DEMONSTRATIO MATHEMATICA
Vol. XXXIV No 2 2001

Grzegorz A. Rempała

THE MARTINGALE DECOMPOSITION AND APPROXIMATION THEOREMS FOR A GENERALIZED RANDOM PERMANENT FUNCTION

Dedicated to Professor Kazimierz Urbanik

Abstract. The paper presents some recent developments in the theory of permanents for random matrices of independent columns. In particular, it is shown that the theory in a natural way incorporates and extends that of U-statistics of iid real random variables. An extension of the famous martingale decomposition (or H-decomposition) for U-statistics to a certain class of matrix functionals, which includes in particular a classical permanent function, is given.

1. Introduction

Let $A = [a_{ij}]$ be a real $m \times n$ matrix with $m \leq n$. The permanent of matrix A is defined by

$$\text{Per } A \stackrel{def}{=} \sum_{1 \leq i_1 \neq \ldots \neq i_m \leq n} \prod_{s=1}^{m} a_{s,i_s}.$$

The permanent function has a long history, having been first introduced by Cauchy in 1812 in his celebrated memoir on determinants and, almost simultaneously, by Binet (1812). More recently several problems in statistical mechanics, quantum field theory, and chemistry as well as enumeration problems in combinatorics and linear algebra have been reduced to the computation of a permanent. Unfortunately, the fastest known algorithm for computing a permanent of an $n \times n$ matrix runs, as shown by Ryser (1963), in $O(n2^n)$ time. Moreover, strong evidence for the apparent complexity of the problem was provided by Valiant (1979), who showed that evaluating a permanent is $\#P$-complete, even when restricted to $0-1$ matrix.

AMS 2000 Subject Classification: Primary 60F05; Secondary 15A15, 15A52.

In this work we will be concerned with *a random permanent* function which often arises naturally in statistical physics or statistical mechanics problems, when the investigated physical phenomenon is driven by some random process, and thus is stochastic in nature. In the present paper we develop the idea of approximating a certain large class of random functionals (which contains in particular a random permanent function) by a sum of uncorrelated martingales. Based on this representation we further develop a method for the (stochastic) approximation for the members of the class by a sum of certain iid random variables. The method is closely related to that used in approximating generalized averages, or so called *U-statistics* and in fact can be viewed as its natural extension. The paper is organized as follows. In the reminder of this section we briefly recall some basic facts from the general U-statistics theory (for details see, for instance, Lee 1990), in Section 2 we define the class of random functions of interest as well as state the result on their decomposition into a sum of uncorrelated martingales. In Section 3 we use this decomposition to obtain the variance formula which in turn enables us to prove the approximation result.

1.1. U-statistics

Let $Y_1, \dots Y_l$ be equidistributed, independent random elements taking values in some metric space $\mathcal{I}$ and let $h^{(k)}(y_1, \dots, y_k)$ be a measurable and symmetric kernel function $h^{(k)} : \mathcal{I}^k \to \mathbf{R}$. We shall also assume that $k \le l$ and, whenever it doesn't lead to ambiguity, we shall write h, in place of $h^{(k)}$.

For any such h we define a symmetrization operator $\pi_k^l(h)$ as

$$\pi_k^l(h) \stackrel{def}{=} \sum_{1 \le s_1 < \dots < s_k \le l} h(Y_{s_1}, \dots, Y_{s_k}) \tag{1}$$

and a conditional expectation of h with respect to $\sigma(Y_1, \dots, Y_c)$

$$h_c \stackrel{def}{=} E(h | \sigma(Y_1, \dots, Y_c))$$

for $c = 1, \dots, k-1$.

DEFINITION 1. For any kernel function h we shall define a corresponding *U-statistic* $U_l^{(k)}(h)$ as

$$U_l^{(k)}(h) = \binom{l}{k}^{-1} \pi_k^l(h).$$

In the sequel, whenever it is not ambiguous, we shall write $U_l^{(k)}$ for $U_l^{(k)}(h)$.

One of the most important tools in investigating U-statistics is the so called *H-decomposition.* We give its quick overview below. Let $\mathcal{B}(\mathcal{I})$ denote the Borel σ-field in the metric space $\mathcal{I}$. For any probability measure Q_i $(1 \le i \le l)$ on $(\Omega, \mathcal{B}(\mathcal{I}))$ we define

$$Q_1 \cdots Q_k h \stackrel{def}{=} \int \cdots \int h(y_1, \ldots, y_k)\, dQ_1(y_1) \cdots dQ_k(x_k).$$

Now, by expanding the simple identity (where δ_x is a Dirac probability measure at x and P is the distribution law of the Y's)

$$h(y_1, \ldots, y_k) = \prod_{i=1}^{k} (\delta_{y_i} - P + P)h$$

in a "binomial-like" fashion and noting that the appropriate operators commute, we arrive at the identity

$$h(Y_1, \ldots, Y_k) - Eh(Y_1, \ldots, Y_k) = \sum_{c=1}^{k} \pi_c^k(g_c) \tag{2}$$

where

$$g_c(y_1, \ldots, y_c) \stackrel{def}{=} \int_{\mathcal{I}} \cdots \int_{\mathcal{I}} h_c(z_1, \ldots, z_c) \prod_{s=1}^{c} (\delta_{y_s}(dz_s) - P(dz_s)).$$

Let us note that for $c = 1 \ldots k$ the g_c's are symmetric functions (often called *canonical*), satisfying

$$\int_{\mathcal{I}} g_c(y_1, \ldots, y_{c-1}, z) P(dz) = 0. \tag{3}$$

The direct application of the identity (2) to the kernel h along with the change of the order of summation gives the H-decomposition formula of Hoeffding (cf. e.g., Lee 1990)

$$U_l^{(k)} - E\, U_l^{(k)} = \sum_{c=r}^{k} \binom{k}{c} U_{c,l} \tag{4}$$

where

$$\begin{aligned} U_{c,l} &\stackrel{def}{=} \binom{l}{c}^{-1} \pi_c^l(g_c) \\ &= \binom{l}{c}^{-1} \sum_{1 \le i_1 < \ldots < i_c \le l} g_c(Y_{i_1}, \cdots, Y_{i_c}) \end{aligned}$$

and $1 \le r \le k$. The number $r - 1$ is usually called *the degree of degeneration* of $U_l^{(k)}$, whereas $k - r + 1$ is known as *the order* of $U_l^{(k)}$. We say that $U_l^{(k)}$ is of *infinite order*, if $k - r + 1 \to \infty$ as $l \to \infty$.

The most important features of the H-decomposition are the properties of its components $U_{c,l}$. More precisely, under our assumptions on the $Y's$, for a fixed $c \ge 1$,

(i) $U_{c,l}$ is itself a U-statistic with kernel function g_c.

(ii) Define $\mathcal{F}_{c,l} = \sigma\{U_{c,l}, U_{c,l+1}, \ldots\}$. Then

$$E(U_{c,l}|\mathcal{F}_{c,l+1}) = U_{c,l+1}\ , \quad \forall\, l = l_0, l_0+1, \ldots$$

That is, $(U_{c,l}, \mathcal{F}_{c,l})_{l=l_0,l_0+1,\ldots}$ is a backward martingale.

(iii) Additionally, we also have

$$Cov(U_{c_1,l}, U_{c_2,l}) = 0 \qquad \text{for } c_1 \neq c_2.$$

For the proof of the last two properties, see e.g., Lee (1990).

1.2. U-statistics and random permanents

An obvious connection between the permanent and the U-statistic has been noted first by Borovskikh and Koroljuk (1994) and may be described as follows. Let $X_1, \ldots, X_l$ be iid real random variables with a finite coefficient of variation γ.

Observe that for $k \times l$ matrix obtained from k row-replicas of $[X_1, \ldots X_l]$ (*one dimensional projection matrix*)

$$Per\,\underline{X} \stackrel{def}{=} Per \begin{bmatrix} X_1 & X_2 & \ldots & X_l \\ X_1 & X_2 & \ldots & X_l \\ \ldots & \ldots & \ldots & \ldots \\ X_1 & X_2 & \ldots & X_l \end{bmatrix} = k\,!\,\pi_k^l\Big(\prod_{i=1}^{k} X_i\Big).$$

Equivalently,

$$(5) \qquad k\,!^{-1}\binom{l}{k}^{-1} Per\,\underline{X} \;=\; P_l^{(k)}$$

$$\stackrel{def}{=} \binom{l}{k}^{-1} \pi_k^l\Big(\prod_{i=1}^{k} X_i\Big).$$

Thus, a normalized permanent of a one dimensional projection matrix may be viewed as a U-statistic $P_l^{(k)}$ corresponding to the product kernel $h = x_1 \cdots x_k$.

Based on the above observation, one can show with the help of the H-decomposition that for permanents of one dimensional projection matrices we have the following

PROPOSITION 1. *(van Es and Helmers 88 and Borovskikh and Koroljuk 94) Let $\underline{X}$ be an $k \times l$ projection matrix and let $\mathcal{N}(\alpha,\beta)$ and $\mathcal{LN}(\alpha,\beta)$ denote respectively, the normal and the lognormal laws with mean α and variance β. Furthermore, let $\stackrel{D}{\to}$ denote convergence in law:*

(i) If $k^2/l \to 0$ then

$$\frac{\sqrt{l}}{k}\left(\frac{Per\,\underline{X}}{E\,Per\,\underline{X}} - 1\right) \stackrel{D}{\to} \mathcal{N}(0,\gamma^2).$$

(ii) If $k^2/l \to \lambda > 0$ then

$$\frac{Per\,\underline{X}}{E\,Per\,\underline{X}} \xrightarrow{D} \mathcal{LN}(-\lambda\gamma^2/2, \lambda\gamma^2).$$

(iii) If $k^2/l \to \infty$ then no limit exists.[1] ∎

The above theorem indicates that the limiting behavior of the random permanent function depends on the relationships between the dimensions of the matrix. Further, it also points out that for some special types of matrices the theory of U-statistics of infinite order may provide some insight into the behavior of the random permanent. This was first noticed by Rempała and Wesołowski (1999) who have used the idea to derive an analogue of Proposition 1 for the random permanents based on matrices of iid entries. We shall further expand on this concept in the next section.

2. P-statistics

DEFINITION 1. For the sake of simplicity consider in the sequel $\mathcal{I} = \mathbf{R}$ and let $\mathbf{X}$ be a real random matrix of m rows and n columns with $m \leq n$. Denoting the i-th row vector by $X_{(i)}$ and the j column vector by $X^{(j)}$ we have

$$\begin{aligned}\mathbf{X} &= [X_{(1)}, \dots, X_{(m)}] \\ &= [X^{(1)}, \dots, X^{(n)}].\end{aligned}$$

Throughout the reminder of the paper we shall assume the following about the structure of $\mathbf{X}$:

(A1) $X_{(1)}, \dots, X_{(m)}$ are obtained from the first n elements of exchangeable sequences $(X_{1,j})_{j\geq 1}, \dots, (X_{m,j})_{j\geq 1}$, equidistributed with the probability law $\mathcal{L}_r$ and

(A2) $X^{(1)}, \dots, X^{(n)}$ are obtained from the first m elements of independent sequences $(X_{i,1})_{i\geq 1}, \dots, (X_{i,n})_{i\geq 1}$, equidistributed with the probability law $\mathcal{L}_c$,

where $\mathcal{L}_c$ and $\mathcal{L}_r$ are some (possibly different) probability laws in the space of infinite sequences.

DEFINITION 2. For a given kernel function h *the generalized permanent function* $Per_h\,\mathbf{X}$ is defined as

$$Per_h\,\mathbf{X} = \sum_{1\leq j_1 \neq \dots \neq j_m \leq n} h(X_{1,j_1}, X_{2,j_2}, \dots, X_{m,j_m}).$$

[1]That, is there exist no sequences of real numbers a_n and b_n such that $a_n(Per\,\underline{X} - b_n)$ converges in law to a non-degenatate random variable.

(If $h = \prod_{i=1}^m y_i$ then $Per_h \mathbf{X} = Per\mathbf{X}$). A corresponding *P-statistic* is then defined as $\binom{n}{m}^{-1} Per_h \mathbf{X}/m!$.

Note that in the case of the one dimensional projection matrices the above definition reduces to that of a U-statistic from Definition 1, in view of the relation similar to (5) extended now to an arbitrary kernel h.

2.1. Martingale decomposition of a P-statistic

In the sequel let us denote by $\mathbf{X}(i_1, \dots, i_p|j_1, \dots, j_q)$ a sub-matrix of $\mathbf{X}$ consisting of the entries at the intersections of the rows $i_1, \dots, i_p$ and the columns $j_1, \dots, j_q$.

Let us note that since $Per_h \mathbf{X}$ is a symmetric real function with respect to both the rows and the columns of $\mathbf{X}$ we may define the row and column versions of the symmetrization operator (1)

$$\pi_p^m Per_h \mathbf{X}(p|\cdot) \stackrel{def}{=} \sum_{1\le i_1<\ldots<i_p\le m} Per_h \mathbf{X}(i_1, \dots, i_p|\cdot),$$

$$\pi_q^n Per_h \mathbf{X}(\cdot|q) \stackrel{def}{=} \sum_{1\le j_1<\ldots<j_q\le n} Per_h \mathbf{X}(\cdot|j_1, \dots, j_q).$$

By the Laplace expansion formula for permanents (cf Minc 1978 chapter 2)

$$Per_h \mathbf{X} = \pi_m^n Per_h \mathbf{X}(1, 2, \dots, m|m). \tag{6}$$

Thus, in order to decompose $Per_h \mathbf{X}$, we will first decompose $Per_h \mathbf{X}(1, 2, \dots, m|j_1, j_2, \dots, j_m)$ Assuming for convenience that $Eh = 0$ (or taking $\tilde{h} = h - Eh$ in the sequel), let us consider $\sigma = (\sigma(j_1), \dots, \sigma(j_m))$, some fixed permutation of the set of m column numbers $\{j_1, \dots, j_m\}$.

For given σ, denoting $h(X_{1,\sigma(j_1)}, \dots, X_{m,\sigma(j_m)})$ by, say, $h(W_1, \dots, W_m|\sigma)$ we have by Hoeffding's theorem

$$h(X_{1,\sigma(j_1)}, \dots, X_{m,\sigma(j_m)}) = h(W_1, \dots, W_m|\sigma) = \sum_{c=1}^{m} \pi_c^m (g_c|\sigma)$$

where

$$g_c(w_1, \dots, w_c|\sigma) = \int_{\mathbf{R}} \cdots \int_{\mathbf{R}} h(z_1, \dots, z_m|\sigma) \prod_{s=1}^{c} (\delta_{w_s}(dz_s) - P(dz_s)) \prod_{s=c+1}^{m} P(dz_s).$$

In view of the above, the decomposition of $Per_h \mathbf{X}(1, 2, \dots, m| j_1, j_2, \dots, j_m)$ can be obtained as follows

$$\begin{aligned}Per_h\,\mathbf{X}(1,2,\ldots,m|j_1\ldots,j_m) &= \sum_\sigma h(X_{1,\sigma(j_1)},X_{2,\sigma(j_2)},\ldots,X_{m,\sigma(j_m)})\\ &= \sum_\sigma\sum_{c=1}^m \pi_c^m(g_c|\sigma) = \sum_{c=1}^m \pi_c^m\Big(\sum_\sigma (g_c|\sigma)\Big)\\ &= \sum_{c=1}^m \pi_c^m Per_{g_c}\,\mathbf{X}(c|j_1\ldots,j_m).\end{aligned}$$

Thus,

$$\begin{aligned}Per_h\,\mathbf{X} &= \pi_m^n Per_h\,\mathbf{X}(1,2,\ldots,m|m)\\ &= \sum_{1\le j_1<\ldots<j_m\le n} Per_h\,\mathbf{X}(1,2,\ldots,m|j_1,\ldots,j_m)\\ &= \sum_{1\le j_1<\ldots<j_m\le n}\sum_{c=1}^m \pi_c^m Per_{g_c}\,\mathbf{X}(c|j_1,\ldots,j_m)\\ &= \sum_{1\le j_1<\ldots<j_m\le n}\sum_{1\le i_1<\ldots<i_c\le m}\sum_{c=1}^m Per_{g_c}\,\mathbf{X}(i_1,\ldots,i_c|j_1,\ldots,j_m)\\ &= \sum_{c=1}^m\binom{n-c}{m-c}(m-c)!\sum_{1\le j_1<\ldots<j_m\le n}\sum_{1\le i_1<\ldots<i_c\le m} Per_{g_c}\,\mathbf{X}(i_1,\ldots,i_c|j_1,\ldots,j_c)\\ &\stackrel{def}{=} \sum_{c=1}^m\binom{n-c}{m-c}(m-c)!\,\pi_c^n\pi_c^m Per_{g_c}\,\mathbf{X}(c|c),\end{aligned}$$

where the last expression is a definition. Re-writing the above as

$$Per_h\,\mathbf{X} = \sum_{c=1}^m m!\binom{n}{m}\binom{n}{c}^{-1}\pi_c^n\pi_c^m Per_{g_c}\,\mathbf{X}(c|c)/c!$$

and denoting

$$\begin{aligned}U_{g_c}^{(m,n)} &\stackrel{def}{=} \binom{m}{c}^{-1}\binom{n}{c}^{-1}\pi_c^m\pi_c^n Per_{g_c}\,\mathbf{X}(c|c)/c!\\ &= \binom{m}{c}^{-1}\binom{n}{c}^{-1}\pi_c^n\pi_c^m Per_{g_c}\,\mathbf{X}(c|c)/c!\end{aligned}\tag{7}$$

we obtain

THEOREM 1. *Under our assumptions on* $\mathbf{X}$, *suppose that a kernel* h *satisfies* $E|h|<\infty$. *Then*

$$\binom{n}{m}^{-1}(Per_h\,\mathbf{X} - EPer_h\,\mathbf{X})/m! = \sum_{c=1}^m\binom{m}{c}U_{g_c}^{(m,n)}. \qquad \blacksquare$$

REMARK. If $\mathbf{X}$ is a one dimensional projection matrix then the P-statistic is simply a U-statistic and the above reduces to the H-representation theorem.

Let us also note that for any fixed $c \geq 1$ the component $U_{g_c}^{(m,n)}$ is simply a symmetrization (with respect to the rows and the columns of $\mathbf{X}$) of a P-statistic with a kernel function g_c based on a $c \times c$ submatrix of $\mathbf{X}$.

Our next result is the P-statistic version of the result for classical permanent function obtained in Rempała and Wesołowski (2000b).

PROPOSITION 2. *Assume $m = m_n$ is a non-decreasing sequence in n. Under our assumptions on the entries of the matrix $\mathbf{X}$, for a fixed $c \geq 1$*

(i) if we define $\mathcal{F}_c^{(n)} = \sigma\{U_{g_c}^{(m_n,n)}, U_{g_c}^{m_{n+1},n+1}, \ldots\}$, then

$$E(U_{g_c}^{(m_n,n)} | \mathcal{F}_c^{(n+1)}) = U_{g_c}^{m_{n+1},n+1}, \quad \forall\, n = n_0, n_0+1, \ldots$$

that is, $(U_{g_c}^{(m,n)}, \mathcal{F}_c^{(n)})_{n=n_0,n_0+1,\ldots}$ is a backward martingale.

(ii) Additionally, we have also

$$Cov(U_{g_{c_1}}^{(m,n)}, U_{g_{c_2}}^{(m,n)}) = 0 \qquad \text{for } c_1 \neq c_2.$$

Proof. *(i)* Let us denote an element of the Hoeffding-like decomposition (7) of an $m_n \times n$ matrix obtained from the $m_{n+1} \times (n+1)$ matrix by deleting the $m_{n+1} - m_n$ rows, say, $l_1, \ldots, l_{m_{n+1}-m_n}$ and deleting the k-th column by $U_{g_c}^{(m_n,n)}(l_1, \ldots, l_{m_{n+1}-m_n}; k)$, i.e.,

$$U_{g_c}^{(m_n,n)}(l_1, \ldots, l_{m_{n+1}-m_n}; k) \stackrel{def}{=} \binom{m_n}{c}^{-1} \binom{n}{c}^{-1} (\mu^c c!)^{-1} \times$$
$$\times \sum_{\substack{1 \leq i_1 < \ldots < i_c \leq m_{n+1} \\ \{i_1,\ldots,i_c\} \cap \{l_1,\ldots,l_{m_{n+1}-m_n}\} = \emptyset}} \; \sum_{\substack{1 \leq j_1 < \ldots j_c \leq n+1 \\ k \notin \{j_1,\ldots,j_c\}}} Per_{g_c} \mathbf{X}(i_1, \ldots, i_c | j_1, \ldots, j_c).$$

Observe that

$$\sum_{k=1}^{n+1} \sum_{1 \leq l_1 < \ldots < l_{m_{n+1}-m_n} \leq m_{n+1}} U_{g_c}^{(m_n,n)}(l_1, \ldots, l_{m_{n+1}-m_n}; k) =$$
$$= \binom{m_{n+1}-c}{m_n - c} \frac{n+1-c}{\binom{m_n}{c}\binom{n}{c}} \pi_c^{n+1} \pi_c^{m+1} Per_{g_c} \mathbf{X}(c|c)/c!.$$

Since $\binom{m_{n+1}-c}{m_{n+1}-m_n} \binom{m_n}{c}^{-1} = \binom{m_{n+1}}{m_n} \binom{m_{n+1}}{c}^{-1}$, the above entails

$$\sum_{k=1}^{n+1} \sum_{1 \leq l_1 < \ldots < l_{m_{n+1}-m_n} \leq m_{n+1}} U_{g_c}^{(m_n,n)}(l_1, \ldots, l_{m_{n+1}-m_n}; k) \tag{8}$$
$$= (n+1) \binom{m_{n+1}}{m_n} U_{g_c}^{(m_{n+1},n+1)}.$$

Let us note that in view of our assumptions (A1)-(A2) it follows that the conditional distribution of $U_{g_c}^{(m_n,n)}(l_1,\dots,l_{m_{n+1}-m_n};k)$ given $\mathcal{F}_c^{(n+1)}$ is the same for any particular choice of $k\in\{1,\dots,n+1\}$ and $\{l_1,\dots,l_{m_{n+1}-m_n}\}\subset\{1,\dots,m_{n+1}\}$. Consequently,

$$E[U_{g_c}^{(m_n,n)}(l_1,\dots,l_{m_{n+1}-m_n};k)|\mathcal{F}_c^{(n+1)}]=$$
$$=E[U_{g_c}^{(m_n,n)}(n+1,n+2,\dots,m_{n+1};n+1)|\mathcal{F}_c^{(n+1)}]=E[U_{g_c}^{(m,n)}|\mathcal{F}_c^{(n+1)}],$$

which entails

$$(9)\quad E\Big[\sum_{k=1}^{n+1}\sum_{1\le l_1<\dots<l_{m_{n+1}-m_n}\le m_{n+1}}U_{g_c}^{(m_n,n)}(l_1,\dots,l_{m_{n+1}-m_n};k)|\mathcal{F}_c^{(n+1)}\Big]$$

$$(10)\quad =(n+1)\binom{m_{n+1}}{m_n}E(U_{g_c}^{(m_n,n)}|\mathcal{F}_c^{(n+1)}).$$

But on the other hand, in view of the identity (8) we have that (9) is equal to

$$(11)\quad (n+1)\binom{m_{n+1}}{m_n}E(U_{g_c}^{(m_{n+1},n+1)}|\mathcal{F}_c^{(n+1)})=(n+1)\binom{m_{n+1}}{m_n}U_{g_c}^{(m_{n+1},n+1)}$$

Comparing the expression (10) with that at the right-hand side of (11), we arrive at *(i)*.

In order to show *(ii)* it is enough to show that for $c_1\neq c_2$ we have

$$Cov\,[\pi_{c_1}^m\pi_{c_1}^n Per_{g_c}\mathbf{X}(c_1|c_1),\pi_{c_2}^m\pi_{c_2}^n Per_{g_c}\mathbf{X}(c_2|c_2)]=0$$

which will follow if we can show that for any pairs of fixed sets of rows $\{i_1,\dots,i_{c_1}\},\{k_1,\dots,k_{c_2}\}$ and columns $\{j_1,\dots,j_{c_1}\},\{l_1,\dots,l_{c_2}\}$ we have

$$(12)\quad Cov\,[Per_{g_c}\mathbf{X}(i_1,\dots,i_{c_1}|j_1,\dots,j_{c_1}),Per_{g_c}\mathbf{X}(k_1,\dots,k_{c_2}|l_1,\dots,l_{c_2})]=0.$$

Consider an arbitrary pair of fixed sets of rows $\{i_1,\dots,i_{c_1}\},\{k_1,\dots,k_{c_2}\}$ and columns $\{j_1,\dots,j_{c_1}\},\{l_1,\dots,l_{c_2}\}$. Let us note that by the linearity

$$Cov\,(Per_{g_c}\mathbf{X}(i_1,\dots,i_{c_1}|j_1,\dots,j_{c_1}),Per_{g_c}\mathbf{X}(k_1,\dots,k_{c_2}|l_1,\dots,l_{c_2}))$$
$$=\sum_{\sigma_{c_1},\,\sigma_{c_2}}E\left(g_c(X_{i_1,\sigma(j_1)},\dots,X_{i_{c_1},\sigma(j_{c_1})})\,g_c(X_{k_1,\sigma(l_1)},\dots,X_{k_{c_2},\sigma(l_{c_2})})\right)$$

where the summation is taken over all respective permutations σ_{c_1} of $\{j_1,\dots,j_{c_1}\}$ and σ_{c_2} of $\{l_1,\dots,l_{c_2}\}$. Since $c_1\neq c_2$ we may assume without loosing generality that $c_1>c_2$. It follows that there exist at least one column $j_s\in\{j_1,\dots,j_{c_1}\}$ such that $j_s\notin\{l_1,\dots,l_{c_2}\}$. But by the assumption of the independence of columns and the property of canonical functions (3) the standard conditioning argument implies now that all the summands above are zero and thus (12) follows. The proof of the proposition is complete. ■

Let us provide a simple example of the application of Theorem 1.

EXAMPLE *(Random permanent decomposition).* Let $EX_{ij} = \mu \neq 0$ and consider the kernel $h(y_1, \ldots, y_m) = \prod_{i=1}^m y_i$. Then

$$g_c(y_1, \ldots, y_c) = \prod_{i=1}^{c} (y_i - \mu)\, \mu^{m-c}$$

and by Theorem 1 we have

$$\frac{Per\, \mathbf{X}}{\binom{n}{m}\, m! \mu^m} = 1 + \sum_{c=1}^{m} \binom{m}{c} U_c^{(m,n)},$$

where

$$U_c^{(m,n)} = \binom{n}{c}^{-1} \binom{m}{c}^{-1} (\mu^c c!)^{-1} \sum_{1 \le i_1 < \ldots < i_c \le m} \; \sum_{1 \le j_1 < \ldots < j_c \le n} Per\, [\tilde{X}_{i_u j_v}]_{\substack{u = 1, \ldots, c \\ v = 1, \ldots, c}},$$

for $\tilde{X}_{ij} = (X_{ij} - \mu)/\mu$, $(i = 1, \ldots, m,\ j = 1, \ldots, n)$.

3. Approximation of a P-statistic

3.1. The variance formula

Using the representation of Theorem 1 we shall prove here a general approximation theorem for P-statistics. In order to do so we shall need the expression for the variance of a P-statistic given by the following

THEOREM 2. *(The variance formula for P-statistics) Under our assumptions on* $\mathbf{X}$ *let us suppose* $Eh^2 < \infty$ *then for* $c = 1, \ldots, m$

$$Var\, U_c^{(m,n)} = \binom{n}{c}^{-1} \binom{m}{c}^{-1} \sum_{r=0}^{c} \binom{m-r}{c-r} \frac{D(c,r)}{r!}$$

where

$$D(c,r) \stackrel{def}{=} \sum_{l=0}^{r} \binom{r}{l} (-1)^{r-l} \rho_{c,l}$$

for $r = 0, 1, \ldots, c$ *and*

$$\rho_{c,l} \stackrel{def}{=} E\Big[g_c(X_{11}, \ldots, X_{ll}, X_{i_{l+1}, l+1} \ldots, X_{i_c, c}) \times g_c(X_{11}, \ldots, X_{ll}, X_{j_{l+1}, l+1} \ldots, X_{j_c, c})\Big]$$

where $i_k \neq j_k$ *for* $k = 1, \ldots, c - l + 1$ *and* $l = 0, \ldots, c$.

Let us note that for the classical permanent function the above result reduces to that obtained in Rempała and Wesołowski (2000a). Let us also note that it implies, in particular, that if the rows of $\mathbf{X}$ are also independent, then

$$D(c,r) = 0 \quad \text{unless} \quad r = c$$

and, since $\rho(c,c) = Var g_c$, that

$$Var\left(U_{g_c}^{(m,n)}\right) = \binom{m}{c}^{-1}\binom{n}{c}^{-1}\frac{Var g_c}{c!},$$

which entails (by orthogonality)

$$Var\left(\frac{Per_h \mathbf{X}}{\binom{n}{m}\, m!}\right) = \sum_{c=1}^{m}\frac{\binom{m}{c}}{\binom{n}{c}}\frac{Var g_c}{c!}.$$

Again, in a special case of random permanent function, this particular formula has been obtained by Rempała and Wesołowski (1999) by means of quite involved combinatorial calculations.

On the other hand if $\mathbf{X} = \underline{X}$ (one dimensional projection matrix) then we have

$$D(c,r) = 0 \quad \text{unless} \quad r = 0$$

and we obtain the standard variance formula for U-statistics (cf. e.g., Lee 1990)

$$Var\left(\frac{Per_h \underline{X}}{\binom{n}{m}}\right) = \sum_{c=1}^{m}\frac{\binom{m}{c}^2}{\binom{n}{c}}\, Var\, g_c.$$

Let us now prove Theorem 2.

P r o o f. First, let us note that by Proposition 2 it follows that

$$(13) \quad Var\,\frac{Per_h \mathbf{X}}{\binom{n}{m}\, m!}$$
$$= \sum_{c=1}^{m}\binom{m}{c}^2 Var\, U_{g_c}^{(m,n)} + \sum_{1\le c_1\ne c_2\le m}\binom{m}{c_1}\binom{m}{c_2} Cov\left(U_{g_{c_1}}^{(m,n)}, U_{g_{c_2}}^{(m,n)}\right)$$
$$= \sum_{c=1}^{m}\binom{m}{c}^2 Var\, U_{g_c}^{(m,n)}.$$

Now, for $c = 1, \dots m$

$$\binom{n}{c}^2\binom{m}{c}^2 c!^2\, Var\, U_{g_c}^{(m,n)} = Var\ \pi_c^n \pi_c^m Per_{g_c}\mathbf{X}(c|c)$$
$$= \sum_{1\le j_1<\dots<j_c\le n} Var\Big(\sum_{1\le i_1<\dots<i_c\le m} Per_{\,g_c}[X_{i_u j_v}]_{\substack{u=1,\dots,c\\ v=1,\dots,c}}\Big)$$

since, by independence of columns of $\mathbf{X}$, we have (cf. also the proof of part *(ii)* of Proposition 2)

$$Cov\Big(\sum_{1\le i_1<\ldots<i_c\le m} Per_{\,g_c}[X_{i_u j_v}]_{\substack{u=1,\ldots,c\\ v=1,\ldots,c}}, \sum_{1\le i_1<\ldots<i_c\le m} Per_{\,g_c}[X_{i_u l_v}]_{\substack{u=1,\ldots,c\\ v=1,\ldots,c}}\Big)=0$$

if only $\{j_1,\ldots,j_c\}\neq\{l_1,\ldots,l_c\}$.

Since the columns of $\mathbf{X}$ are identically distributed, the latest expression simplifies to

$$\binom{n}{c}\binom{m}{c}^2 c!^2\, Var\, U_{g_c}^{(m,n)} = Var\Big(\sum_{1\le i_1<\ldots<i_c\le m} Per_{\,g_c}[X_{i_u j}]_{\substack{u=1,\ldots,c\\ j=1,\ldots,c}}\Big).$$

Let us note that the number of pairs of $c\times c$ submatrices of $\mathbf{X}$ having exactly k rows in common equals $\binom{m}{c}\binom{c}{k}\binom{m-c}{c-k}$, for $\max(0,2c-m)\le k\le c$, and each such pair has equal covariance (since the row vectors of $\mathbf{X}$ are identically distributed). Hence, for given c, the above right-hand side can be written as

$$\binom{m}{c}\sum_{k=\max(0,2c-m)}^{c}\binom{c}{k}\binom{m-c}{c-k}\times$$

$$\times\, Cov\left(Per_{\,g_c}[X_{ij}]_{\substack{i=1,\ldots,k,i_{k+1},\ldots,i_c\\ j=1,\ldots,c}}, Per_{\,g_c}[X_{ij}]_{\substack{i=1,\ldots,k,l_{k+1},\ldots,l_c\\ j=1,\ldots,c}}\right)$$

where $\{i_{k+1},\ldots,i_c\}\cap\{l_{k+1},\ldots,l_c\}=\emptyset$.

Observe that each term of the above sum is itself a sum of the partial covariance elements of the form

$$E\left[g_c(X_{i_1,1},\ldots,X_{i_l,l},X_{i_{l+1},l+1}\ldots,X_{i_c,c})\, g_c(X_{i_1,1},\ldots,X_{i_l,l},X_{j_{l+1},l+1}\ldots,X_{j_c,c})\right]$$

where $i_{l+1},\ldots,i_c$ and $j_{l+1},\ldots,j_c$ are fixed non-overlapping subsets of $\{l+1,\ldots,m\}$ for some $0\le l\le k$. By the assumptions about the entries of the matrix $\mathbf{X}$ it follows that the partial covariances having exactly l $(0\le l\le k)$ elements in common are the same and equal to

$$\rho_{c,l} =$$
$$E\left[g_c(X_{11},\ldots,X_{ll},X_{i_{l+1},l+1}\ldots,X_{i_c,c})\, g_c(X_{11},\ldots,X_{ll},X_{j_{l+1},l+1}\ldots,X_{j_c,c})\right].$$

Now, to compute the covariance of such $k\times c$ permanents it suffices to find the number of pairs of c-tuples of arguments of the function g_c with exactly l elements in common, $(0\le l\le k\le c)$. Observe that it equals to the number of pairs of c-tuples having exactly l common elements in a permanent of the matrix $k\times c$, multiplied by $(c-k)!^2$ – the number of all possible permutations of $i_{k+1},\ldots,i_c$ and $l_{k+1},\ldots,l_c$.

To compute the number of pairs of c-tuples with exactly l elements in common let us start with finding the number of c-tuples present in the defining formula for $Per_{g_c}\mathbf{Y}[k,c]$, where $\mathbf{Y}[k,c]$ is a $k \times c$ matrix, having exactly l elements in common with the diagonal entries $y_{11}\dots y_{kk}$. First, we fix l factors in $\binom{k}{l}$ ways. If we assume that $y_{11},\dots,y_{ll}$ are fixed, then the remaining factors, in the c-tuples we are looking for, have to be of the form $y_{l+1,j_{l+1}},\dots,y_{k,j_k}$, where $j_r \neq r$, $r = l+1,\dots,k$. Finding the number of such c-tuples (say, $\mathcal{R}_l(k,c)$) is equivalent to computing the number of summands in a permanent of the matrix of dimensions $(k-l)\times(c-l)$ which do not contain any diagonal entry. To this end, we subtract the number of all summands having at least one factor being the diagonal entry, from the total number of all summands in that permanent. Using the exclusion-inclusion formula we get that

$$\mathcal{R}_l(k,c) = \binom{c-l}{k-l}(k-l)! - \sum_{j=1}^{k-l}(-1)^{j+1}\binom{k-l}{j}\binom{c-l-j}{k-l-j}(k-l-j)!$$

where the absolute value of the j-th member of the above sum denotes the number of c-tuples having exactly j factors being the diagonal entries (equal to the number of choices of j positions on the diagonal) multiplied by the number of c-tuples of $k-l-j$ factors from the outside of the diagonal (equal to number of c-tuples in the permanent of the matrix of dimensions $(k-l-j)\times(c-l-j)$). Thus, in a slightly more compact form,

$$\mathcal{R}_l(k,c) = \sum_{j=0}^{k-l}(-1)^j\binom{k-l}{j}\binom{c-l-j}{k-l-j}(k-l-j)!. \tag{14}$$

Consequently, the number of pairs of c-tuples in $Per\,\mathbf{Y}[k,c]$ with exactly l factors in common equals to

$$\binom{c}{k}k!\binom{k}{l}\mathcal{R}_l(k,c).$$

Hence, combining the above formula with an earlier one for the number of pairs of c-tuples with l identical factors we arrive at

$$Cov\Big(Per_{g_c}[X_{ij}]_{\substack{i=1,\dots,k,i_{k+1},\dots,i_c\\ j=1,\dots,c}}, Per_{g_c}[X_{ij}]_{\substack{i=1,\dots,k,l_{k+1},\dots,l_c\\ j=1,\dots,c}}\Big) =$$

$$= (c-k)!^2\sum_{l=0}^{k}\binom{c}{k}k!\binom{k}{l}\rho_{c,l}\mathcal{R}_l(k,c).$$

Now, returning to the formula for the variance of $U_{g_c}^{(m,n)}$ we obtain by (14)

$$(15)\quad \binom{n}{c}\binom{m}{c}\, Var\, U_{g_c}^{(m,n)}$$

$$= \frac{1}{c!^2} \sum_{k=\max(0,2c-m)}^{c} \binom{c}{k}^2 \binom{m-c}{c-k} k!(c-k)!^2 \sum_{l=0}^{k} \binom{k}{l} \rho_{c,l} \mathcal{R}_l(k,c)$$

$$= \sum_{k=\max(0,2c-m)}^{c} \binom{m-c}{c-k} \sum_{r=0}^{k} \binom{c-r}{k-r} \frac{1}{r!} D(c,r),$$

since

$$\sum_{l=0}^{k} \binom{k}{l} \rho_{c,l}\mathcal{R}_l(k,c) = \sum_{l=0}^{k} \binom{k}{l} \rho_{c,l} \sum_{r=l}^{k} (-1)^{r-l} \binom{k-l}{r-l} \binom{c-r}{k-r} (k-r)!$$

$$= \sum_{r=0}^{k} \binom{c-r}{k-r} (k-r)! \binom{k}{r} \sum_{l=0}^{r} \binom{r}{l} (-1)^{r-l} \rho_{c,l}$$

$$= \sum_{r=0}^{k} \binom{c-r}{k-r} \frac{k!}{r!} D(c,r).$$

Observe that we can further simplify the expression (15), since

$$\sum_{k=\max(0,2c-m)}^{c} \binom{m-c}{c-k} \sum_{r=0}^{k} \binom{c-r}{k-r} \frac{D(c,r)}{r!}$$

$$= \sum_{r=0}^{c} \frac{D(c,r)}{r!} \sum_{k=\max(r,2c-m)}^{c} \binom{m-c}{c-k} \binom{c-r}{k-r} = \sum_{r=0}^{c} \binom{m-r}{c-r} \frac{D(c,r)}{r!},$$

where the last equality follows by applying the hypergeometric summation rule for the inner sum. Thus, we can rewrite (15) as

$$Var\, U_{g_c}^{(m,n)} = \binom{n}{c}^{-1} \binom{m}{c}^{-1} \sum_{r=0}^{c} \binom{m-r}{c-r} \frac{D(c,r)}{r!},$$

which along with (13) completes the proof. ∎

3.2. Approximation theorem

With the just derived variance formulas for the components of the orthogonal decomposition, we are now finally in a position to state as our last result the following simple approximation theorem for P-statistics.

THEOREM 3. *(Approximation Theorem for P-statistics). Suppose* $m/\sqrt{n} \to 0$ *as* $n \to \infty$, $Eh^2 < \infty$ *and* $g_1 = h_1 - Eh \neq 0$ *and* $\exists M$ *st. for* $c = 1, 2, 3, \ldots$

$$\forall_{r \le c}\; D(c,r) \le M^c < \infty.$$

Then

$$\frac{Per_h \mathbf{X}}{\binom{n}{m} m!} - Eh = \frac{1}{n}\sum_{ij} g_1(X_{ij}) + o_p\left(\frac{m}{\sqrt{n}}\right).$$

REMARK. In particular, if $h(y_1, \ldots, y_m) = \prod_{i=1}^m y_i$ and $Eh = \mu \neq 0$ then $[g_1(X_{ij})] = \mu^{m-1}[X_{ij} - \mu]$ and thus the above formula provides the approximation for the classical permanent function since then we may take $M = Var\,(X_{11})$.

Proof. The result is a simple consequence of Theorems 1 and 2 since for $c = 1$

$$\begin{aligned} U_{g_1}^{(m,n)} &\stackrel{def}{=} \binom{m}{1}^{-1}\binom{n}{1}^{-1} \pi_1^m \pi_1^n Per_{g_1}\mathbf{X}(1|1) \\ &= \frac{1}{mn}\sum_{ij} g_1(x_{ij}) \end{aligned}$$

and thus by Theorem 1

$$\frac{Per_h \mathbf{X}}{\binom{n}{m} m!} - Eh = \frac{1}{n}\sum_{ij} g_1(x_{ij}) + R_{m,n}.$$

To complete the proof it is enough then to show $R_{m,n} = o_p(m/\sqrt{n})$ which will follow if we can argue that

$$\frac{n}{m^2} Var\, R_{m,n} \to 0$$

as $n \to \infty$ and $m^2/n \to 0$. Since under the assumptions of the theorem we have

$$Var\, U_{g_c}^{(m,n)} = \binom{n}{c}^{-1}\binom{m}{c}^{-1}\sum_{r=0}^{c}\binom{m-r}{c-r}\frac{D(c,r)}{r!} \leq \binom{n}{c}^{-1}\exp(M)$$

as $\binom{m-r}{c-r} \leq \binom{m}{c}$ for $(1 \leq r \leq c \leq m)$ and by Theorem 2

$$\begin{aligned} \frac{n}{m^2} Var\, R_{m,n} &= \frac{n}{m^2}\sum_{c=2}^{m}\binom{m}{c}^2 Var\, U_{g_c}^{(m,n)} \\ &\leq \frac{n}{m^2}\exp(M)\sum_{c=2}^{m}\binom{m}{c}^2\binom{n}{c}^{-1} \\ &\leq \frac{n}{m^2}\exp(M)\sum_{c=2}^{m}\left(\frac{m^2}{n}\right)^c\frac{1}{c!}, \end{aligned}$$

in view of

$$c!\frac{\binom{m}{c}^2}{\binom{n}{c}} \leq \left(\frac{m^2}{n}\right)^c.$$

Consequently,

$$\frac{n}{m^2} Var\, R_{m,n} \leq \exp(M) \frac{m^2}{n} \sum_{c=2}^{m} \left(\frac{m^2}{n} \right)^{c-2} \frac{1}{c!} \to 0,$$

since the sum above is bounded by a constant and $m^2/n \to 0$. ∎

References

Binet, J. P. M. (1812). *Memoire sur un systeme de formules analytiques, et leur application à des considérations géometriques.[Memoirs on the System of Analytic Formulae and Their Application to Geometric Considerations.].* L' École Politechnique, Paris, France.

Borovskikh, Y. V. and Korolyuk, V. S. (1994). *Random permanents and symmetric statistics*, Acta Appl. Math. 36(3), 227–288.

Lee, A. J. (1990). *U-statistics.Theory and Practice*, volume 110 of *Statistics: Textbooks and Monographs.* Marcel Dekker Inc., New York.

Minc, H. (1978). *Permanents.* Addison-Wesley Publishing Co., Reading, Mass. With a foreword by Marvin Marcus, Encyclopedia of Mathematics and its Applications, Vol. 6.

Rempała, G. and Wesołowski, J. (2000a). *Central limit theorems for random permanents with correlation structure.* Manuscript, to appear in Journal of Theoretical Probability.

Rempała, G. and Wesołowski, J. (2000b). *Strong laws of large numbers for random permanents.* Manuscript.

Rempała, G. and Wesołowski, J. (1999). *Limiting behavior of random permanents*, Statist. Probab. Lett. 45(2), 149–158.

Ryser, H. J. (1963). *Combinatorial Mathematics.* Published by The Mathematical Association of America. The Carus Mathematical Monographs, No. 14.

Valiant, L. (1979). *The complexity of computing the permanent*, Theoret. Comput. Sc. 8(1), 189–201.

van Es, A. J. and Helmers, R. (1988). *Elementary symmetric polynomials of increasing order*, Probab. Theory Related Fields 80(1), 21–35.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF LOUISVILLE
Natural Sciences Bldg.
LOUISVILLE, KY 40292, USA

Received July 17, 2000; revised version January 24, 2001.