretation. We first recall the van der Waerden conjecture (now a theorem due to Egorychev and Falikman), which asserts that the minimum permanent over the set of $n \times n$ doubly stochastic matrices is attained only at J_n, the $n \times n$ matrix will all entries $1/n$.

Consider n urns, and balls of n colors distributed in the urns in such a way that the probability of drawing a ball of color j from urn i is p_{ij}. Then clearly $P=(p_{ij})$ is a row-stochastic $n \times n$ matrix. If a ball is drawn at random from each urn, then the probability that the balls are all of different colors is $\operatorname{per} P$. The expected number of balls of color j is given by the jth column sum of P. Now suppose P is doubly stochastic. This amounts to the assumption that the expected number of balls of each color is one. Now according to the van der Waerden conjecture the probability that the balls are all of different colors is minimized when and only when the compositions of the urns are identical. Gleason [10] also gives an analogous statement when balls of only r ($\leqslant n$) colors are available.

Gyires [12, 13] has used permanents to define a class of distributions more general than the multiparameter multinomial (3.3). Before defining the class we recall the Cauchy-Binet formula for the permanent [20, p. 17].

Lemma 3.4. *Let A, B be $n \times r$ matrices. Then*

$$\operatorname{per} AB' = \sum_{k \in \mathscr{K}_{n,r}} \frac{1}{k!} \operatorname{per} A(k) \operatorname{per} B(k).$$

Definition. Let A, B be $n \times r$ nonnegative matrices, and suppose $AB' = E$, the matrix with all 1's. We say that the random vector $X = (X_1, \dots, X_r)$ has the multinomial distribution generated by A, B if

$$\Pr(X = k) = \frac{1}{n!k!} \operatorname{per} A(k) \operatorname{per} B(k), \qquad k \in \mathscr{K}_{n,r}. \tag{3.6}$$

Note that by Lemma 3.4 $\Pr(X=k)$ defined in (3.6) is a density. If A is a row-stochastic $n \times r$ matrix and if B has all entries equal to 1, then the density (3.6) reduces to the multiparameter multinomial density (3.3).

If X has density (3.6), then Gyires [12] has shown that the characteristic function of X is given by

$$\phi(t_1, \dots, t_r) = \frac{1}{n!} \operatorname{per} A \begin{bmatrix} e^{it_1} & & \\ & \ddots & \\ & & e^{it_r} \end{bmatrix} B'.$$

The proof again uses Lemma 3.4. It is then possible to obtain the moments. For example, it turns out that

$$E(X_l) = \frac{1}{n} \sum_{j=1}^{n} a_{jl} \sum_{j=1}^{n} b_{jl}, \qquad l = 1, 2, \dots, r$$

In [13] Gyires has studied the asymptotic distribution of $X^1,\dots,X^m,\dots,$ where each X^m has a density of the form (3.6). Here we give one result as an example and refer to [13] for the proof and for other related results.

Let A be a positive row-stochastic matrix with infinitely many rows and r columns. Let A^m denote the submatrix of A formed by the first m rows, $m=1,2,\dots$. Suppose for each $k=1,2,\dots,r$; the limit of a_{jk} as $j\to\infty$ exists and is positive. Let $X^m=(X_1^m,\dots,X_r^m)$ have the multiparameter multinomial distribution with parameter matrix A^m, $m=1,2,\dots$.

Let

$$B_m=\begin{bmatrix}\left(\sum_{j=1}^m a_{j1}\right)^{-1/2} & & \\ & \ddots & \\ & & \left(\sum_{j=1}^m a_{jr}\right)^{-1/2}\end{bmatrix},\qquad m=1,2,\dots.$$

Then [13, Theorem 5] the asymptotic distribution of $B_m[X^m-E(X^m)]'$ as $m\to\infty$ is multivariate normal.

We conclude this section by describing a recent result due to Boland and Proschan [7, p. 292]. A real-valued function $f(x_1,\dots,x_s)$ of n-dimensional vector arguments $x_1,\dots,x_s$ is said to be arrangement-decreasing if the function decreases in value as the components of the vectors $x_1,\dots,x_s$ become "more similarly arranged." We refer to [7] for a formal definition.

The next result is equivalent to the assertion that the permanent of a nonnegative $n\times n$ matrix is an arrangement-decreasing function of its columns (equivalently, its rows).

THEOREM 3.5. *Let A be a nonnegative $n\times n$ matrix. Let $1\leqslant l<m\leqslant n$ be fixed, and let $B=(b_{ij})$ be obtained from A by the following rule: for $k=1,2,\dots,n$, interchange a_{lk} and a_{mk} if $a_{lk}>a_{mk}$. Then* $\operatorname{per}B\leqslant\operatorname{per}A$.

Proof. The proof given here is somewhat easier than the one in [7].

First let $n=2$, so that

$$A=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}.$$

If $a_{11}\leqslant a_{21}$ and $a_{12}\leqslant a_{22}$ or if $a_{11}\geqslant a_{21}$ and $a_{12}\geqslant a_{22}$, then clearly $\operatorname{per}B=$

$\operatorname{per} A$. So suppose, without loss of generality, that $a_{11} \leqslant a_{21}$ and $a_{12} \geqslant a_{22}$. Then

$$B = \begin{bmatrix} a_{11} & a_{22} \\ a_{21} & a_{12} \end{bmatrix}.$$

We have

$$\begin{aligned} \operatorname{per} A - \operatorname{per} B &= a_{11}(a_{22} - a_{12}) + a_{21}(a_{12} - a_{22}) \\ &= (a_{11} - a_{21})(a_{22} - a_{12}) \geqslant 0. \end{aligned}$$

For $n > 2$ we may assume, without loss of generality, that $l = 1$, $m = 2$. Now expand both $\operatorname{per} A, \operatorname{per} B$ in terms of the last $n - 2$ rows and use the result obtained for 2×2 matrices. That completes the proof. ■

A probabilistic interpretation of the result, given by Boland and Proschan [7], but slightly rephrased here, is as follows. As before, consider n urns containing balls of n different colors. Let the probability of drawing a ball of color j from urn i be given by p_{ij}. Then $P = (p_{ij})$ is row-stochastic, and the probability of getting balls of n distinct colors when one ball is drawn from each urn is $\operatorname{per} P$. Let $Q = (q_{ij})$ be defined by letting $q_{i1} \geqslant \cdots \geqslant q_{in}$ be a rearrangement of $p_{i1}, \dots, p_{in}$ for each i. Since the permanent is an arrangement-decreasing function of the rows as well, $\operatorname{per} Q \leqslant \operatorname{per} P$. Thus, as we intuitively expect, the probability of getting balls of n distinct colors cannot increase.

4 THE MULTIPARAMETER NEGATIVE-MULTINOMIAL DISTRIBUTION

Suppose we have m dice, each with $r + 1$ faces. We assume that each die has one face marked with an asterisk $*$, whereas the remaining faces carry $1, 2, \dots, r$ spots. Suppose $p_{ij} > 0$ is the probability of getting j spots when the ith die is rolled, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, r$. Let P denote the $m \times r$ matrix (p_{ij}). Let p_{i0} be the probability of getting an $*$ when the ith die is rolled, $i = 1, 2, \dots, m$. Then clearly

$$p_{i0} = 1 - \sum_{j=1}^{r} p_{ij}, \qquad i = 1, 2, \dots, m.$$

The following experiment is conducted. The first die is rolled until it shows $*$. Then we switch over to the second die and roll it until an $*$ is obtained, whence we take up the third die. The process is repeated, and the experiment stops when we have rolled all the m dice, obtaining m $*$'s. Let X_j denote the number of times we get j spots in the experiment, $j = 1, 2, \dots, r$, and let $X = (X_1, \dots, X_r)$. In this setup X is said to have the multiparameter negative-multinomial distribution with the $m \times r$ parameter matrix P. Just like (3.3), we have a permanent representation for the density function of X as follows. Let $k = (k_1, \dots, k_r) \in \mathscr{K}_{n,r}$. Then

$$\Pr(X = k) = \frac{p_{10} \cdots p_{m0}}{k!} \sum \operatorname{per}\{[P(k)]'(l)\}, \tag{4.1}$$

where the summation is over all vectors $l = (l_1, \dots, l_m) \in \mathscr{K}_{n,m}$. We refer to [2] for a proof.

In this situation a result analogous to Theorem 3.1 is not yet proved. It has been conjectured by Karlin and Rinott [16] that the entropy function in this case is Schur-convex. To settle this conjecture it must be shown that $-\log \phi(k)$ is conditionally positive definite on $\mathscr{K}_{n,r}$ for any n, where $\phi(k) = \Pr(X = k)$ (see [16, Conjecture 3.2]). We now give a proof of a slightly different assertion. It will be shown that $\phi(k) = k! \Pr(X = k)$ is positive semidefinite on $\mathscr{K}_{n,r}$ for any n. The result appears to be new.

If A is an $n \times r$ matrix, let $\overline{\operatorname{per}} A$ denote the sum of all $n \times n$ subpermanents of A, where repetition of columns is permitted. More formally, we define

$$\overline{\operatorname{per}}\, A = \sum_{l \in \mathscr{K}_{n,r}} \operatorname{per} A(l).$$

Then (4.1) can be expressed as

$$\Pr(X = k) = \frac{p_{10} \cdots p_{m0}}{k!} \overline{\operatorname{per}}\, P(k)'. \tag{4.2}$$

THEOREM 4.1. *Let $\phi \colon \mathscr{K}_{n,r} \to (-\infty, \infty)$, and for any $n \times r$ matrix A, let*

$$\psi(A) = \sum_{k \in \mathscr{K}_{n,r}} \frac{\phi(k)}{k!} \operatorname{per} A(k).$$

(i) *If ϕ is conditionally positive definite on $\mathscr{K}_{n,r}$, then ψ is Schur-concave on the set of $n \times r$ row-stochastic matrices.*

(ii) *If ϕ is positive semidefinite on $\mathscr{K}_{n,r}$, then ψ is Schur-concave on the set of $n \times r$ nonnegative matrices.*

Theorem 4.1(i) is the same as Theorem 2.1 in [16], and (ii) can be proved along similar lines.

THEOREM 4.2. *Let X have the multiparameter negative multinomial distribution with the $m \times r$ parameter matrix P. Then for any positive integer $n \geq 2$ and for any $k \in \mathscr{K}_{n-2,r}$, the matrix $(k_{ij}!\Pr(X = k_{ij}))$ is positive semidefinite.*

Proof. We will first show that the quadratic form

$$\xi(x) = \overline{\text{per}}\,(A, x, x)', \qquad x \in R^m,$$

is positive semidefinite for any nonnegative $m \times (n-2)$ matrix A.

A simple calculation shows that $\phi(\underline{k}) = k!$ is positive semidefinite on $\mathscr{K}_{n,r}$. Hence by Theorem 4.1(ii), $\psi(b) = \overline{\text{per}}\,B$ is multivariate Schur-concave on the set of nonnegative matrices. Therefore for any nonnegative $m \times (n-2)$ matrix A and for any nonnegative vectors x, y,

$$\overline{\text{per}}\left(A, \frac{x+y}{2}, \frac{x+y}{2}\right)' \geq \overline{\text{per}}\,(A, x, y)'. \tag{4.3}$$

Since $\overline{\text{per}}$ is a multilinear function of the rows, (4.3) simplifies to

$$\overline{\text{per}}\left(A, \frac{x-y}{2}, \frac{x-y}{2}\right)' \geq 0,$$

and this is equivalent to showing that ξ is positive semidefinite.

Now let P_j denote the jth column of P, so that $P = (P_1, \ldots, P_r)$. Let e_i denote the ith column of the $m \times m$ identity matrix, $i = 1, 2, \ldots, m$, and let $H = (h_{ij})$ be the $m \times m$ matrix defined as

$$h_{ij} = \text{per}\big(P(k), e_i, e_j\big)', \qquad i, j = 1, 2, \ldots, m.$$

Then clearly $\xi(x) = x'Hx$, and hence H is positive semidefinite.

Now observe that

$$\begin{aligned}
\big(k_{ij}!\Pr(X = k_{ij})\big) &= p_{10} \cdots p_{m0}\, \overline{\text{per}}\, P(k_{ij})' \\
&= p_{10} \cdots p_{m0}\, \overline{\text{per}}\,\big(P(k), P_i, P_j\big)' \\
&= p_{10} \cdots p_{m0}(PHP').
\end{aligned}$$

Since H is positive semidefinite, so is PHP′, and the proof is complete. ∎

By Theorem 4.2 any 2×2 principal minor of $(k_{ij}! \Pr(X = k_{ij}))$ must be nonnegative, and hence we obtain the inequality

$$\left\{k_{ij}! \Pr(X = k_{ij})\right\}^2 \leqslant \left\{k_{ii}! \Pr(X = k_{ii})\right\}\left\{k_{jj}! \Pr(X = k_{jj})\right\}. \quad (4.4)$$

Suppose Z_i, $i = 1, 2, \dots, m$, are independent random variables following a simple geometric distribution such that

$$P(Z_i = k) = p_i^k q_i, \qquad k = 0, 1, 2, \dots, \qquad q_i = 1 - p_i.$$

Then $Z = Z_1 + \cdots + Z_m$ has the multiparameter negative binomial distribution. It is known [15, p. 104; 14, p. 164] that

$$\Pr(Z = k)^2 \leqslant \frac{(k+1)(n+k-1)}{k(n+k)} \Pr(Z = k-1) \Pr(Z = k+1).$$

In particular,

$$\left\{k! \Pr(Z = k)\right\}^2 \leqslant \left\{(k-1)! \Pr(Z = k-1)\right\}\left\{(k+1)! \Pr(Z = k+1)\right\}. \quad (4.5)$$

Thus (4.4) can be thought of as a multivariate analog of (4.5).

5. ORDER STATISTICS

Permanents provide an effective tool in dealing with order statistics corresponding to random variables which are independent but possibly nonidentically distributed. Let $X_1, \dots, X_n$ be independent random variables with distribution functions $F_1, \dots, F_n$ respectively, and let $Y_1 \leqslant \cdots \leqslant Y_n$ denote the corresponding order statistics. Suppose the densities $f_1, \dots, f_n$ of $X_1, \dots, X_n$ exist. Vaughan and Venables [26] have shown that the density of any order statistic or the joint density of several order statistics is conveniently expressed in terms of a permanent. For example, the density of Y_r is given by

$$g_r(y) = \frac{1}{(r-1)!(n-r)!}$$

$$\times \operatorname{per}\left[\begin{matrix} f_1(y) & F_1(y) & \cdots & F_1(y) & 1-F_1(y) & \cdots & 1-F_1(y) \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ f_n(y) & \underbrace{F_n(y)}_{} & \underbrace{\cdots}_{r-1} & \underbrace{F_n(y)}_{} & \underbrace{1-F_n(y)}_{} & \underbrace{\cdots}_{n-r} & \underbrace{1-F_n(y)}_{} \end{matrix}\right]. \tag{5.1}$$

Similarly the distribution function of Y_r or that of a subset of $Y_1, \ldots, Y_n$ may be expressed in terms of permanents. For example, as indicated in [4], the distribution function of Y_r is given by

$$\Pr(Y_r \leqslant y) = \sum_{i=r}^{n} \frac{1}{i!(n-i)!}$$

$$\times \operatorname{per}\left[\begin{matrix} F_1(y) & \cdots & F_1(y) & 1-F_1(y) & \cdots & 1-F_1(y) \\ \vdots & & \vdots & \vdots & & \vdots \\ \underbrace{F_n(y)}_{} & \underbrace{\cdots}_{i} & \underbrace{F_n(y)}_{} & \underbrace{1-F_n(y)}_{} & \underbrace{\cdots}_{n-i} & \underbrace{1-F_n(y)}_{} \end{matrix}\right]. \tag{5.2}$$

There are a number of recurrence relations for order statistics in the literature. Almost all of these are for the case of i.i.d. $X_1, \ldots, X_n$. Using (5.1), (5.2) it is possible to get refined versions of some of these recurrence relations corresponding to the case of nonidentical $X_1, \ldots, X_n$. The proof usually involves simple manipulations with the permanent using the Laplace expansion. We refer to [4] for some examples.

If it is desired to incorporate the possibility of one or more outliers being present in $X_1, \ldots, X_n$, then one naturally arrives at the situation where $X_1, \ldots, X_n$ are nonidentically distributed. It is common practice to restrict the analysis to the case of one outlier, since for more outliers the treatment becomes complicated. The permanent representations may be of help in making further progress. In some instances $F_1, \ldots, F_n$ may be believed to be of the same functional form but with different values of the parameters involved. For example, $X_1, \dots, X_n$ may all be exponential random variables with parameters $\lambda_1, \dots, \lambda_n$ respectively. Then some further simplification results while manipulating the permanents in (5.1), (5.2). We refer to Gross, Hunt, and Odeh [11] for an example.

The next result from [4] illustrates more applications of Alexandroff's inequality.

THEOREM 5.1. *Let $X_1, \dots, X_n$ be independent random variables with distribution functions $F_1, \dots, F_n$ respectively. Let $Y_1 \leqslant \cdots \leqslant Y_n$ be the corresponding order statistics with respective distribution functions $G_1, \dots, G_n$. Let $y \in (-\infty, \infty)$ be fixed. Then $G_1(y), \dots, G_n(y)$ and $1 - G_1(y), \dots, 1 - G_n(y)$ are* log-*concave. Furthermore if $X_1, \dots, X_n$ are continuous with respective densities $f_1, \dots, f_n$, then $g_1(y), \dots, g_n(y)$ is* log-*concave, where $g_1, \dots, g_n$ are the densities of $Y_1, \dots, Y_n$ respectively.*

Proof. Let

$$\alpha_i = \operatorname{per}\left[\underbrace{\begin{matrix} F_1(y) & \cdots & F_1(y) \\ \vdots & & \vdots \\ F_n(y) & \cdots & F_n(y) \end{matrix}}_{i} \quad \underbrace{\begin{matrix} 1-F_1(y) & \cdots & 1-F_1(y) \\ \vdots & & \vdots \\ 1-F_n(y) & \cdots & 1-F_n(y) \end{matrix}}_{n-i}\right],$$

$$i = 0, 1, \dots, n.$$

By Alexandroff's inequality $\alpha_0, \alpha_1, \dots, \alpha_n$ is log-concave. Since $[i!(n-i)!]^{-1}$, $i = 0, 1, \dots, n$, is also log-concave, by Lemma 2.1(iii)

$$\frac{1}{i!(n-i)!}\alpha_i, \qquad i = 0, 1, \dots, n,$$

is log-concave. It follows that $G_1(y), \dots, G_n(y)$ and $1 - G_1(y), \dots, 1 - G_n(y)$ are log-concave in view of (5.2) and Lemma 2.1(v). Similarly, (5.1) and Alexandroff's inequality lead to the log-concavity of $g_1(y), \dots, g_n(y)$. ∎

The log-concavity of $G_1(y), \dots, G_n(y)$ has another interpretation. Suppose $G_i(y) > 0$, $i = 1, 2, \dots, n$. Then

$$\frac{G_i(y)}{G_{i-1}(y)} = \frac{\Pr(Y_i \leqslant y)}{\Pr(Y_{i-1} \leqslant y)} = \Pr(Y_i \leqslant y \mid Y_{i-1} \leqslant y).$$

Thus by Lemma 2.1(i) we conclude that the sequence of conditional probabilities $\Pr(Y_i \leqslant y \mid Y_{i-1} \leqslant y)$, $i = 1,2,\dots,n$, is nonincreasing. This observation may be of some interest in the context of reliability theory.

6. SEQUENTIAL EXPERIMENTS WITH FEEDBACK

Consider a deck of n cards containing a_i identical cards of type i, $i = 1,2,\dots,r$, where $a_1 + \cdots + a_r = n$. The deck is well shuffled, and the cards are presented to a person, face down, in sequence. The person tries to guess the type of each card. The number of correct guesses can then be used to test whether the person has ESP (extrasensory perception).

Various authors have tried to introduce some kind of feedback in the process. A kind of partial feedback, which we consider here, is to tell the person whether his guess was right or wrong immediately after each guess. Diaconis and Graham [9] and Chung, Diaconis, Graham, and Mallows [8] deal with this setup. One of the main results in [8] can be informally stated as follows. The probability that the next card is type i, at any stage of the experiment, does not decrease if it was guessed that the previous card is type i and the guess is incorrect. We wish to indicate that this result is an immediate consequence of Alexandroff's inequality.

We in fact consider a more general situation. Consider a deck of n cards marked with integers $1,2,\dots,n$. The cards are presented to the subject, face down, in a random order. At each stage the subject is required to produce a subset of $\{1,2,\dots,n\}$ and guess that the card is in that subset. In the next result we show the following. Let $S \subset \{1,2,\dots,n\}$. Suppose at any given stage of the experiment the probability that the next card is in S, given the feedback, is α. Now if the subject guesses that the card is in S and is proved wrong, then the probability that the following card is in S cannot be less than α. The result of Chung et al. stated earlier is clearly a consequence of this result obtained by restricting S to the set of cards of type i.

THEOREM 6.1. *Let $S_1, S_2,\dots$ be arbitrary nonempty subsets of $\{1,2,\dots,n\}$, and let S_i^j be the event that the jth card is in S_i. Then*

$$\Pr\left(S_k^k \middle| S_1^1,\dots,S_{k-1}^{k-1}\right) \geqslant \Pr\left(S_k^{k+1} \middle| S_1^1,\dots,S_k^k\right).$$

Proof. For any $S \subset \{1,2,\dots,n\}$ let $e(S)$ denote the 0-1 vector of length n such that its ith entry is 1 if and only if $i \in S$, $i = 1,2,\dots,n$. Also let e denote the vector of all 1's. Then

$$\Pr\left(S_k^k \middle| S_1^1,\dots,S_k^{k-1}\right) = \frac{\Pr\left(S_1^1,\dots,S_k^k\right)}{\Pr\left(S_1^1,\dots,S_{k-1}^{k-1}\right)}$$

and

$$\Pr\left(S_k^{k+1} \middle| S_1^1,\dots,S_k^k\right) = \frac{\Pr\left(S_1^1,\dots,S_k^k,S_k^{k+1}\right)}{\Pr\left(S_1^1,\dots,S_k^k\right)}.$$

Note that

$$\Pr\left(S_1^1,\dots,S_{k-1}^{k-1}\right) = \frac{1}{n!}\operatorname{per}\big(e(S_1),\dots,e(S_{k-1}),\underbrace{e,\dots,e}_{n-k+1}\big),$$

$$\Pr\left(S_1^1,\dots,S_k^k\right) = \frac{1}{n!}\operatorname{per}\big(e(S_1),\dots,e(S_k),\underbrace{e,\dots,e}_{n-k}\big),$$

$$\Pr\left(S_1^1,\dots,S_k^k,S_k^{k+1}\right) = \frac{1}{n!}\operatorname{per}\big(e(S_1),\dots,e(S_k),e(S_k),\underbrace{e,\dots,e}_{n-k-1}\big)$$

Hence the result follows by Alexandroff's inequality. ■

In the inequality of Theorem 6.1, if we subtract both sides from 1, we get

$$\Pr\left(k\text{th card not in } S_k \middle| S_1^1,\dots,S_{k-1}^{k-1}\right) \leqslant \Pr\left((k+1)\text{th card not in } S_k \middle| S_1^1,\dots,S_k^k\right).$$

Now setting S_k equal to the complement of S in $\{1,2,\dots,n\}$, we get the interpretation of the result given before Theorem 6.1.

Consider two decks of n cards each: deck 1 containing a_i cards of type i, deck 2 containing b_i cards of type i, $i = 1,2,\dots,r$. Both decks are shuffled, and cards turned up in pairs simultaneously. Let X denote the resulting number of matches. Let $A(x)$ denote the $n \times n$ matrix constructed as follows. There are r disjoint principal blocks in $A(x)$ of size $a_i \times b_i$, $i = 1,2,\dots,r$, consisting of all x's; the remaining entries of $A(x)$ are all 1's. It has been observed by Olds [23] that the probability generating function of X is given by

$$\phi(x) = \frac{1}{n!}\operatorname{per}\mathrm{A}(\mathrm{x}).$$

Olds [23] has used the permanent representation to find the first few moments of X.

Some properties of $\operatorname{per} A(0)$ have been obtained by Chung et al. [8]. The results include exact expressions for $\operatorname{per} A(0)$, recurrence relations, and certain inequalities.

7. SAMPLING

Consider a population of N units denoted by $\mathscr{P} = \{1,2,\dots,N\}$. We consider a sample of size $n \leqslant N$ drawn from the population and denote the observations $X_1,\dots,X_n$. We will denote by E_0 and E_1 expectations under sampling without replacement and under sampling with replacement respectively.

For $i = 1,2,\dots,n$, let $\phi_i : \mathscr{P} \to (-\infty,\infty)$. Let $A = (a_{ij})$ be the $N \times N$ matrix defined as follows.

$$a_{ij} = \begin{cases} \phi_i(j), & 1 \leqslant i \leqslant n, \\ 1, & n < i \leqslant N. \end{cases} \tag{7.1}$$

It can be easily verified that

$$E_0\left\{ \prod_{i=1}^{n} \phi_i(X_i) \right\} = \frac{1}{N!} \operatorname{per} A, \qquad E_1\left\{ \prod_{i=1}^{n} \phi_i(X_i) \right\} = \frac{1}{N^N} \prod_{i=1}^{N} \left(\sum_{j=1}^{N} a_{ij} \right). \tag{7.2}$$

We now prove a preliminary result.

LEMMA 7.1. *Let $B = (b_1,\dots,b_N)$ be a nonnegative $N \times N$ matrix such that $b_1 \geqslant b_2$. Let $0 \leqslant t \leqslant 1$, and define C as*

$$C = \left(tb_1 + (1-t)b_2, (1-t)b_1 + tb_2, b_3,\dots,b_N\right).$$

Then $\operatorname{per} B \leqslant \operatorname{per} C$.

Proof. We have

$$\begin{aligned} \operatorname{per} C - \operatorname{per} B &= t(1-t)\left\{\operatorname{per}(b_1,b_1,b_3,\dots,b_N) + \operatorname{per}(b_2,b_2,b_3,\dots,b_N)\right\} \\ &\quad - 2t(1-t)\operatorname{per} B \\ &= t(1-t)\operatorname{per}(b_1 - b_2, b_1 - b_2, b_3,\dots,b_N) \geqslant 0, \end{aligned}$$

since $b_1 \geqslant b_2$. ∎

The next result has been proved by Karlin and Rinott [17, p. 39]. We give a proof using permanents.

THEOREM 7.2. *Let $\phi_i : \mathscr{P} \to [0,\infty)$ be nondecreasing functions, $i = 1,2,\ldots,n$. Then*

$$E_0\left\{\prod_{i=1}^{n}\phi_i(X_i)\right\} \leqslant E_1\left\{\prod_{i=1}^{n}\phi_i(X_i)\right\}.$$

Proof. Define the $N \times N$ matrix A as in (7.1). In view of (7.2) we must show that

$$\operatorname{per} A \leqslant \operatorname{per}(\theta,\ldots,\theta)$$

where θ is a vector with $\theta_i = (1/N)\sum_{j=1}^{N}a_{ij}$, $i = 1,2,\ldots,N$.

Let P^k denote the $N \times N$ permutation matrix corresponding to the transposition which interchanges coordinates k and $k+1$, $k = 1,2,\ldots,N-1$. Then it can be verified that

$$M = \prod_{k=1}^{N}\tfrac{1}{2}(I + P^k)$$

$$= \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \cdots & \frac{1}{2^{N-2}} & \frac{1}{2^{N-1}} & \frac{1}{2^{N-1}} \\ \frac{1}{2} & \frac{1}{4} & \cdots & \frac{1}{2^{N-2}} & \frac{1}{2^{N-1}} & \frac{1}{2^{N-1}} \\ & \ddots & \ddots & \vdots & \vdots & \vdots \\ & & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \\ & & & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & & & & \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Then $M \in \mathscr{D}_N$ and, as shown by Marcus and Newman [18, p. 67],

$$\lim_{s\to\infty} M^s = J_N,$$

where J_N is the matrix with all entries 1/N.

Since ϕ_i are nondecreasing, A satisfies $a_1 \leqslant \cdots \leqslant a_N$, and the same holds for all AM^s. Hence by Lemma 7.1

$$\operatorname{per} A \leqslant \operatorname{per} AM^s, \qquad s = 1, 2 \ldots .$$

Letting $s \to \infty$, the result is proved. ■

Karlin and Rinott [17] also discuss the following generalized birthday problem. Consider a group if n individuals. Let $A = (a_{ij})$ be the $n \times 365$ matrix where a_{ij} is the probability that the ith person's birthday is on the jth day of the year. Let $\psi(A)$ denote the probability that the n persons have n distinct birthdays. Suppose that, after renumbering the days if necessary, $A = (a_1, \ldots, a_{365})$ satisfies

$$a_1 \geqslant a_2 \geqslant \cdots \geqslant a_{365}.$$

Then $\psi(A)$ is maximized if all days are equally likely birthdays for all individuals. This result may also be proved using permanents. Augment A to a 365×365 matrix by adding rows of all 1's. Then

$$\psi(A) = \frac{1}{(365 - n)!} \operatorname{per} A. \tag{7.3}$$

Now the proof is similar to that of Theorem 7.2.

We conclude by pointing out an error which occurs in Karlin and Rinott's [17] paper. Let us say that the columns of an $n \times n$ matrix A are *similarly ordered* if, after a renumbering of the columns if necessary, they satisfy $a_1 \leqslant a_2 \leqslant \cdots \leqslant a_n$. It is clear from Lemma 7.1 that if the columns of a nonnegative matrix A are similarly ordered and if

$$B = A[tI + (1 - t)T], \tag{7.4}$$

where $0 \leqslant t \leqslant 1$ and T is a permutation matrix that interchanges only two coordinates, then

$$\operatorname{per} A \leqslant \operatorname{per} B. \tag{7.5}$$

However, note that the columns of B may no longer be similarly ordered, and thus it cannot be concluded, as has been erroneously done in Theorem 3 of [17], that $A \prec C \Rightarrow \operatorname{per} A \leqslant \operatorname{per} C$ if A is a nonnegative matrix with similarly ordered columns. Of course, the proof of Theorem 3 given in [17] can be used to conclude (7.5) provided B is as in (7.4). A similar remark applies to the discussion of the birthday problem. Thus the assertion [17, p. 40, line 10] that $\psi(A)$ of (7.3) is multivariate Schur concave is not valid.

REFERENCES

1 A. D. Alexandroff, Zur Theorie der gemischten Volumina von konvexen Korpern IV (in Russian with German summary), *Mat. Sb.* 3(45):227–251 (1938).
2 R. B. Bapat, Discrete multivariate distributions and generalized log-concavity, *Sankhyā Ser. A* 50:98–110 (1988).
3 R. B. Bapat, Multinomial probabilities, permanents and a conjecture of Karlin and Rinott, *Proc. Amer. Math. Soc.* 102(3):467–472 (1988).
4 R. B. Bapat and M. I. Beg, Order statistics for nonidentically distributed variables and permanents, *Sankhyā Ser. A* 51:79–93 (1989).
5 LeRoy Beasley and E. E. Underwood, The 1987 Utah State University Department of Mathematics conference, 29, 30, and 31 January 1987, Logan, Utah, *Linear Algebra Appl.* 104:165–249 (1988).
6 Natalia Bebiano, On the evaluation of permanents, *Pacific J. Math.* 101:1–9 (1982).
7 Philip J. Boland and Frank Proschan, Multivariate arrangement increasing functions with applications in probability and statistics, *J. Multivariate Anal.* 25:286–298 (1988).
8 F. R. K. Chung, P. Diaconis, R. L. Graham, and C. L. Mallows, On the permanents of complements of the direct sum of identity matrices, *Adv. in Appl. Math.* 2:121–137 (1981).
9 P. Diaconis and R. L. Graham, The analysis of sequential experiments with feedback to subjects, *Ann. Statist.* 9:3–23 (1981).
10 Andrew M. Gleason, Remarks on the van der Waerden permanent conjecture, *J. Combin. Theory* 8:54–64 (1970).
11 A. J. Gross, H. H. Hunt, and R. E. Odeh, The correlation coefficient between the smallest and largest observations when $(n-1)$ of the n observations are i.i.d. exponentially distributed, *Comm. Statist. Theory Methods* 15:1113–1123 (1986).
12 Bela Gyires, Discrete distribution and permanents, *Publ. Math. Debrecen* 20:93–106 (1973).
13 Bela Gyires, On the asymptotic behaviour of the generalized multinomial distribution, in *Recent Developments in Statistics*, North-Holland, Amsterdam, 1977, pp. 461–470.
14 G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities*, Cambridge U.P., 1934.
15 Samuel Karlin and Yosef Rinott, Entropy inequalities for classes of probability distributions I. The univeriate case, *Adv. Appl. Probab.* 13:93–112 (1981).
16 Samuel Karlin and Yosef Rinott, Entropy inequalities for classes of probability distributions II. The multivariate case, *Adv. Appl. Probab.* 13:325–351 (1981).
17 Samuel Karlin and Yosef Rinott, Random replacement schemes and multivariate majorization, in *Inequalities in Statistics and Probability*, IMS Lecture Notes—Monograph Ser. 5, 1984, pp. 35–40.

18 Marvin Marcus and Morris Newman, On the minimum of the permanent of a doubly stochastic matrix, *Duke Math. J.* 26:61–72 (1959).
19 A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic, New York, 1979.
20 H. Minc, *Permanents*, Addison-Wesley, Reading, Mass., 1978.
21 H. Minc. Theory of permanents 1978–1981, *Linear and Multilinear Algebra* 12:227–263 (1983).
22 H. Minc. Theory of permanents 1982–1986, *Linear and Multilinear Algebra* 21:109–148 (1987).
23 E. G. Olds, A moment generating function which is useful in solving certain matching problems, *Bull. Amer. Math. Soc.* 44:407–413 (1928).
24 J. K. Percus, *Combinatorial Methods*, Springer-Verlag, New York, 1969.
25 J. H. van Lint, Notes on Egoritsjev's proof of the van der Waerden conjecture, *Linear Algebra Appl.* 39:1–8 (1981).
26 R. J. Vaughan and W. N. Venables, Permanent expressions for order statistics densities, *J. Roy. Statist. Soc.* 34:308–310 (1972).

Received 15 June 1988; final manuscript accepted 11 May 1989