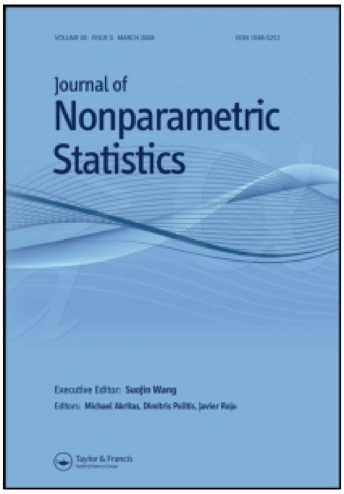

Journal of Nonparametric Statistics
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713645758

INCOMPLETE U -STATISTICS OF PERMANENT DESIGN

Grzegorz Rempala [a]; Jacek Wesolowski [a]
[a] Department of Mathematics, University of Louisville, Louisville, KY, USA.

Online Publication Date: 01 April 2003

To cite this Article Rempala, Grzegorz and Wesolowski, Jacek(2003)'INCOMPLETE U -STATISTICS OF PERMANENT DESIGN',Journal of Nonparametric Statistics,15:2,221 — 236

To link to this Article: DOI: 10.1080/1048525031000089284

URL: http://dx.doi.org/10.1080/1048525031000089284

Nonparametric Statistics, 2003, Vol. 15(2), pp. 221–236

INCOMPLETE *U*-STATISTICS OF PERMANENT DESIGN

GRZEGORZ REMPALA[a,]* and JACEK WESOLOWSKI[b]

[a]*Department of Mathematics, University of Louisville, Louisville, KY, USA;*
[b]*Department of Mathematics, Warsaw Technical University, Warsaw, Poland*

(Received 27 June 2000; In final form 8 September 2000)

A class of incomplete *U*-statistics based on what we call a "permanent design" is investigated. It is shown that the design is balanced and has minimal variance in a certain class of designs as well as is asymptotically efficient, *i.e.*, its asymptotic variance equals that of a complete *U*-statistic. Several examples of applications of our results to obtaining limiting distributions of certain statistics as well as to sub-sampling theory are discussed.

Keywords: Incomplete *U*-statistics; Random permanents; Minimum variance designs; Relative efficiency

1 INTRODUCTION

In his seminal paper Hoeffding (1948) has extended the notion of forming an average of l i.i.d. random variables by averaging a symmetric measurable function (kernel) of $k > 1$ arguments over $\binom{l}{k}$ possible subsets. Since then these generalized averages or "*U*-statistics" have become one of the most intensely studied objects in non-parametric statistics, (see, for instance, Lee, 1990 and references therein) with some aspects of the theory, like *e.g.*, the asymptotic behavior, being a very active area of research even until now (see, for instance, Giné and Zinn, 1994; Rempala, 1998; Latala and Zinn, 2000 or Giné *et al.*, 2001, for some of the most recent results). Even though *U*-statistics are relatively simple and thus theoretically appealing probabilistic objects their practical use is somewhat limited due to the fact that for large values of k and l the number of needed averagings $\binom{l}{k}$ may be very large and the actual calculation of a *U*-statistic can be quite onerous. This drawback is especially apparent when considering so-called *U*-statistics of infinite order (USIO) where the dimension k of a kernel function is allowed to grow with the sample size l (cf. *e.g.*, Székely, 1982; Frees, 1989; Politis and Romano, 1994; Rempala, 1998). In order to address the problem of the computational difficulties Blom (1976) has suggested considering "incomplete" *U*-statistics, where the kernel function is averaged over only some

* Corresponding author.

ISSN 1048-5252 print; ISSN 1029-0311 online
DOI: 10.1080/1048525031000089284

appropriately chosen small subset (a design) of all $\binom{l}{k}$ averagings. The idea of Blom has turned out to be closely related to the general statistical theory of experimental designs and has lead over the next decade to a rapid development of the theory of incomplete U-statistics based on a variety of designs (cf. *e.g.*, Brown and Kildea, 1978; Lee, 1982; Enqvist, 1985; Herrndorf, 1986; Nowicki and Wierman, 1988), including random (with or without replacement) selection of subsets (Janson, 1984).

The main drawback of the theory of incomplete U-statistics seemed to be so far in a lack of a general and simple technique of generating incomplete designs (except perhaps for the random selection) that would work well for a very wide class of U-statistics, including the case when $k \to \infty$.

In this paper we consider such a design for a rectangle of mn i.i.d. random variables based on a definition of a permanent function for $m \times n$ matrix (cf. Minc, 1978). This "permanent" design seems to be quite natural (for instance, it is "equireplicate" in the experimental design sense) and, in particular, under some regularity conditions imposed on the kernel functions of the corresponding U-statistics is always asymptotically efficient, in the usual sense of ARE, when compared with a full design of all $\binom{l}{k}$ averagings, even in the case when both $l, k \to \infty$. In some cases our permanent design can also be shown to be more effective then any other design of the same size, including a random one. The main weakness of our method is in the fact that our permanent designs obtained here are typically still fairly large in terms of the number of subsets, but on the other hand, unlike most of the designs considered so far in the literature can be easily implemented on a computer with a help of a permanent version of the Laplace expansion formula (cf. Mine, 1978).

The paper is organized as follows. The reminder of this section explains our notation used throughout the paper and presents a brief overview of the basic formulae for complete and incomplete U-statistics. In Section 2 we introduce the concept of a "permanent design" and discuss some properties of incomplete U-statistics of permanent design (USPD) as well as provide the variance formula for USPD. Section 3 contains further results on USPD, based on asymptotic considerations. Some examples of applications of the USPD theory, including a theorem on the limiting law for random permanents and a problem of non-parametric interval estimation based on sub-sampling are presented in Section 4.

Throughout the paper, we shall assume that $X_1, \ldots, X_l$ are independent, identically distributed (i.i.d.) random variables taking values in a measurable space $(\mathcal{I}, \mathcal{J})$ with the common probability measure generating the observations denoted by P. We further assume that $h_k^{(l)}(x_1, \ldots, x_k)$ is a measurable and symmetric kernel function $h_k^{(l)}: \mathcal{I}^k \to \mathrm{R}$. In the sequel we shall write h, instead of $h_k^{(l)}$. For $c = 1, \ldots, k-1$ we shall also define

$$h_c = E(h|\sigma(X_1, \ldots, X_c)).$$

Here and elsewhere, $E(\cdot|\sigma(X_1, \ldots, X_c))$ denotes the conditional expectation with respect to the σ-field generated by $X_1, \ldots, X_c$.

For given integers $1 \leq k \leq l$ let $\mathcal{S}_{l,k}$ denote a set of all k-subsets of $\{1, \ldots, l\}$ and let $U_l^{(k)}$ denote the U-statistic associated with h, *i.e.*,

$$U_l^{(k)} = \binom{l}{k}^{-1} \sum_{S \in \mathcal{S}_{l,k}} h(x_{i_1}, \ldots, x_{i_k})$$

where the sum is taken over all possible k-subsets $S = \{1 \le i_1 < \cdots < i_k \le 1\}$ of $\{1, \ldots, l\}$. In the sequel we will often use more convenient notation $h(S)$ instead of $h(x_{i_1}, \ldots, x_{i_k})$. If $E|h| < \infty$, then $U_l^{(k)}$ is an unbiased estimator of the functional

$$\theta_l = \int_{\mathrm{R}} \cdots \int_{\mathrm{R}} h(x_1, \ldots, x_k) P(\mathrm{d}x_1) \cdots P(\mathrm{d}x_k).$$

The traditional martingale representation of $U_l^{(k)}$ is given by

$$U_l^{(k)} - EU_l^{(k)} = \sum_{c=r}^{k} \binom{k}{c} U_{lc} \tag{1}$$

where

$$U_{lc} = \binom{l}{c}^{-1} \sum_{1 \le i_1 < \cdots < i_c \le n} g_c(X_{i_1}, \ldots, X_{i_c}),$$

$$g_c(x_1, \ldots, x_c) = \int_{\mathrm{R}} \cdots \int_{\mathrm{R}} h_c(y_1, \ldots, y_m) \prod_{s=r}^{c} (\delta_{x_s}(\mathrm{d}y_s) - P(\mathrm{d}y_s))$$

and $1 \le r \le k$.

The above formula is known as the H-decomposition. For given l, k the integer $r - 1$ is called *the degree of degeneration* of $U_l^{(k)}$, whereas the integer $k - r + 1$ is called *the order* of $U_l^{(k)}$. We say that $U_l^{(k)}$ is of *infinite order*, if $k - r + 1 \to \infty$ as $l \to \infty$.

It is well known that for any fixed c the U_{lc}'s are martingales with respect to the appropriate sequences of σ-fields (see, for instance, Lee, 1990) and if we additionally assume $Eh_k^2 < \infty$ for $l \ge 1$ that

$$\mathrm{Cov}(U_{l,c_1}, U_{l,c_2}) = 0 \quad \text{for } c_1 \ne c_2. \tag{2}$$

The H-decomposition plays thus a fundamental role in U-statistics theory since it allows us to represent $U_l^{(k)}$ as a linear combination of uncorrelated U-statistics of fixed order (U_{lc}) based on the kernel functions g_c (often called "canonical kernels") for $c = 1, \ldots, m$.

Elementary but somewhat tedious calculation shows that the variance of U-statistic of degree $r - 1$ is given by (cf. *e.g.*, Lee, 1990)

$$\mathrm{Var}(U_l^{(k)}) = \binom{l}{k}^{-1} \sum_{c=r}^{k} \binom{k}{c} \binom{l-k}{k-c} \mathrm{Var}\, h_c. \tag{3}$$

On the other hand, the application of the H-decomposition and the relation (2) gives the equivalent formula

$$\mathrm{Var}(U_l^{(k)}) = \sum_{c=r}^{k} \binom{l}{c}^{-1} \binom{k}{c} \mathrm{Var}\, g_c. \tag{4}$$

Using (3) and (4) one may find the relation between $\mathrm{Var}\, h_c$ and $\mathrm{Var}\, g_c$ for $1 \le c \le k$

$$\mathrm{Var}\, g_c = \sum_{i=r}^{c} (-1)^{c-i} \binom{c}{i} \mathrm{Var}\, h_i. \tag{5}$$

For any class of k-subsets $\mathcal{D}$ s.t. $\mathcal{D} \subseteq \mathcal{S}_{l,k}$ an *incomplete* U-statistic is defined as

$$U_{\mathcal{D}}^{(k)} = \frac{1}{|\mathcal{D}|} \sum_{S \in \mathcal{D}} h(S) \tag{6}$$

where $|\mathcal{D}|$ denotes the cardinality of $\mathcal{D}$. The class $\mathcal{D}$ is often referred to as a *design* of an incomplete U-statistic. The design is called *balanced* (*equireplicate*) if it has the property that every sample point x_i, or equivalently every index i $(1 \leq i \leq l)$, occurs in the same number of k-subsets belonging to $\mathcal{D}$.

The variance formulae (3) and (4) have their analogues for $U_{\mathcal{D}}^{(k)}$. If f_c is the number of pairs S_1, S_2 in the design $\mathcal{D}$ having exactly c elements in common then

$$\operatorname{Var} U_{\mathcal{D}}^{(k)} = \frac{1}{|\mathcal{D}|^2} \sum_{c=r}^{k} f_c \operatorname{Var} h_c. \tag{7}$$

Additionally, if S is a set in $S_{l,v}$ for $1 \leq v \leq k$ and $n(S)$ stands for the number of k-subsets in the design $\mathcal{D}$ which contain S, *i.e.*, $n(S) = |\{S' \in \mathcal{D}\colon S \subset S'\}|$ then

$$\operatorname{Var} U_{\mathcal{D}}^{(k)} = \frac{1}{|\mathcal{D}|^2} \sum_{v=r}^{k} B_v \operatorname{Var} g_v. \tag{8}$$

where

$$B_v = \sum_{c=v}^{k} f_c \binom{c}{v} = \sum_{S \in \mathcal{S}_{l,v}} n^2(S)$$

From (3)–(4) and (7)–(8) it is not difficult to see that for $\mathcal{D} \subseteq \mathcal{S}_{l,k}$ we must always have $\operatorname{Var} U_l^{(k)} \leq \operatorname{Var} U_{\mathcal{D}}^{(k)}$ with equality only if $\mathcal{D} = \mathcal{S}_{l,k}$. Therefore, the incomplete statistic is always less efficient than the complete one. However, for appropriately chosen design $\mathcal{D}$ the increase in the variance of U may be not too large and the loss of the estimation precision may be offset by the considerable simplification of the statistic. The problem of the optimal choice of $\mathcal{D}$ is, therefore, central to the theory of incomplete U-statistics (cf. *e.g.*, Lee, 1982).

2 PERMANENT DESIGN

In order to introduce a permanent design it will be convenient to consider a U-statistic of degree m based on a double-indexed sequence (matrix) of i.i.d. real random variables $\{X_{ij}\}$ with $i = 1, \ldots, m; j = 1, \ldots, n$ and $1 \leq m \leq n$

$$U_{mn}^{(m)} = \binom{mn}{m}^{-1} \sum_{S \in \mathcal{S}_{mn,m}} h(s) \tag{9}$$

where $\mathcal{S}_{mn,m}$ is a set of all m-subsets of the ordered pairs of indices $\{(i,j) | 1 \leq i \leq m;$ $1 \leq j \leq n\}$. The above definition is equivalent to (1) with $k = m$, $l = mn$ and the sequence $X_{11}, \ldots, X_{1n}, X_{21}, \ldots, X_{2n}, \ldots, X_{m1}, \ldots, X_{mn}$ obtained by vectorization of the matrix $[X_{ij}]$.

The following definitions will be useful in the sequel.

DEFINITION 1 *We will say that a design $\mathcal{D} \subseteq \mathcal{S}_{mn,m}$ is meager if for $1 \le v \le m$ the number of elements in $\mathcal{S}_{mn,v}$ being subsets of sets belonging to $\mathcal{D}$ equals* $\binom{n}{v}\binom{m}{v}v!$

In particular, for any meager design $\mathcal{D}$ we have $|\mathcal{D}| = n^{[m]} = \binom{n}{m}m!$. Let us also note that the notion of a meager design in our setting is somewhat related to that of a *strongly regular graph design* known in the experimental design theory.

DEFINITION 2 *A design $\mathcal{D} \subseteq \mathcal{S}_{mn,m}$ is permanent if it consists of m-subsets of the form*

$$\{(1, i_1), (2, i_2), (3, i_3), \ldots, (m, i_m)\} \tag{10}$$

where $1 \le i_1 \ne \cdots \ne i_m \le n$.

In the sequel we shall denote a permanent design by $\mathcal{D}_{\text{per}}$. The design derives its name from the fact that an incomplete version of (9) based on $\mathcal{D}_{\text{per}}$ with the primitive kernel $h(x_1, x_2, \ldots, x_m) = x_1 \cdots x_m$ gives a permanent of matrix $[X_{ij}]$. This connection between permanent designs and permanents will be explored in more detail latter in the paper.

The following theorem describes some of the basic properties of a permanent design.

THEOREM 1 *The design $\mathcal{D}_{\text{per}}$ is balanced (equreplicate) and meager. It is also a minimum variance design in the class of meager designs, that is, for fixed integers $1 \le m \le n$ it minimizes the variance of an incomplete analogue of (9) among all meager designs.*

Proof For given v $(1 \le v \le m)$ consider $S \in \mathcal{S}_{mn,v}$. First, we show that $\mathcal{D}_{\text{per}}$ is meager, that is, the number of different sets S contained in sets belonging to $\mathcal{D}_{\text{per}}$ is $\binom{m}{v}\binom{n}{v}v!$. To see this, consider a set of ordered pairs of indices (coordinates) $\{(i, j) | 1 \le i \le m;\ 1 \le j \le n\}$. Taking any v-subset is equivalent to fixing the values of v first coordinates and v second coordinates, which can be done in $\binom{m}{v}\binom{n}{v}$ different ways; for the selected v pairs of indices there is $v!$ ways to be included into one of the sets of the form (10). Second, we show that $\mathcal{D}_{\text{per}}$ is of minimum variance among meager designs. For given S let us consider $n(S)$ (see Sec. 1 for definition). If S is a subset of one of the m-sets of the form (10) then, from the definition of the permanent design it follows that for such S.

$$n(S) = \binom{n-v}{m-v}(m-v)!. \tag{11}$$

On the other hand, if S is not a subset of one of the m-sets of the form (10) then $n(S) = 0$. From the general theory (cf. *e.g.*, Lee, 1990) it is known that for any design $\mathcal{D} \subseteq \mathcal{S}_{mn,m}$ such that $|D| = \binom{n}{m}m!$ we have $\sum_{S \in \mathcal{S}_{mn,v}} n(S) = \binom{m}{v}\binom{n}{m}m!$. Since the quadratic $\sum x_i^2$. subject to $\sum x_i = c > 0$ and $x \ge 0$ is minimized by taking all the x_i's equal, it follows that for the meager designs $\mathcal{D}_{\text{per}}$ minimizes the quantities B_v in the formula (8) and hence the variance. The fact that $\mathcal{D}_{\text{per}}$ is balanced follows directly from (11) with $v = 1$ and the fact that the design is meager. ■

It is not difficult to show that there exist designs which are meager, balanced, and have strictly larger variance than the permanent design. Therefore, the above theorem is meaningful.

In the sequel we shall refer to any incomplete version of (9) based on $\mathcal{D}_{\mathrm{per}}$ as a U-statistic of permanent design (USPD) and denote it by $U^{(m)}_{\mathcal{D}_{\mathrm{per}}}$.

An obvious connection between a USPD and the U-statistic (9) is given by the following

THEOREM 2 *For given fixed integers* $1 \le m \le n$ *let* Π *be the set of all possible permutations of the elements of* $m \times n$ *matrix* $[X_{ij}]$ *and for given* $\sigma \in \Pi$ *let* $U^{(m)}_{\mathcal{D}_{\mathrm{per}}}(\sigma)$ *denote a corresponding statistic. Then*

$$U^{(m)}_{mn} = \frac{1}{(mn)!}\sum_{\sigma\in\Pi} U^{(m)}_{\mathcal{D}_{\mathrm{per}}}(\sigma).$$

Proof Direct calculation. ∎

The above formula shows that the complete U-statistic (9) is simply a symmetrized version of a USPD. Our next two results show that asymptotically the variances of USPD and its symmetrization coincide for a large class of kernel functions.

THEOREM 3 *The variance of a USPD is given by*

$$\operatorname{Var} U^{(m)}_{\mathcal{D}_{\mathrm{per}}} = \sum_{v=r}^{m} \frac{\binom{m}{v}}{\binom{n}{v}} \frac{1}{v!} \operatorname{Var} g_{v_1} \tag{12}$$

or, equivalently,

$$\operatorname{Var} U^{(m)}_{\mathcal{D}_{\mathrm{per}}} = \frac{1}{\binom{n}{m} m!} \sum_{v=r}^{m} \binom{m}{v} \Psi(m-v, n-v) \operatorname{Var} h_v \tag{13}$$

where

$$\Psi(i,j) = \sum_{c=0}^{i} (-1)^c \binom{i}{c}\binom{j-c}{i-c}(i-c)! \quad \textit{for } 0 \le i \le m-1;\ n-m \le j \le n-1.$$

Proof In the proof of Theorem 1 we have shown that $\mathcal{D}_{\mathrm{per}}$ is meager and that for any v ($1 \le v \le m$) and any $S \in \mathcal{S}_{mn,v}$ being a subset of one of the m-sets belonging to $\mathcal{D}_{\mathrm{per}}$, $n(S)$ is given by (11). This, along with the formula (8) entails (12), since now

$$B_v = \binom{m}{v}\binom{n}{v} v! \binom{n-v}{m-v}^2 (m-v)!^2$$

for $1 \le v \le m$. The formula (13) may be now obtained from (12) with the help of the relation (5) by substituting for the $\operatorname{Var} g_c$'s the appropriate expressions involving only the quantities $\operatorname{Var} h_c$ and changing the order of summation. Equivalently, as in Rempala and Wesolowski (1999), it may also be inferred directly from (7) by verifying that for USPD we have $f_c = \binom{n}{m} m! \binom{m}{c} \Psi(m-c, n-c)$. The drawback of this last approach is in the fact that it requires quite laborious combinatorial calculations. ∎

3 ASYMPTOTIC PROPERTIES OF USPD

Let us first show that USPD's are asymptotically efficient. For any incomplete U-statistic (6) let ARE be its asymptotic relative efficiency as compared with the complete statistic $U_l^{(k)}$; that is,

$$\text{ARE} = \lim_{l\to\infty} \frac{\operatorname{Var} U_l^{(k)}}{\operatorname{Var} U_{\mathcal{D}}^{(k)}}.$$

THEOREM 4 *Suppose that for the U-statistic* (9) *of a fixed* (*i.e.*, *independent of* n, m) *degree of degeneration* $r-1$ *we have* $Eh^2 < \infty$, $0 < \liminf \operatorname{Var} g_r$, *and* $\operatorname{Var} g_v \le c_v$ *for* $v \ge r$ *and some constants* c_v (*independent of* m, n) *which satisfy*

$$\sum_{v=r}^{\infty} \frac{c_v}{v!} < \infty. \tag{14}$$

If $n \ge m \to \infty$ *and* $m/n \to \lambda \ge 0$, *then* ARE *of* USPD *vis a vis the complete U-statistic* (9) *equals one. Furthermore, for* $r = 1$ *the result remains valid if* $m = m_1 \ge 1$ *is a fixed integer and* $n \to \infty$.

Remark 1 The above result appears to be quite useful, since, as we have already noted, the size of any meager design $\mathcal{D}$ is $|\mathcal{D}| = \binom{n}{m} m!$ which even for fixed integers $1 \le m \le n$ is usually a much smaller number than $\binom{mn}{m}$ – the size of a complete design in (9). Asymptotically, this is even more apparent, since by virtue of the Stirling formula we have, for some universal constant $C > 0$,

$$\frac{\binom{n}{m} m!}{\binom{mn}{m}} \le C\sqrt{m}\,\exp(-m) \to 0 \quad \text{as } m \to \infty$$

Remark 2 The condition (14) may be thought of as the requirement of at most exponential-type growth (in v) of the array of the $\operatorname{Var} g_v$'s.

Proof Let us note that by (12) and (4) we have

$$\frac{\operatorname{Var} U_{\mathcal{D}_{\text{per}}}^{(m)}}{\operatorname{Var} U_{mn}^{(m)}} - 1 = \frac{\sum_{v=r}^{m} \left[\left(\binom{m}{v} \Big/ \binom{n}{v} \right) (1/v!) - \left(\binom{m}{v}^2 \Big/ \binom{mn}{v} \right) \right] \operatorname{Var} g_v}{\sum_{v=r}^{m} \left(\binom{m}{v}^2 \Big/ \binom{mn}{v} \right) \operatorname{Var} g_v} \ge 0 \tag{15}$$

as each term in the sum in the numerator is positive, in view of the inequality

$$\binom{mn}{v} \ge \binom{m}{v}\binom{n}{v}(v!)^{-1}$$

which is valid for all integers $1 \leq v \leq m \leq n$. On the other hand, for the expression on the right-hand side of the equality sign in (15), we clearly have for sufficiently large n

$$\frac{\sum_{v=r}^{m}\left[\left(\binom{m}{v}\Big/\binom{n}{v}\right)(1/v!) - \left(\binom{m}{v}^2\Big/\binom{mn}{v}\right)\right]\operatorname{Var} g_v}{\sum_{v=r}^{m}\left(\binom{m}{v}^2\Big/\binom{mn}{v}\right)\operatorname{Var} g_v}$$

$$\leq \frac{\sum_{v=r}^{m}\left[\left(\binom{m}{v}\Big/\binom{n}{v}\right)(1/v!) - \left(\binom{m}{v}^2\Big/\binom{mn}{v}\right)\right]\operatorname{Var} g_v}{\left(\binom{m}{r}^2\Big/\binom{mn}{r}\right)\operatorname{Var} g_r}$$

$$\leq C\frac{\sum_{v=r}^{m}\left[\left(\binom{m}{v}\Big/\binom{n}{v}\right)(1/v!) - \left(\binom{m}{v}^2\Big/\binom{mn}{v}\right)\right]\operatorname{Var} g_v}{\left(\binom{m}{r}^2\Big/\binom{mn}{r}\right)} \tag{16}$$

where $C > 0$ is a universal constant. Let us show that under the assumptions of the theorem the latest expression tends to zero. We will prove this separately for each of the following cases.

Case 1 $\lambda > 0$.

Let $\varepsilon > 0$ be arbitrarily small and let m_0 be an integer such that $\sum_{v=m_0}^{\infty}(v!)^{-1}c_v < \varepsilon$ (the existence of m_0 is guaranteed by (14)). Since for any fixed integer v with $r \leq v \leq m_0$ we have that $\binom{m}{v}^2\Big/\binom{mn}{v} \to \lambda^v/v!$ and $\left(\binom{m}{v}\Big/\binom{n}{v}\right) \to \lambda^v$ as $n \geq m \to \infty$ with $m/n \to \lambda$, then, for sufficiently large m, n

$$\frac{\sum_{v=r}^{m}\left[\left(\binom{m}{v}\Big/\binom{n}{v}\right)(1/v!) - \left(\binom{m}{v}^2\Big/\binom{mn}{v}\right)\right]\operatorname{Var} g_v}{\left(\binom{m}{r}^2\Big/\binom{mn}{r}\right)}$$

$$\leq \frac{2r!}{\lambda^r}\sum_{v=r}^{m_0}\left[\frac{\binom{m}{v}}{\binom{n}{v}}\frac{1}{v!} - \frac{\binom{m}{v}^2}{\binom{mn}{v}}\right]c_v + \varepsilon\frac{2r!}{\lambda^r}$$

$$\leq \varepsilon\frac{2r!}{\lambda^r} + \varepsilon\frac{2r!}{\lambda^r} = \varepsilon\frac{4r!}{\lambda^r}$$

and the assertion follows in view of the relation (14) and the arbitrary choice of ε.

Case 2 $\lambda = 0, m \to \infty$.

In this case,

$$\frac{\sum_{v=r}^{m}\left[\left(\binom{m}{v}\Big/\binom{n}{v}\right)(1/v!) - \left(\binom{m}{v}^2\Big/\binom{mn}{v}\right)\right]\operatorname{Var} g_v}{\left(\binom{m}{r}^2\Big/\binom{mn}{r}\right)}$$

$$= \left[\frac{\binom{m}{r}}{\binom{n}{r}}\frac{\binom{mn}{r}}{\binom{m}{r}^2}\frac{1}{r!} - 1\right]\operatorname{Var} g_r + \sum_{v=r+1}^{m}\left[\frac{\binom{m}{v}}{\binom{n}{v}}\frac{1}{v!}\frac{\binom{mn}{r}}{\binom{m}{r}^2} - \frac{\binom{m}{v}^2}{\binom{mn}{v}}\frac{\binom{mn}{r}}{\binom{m}{r}^2}\right]\operatorname{Var} g_v. \tag{17a, b}$$

Note that

$$\frac{\binom{m}{r}}{\binom{n}{r}}\frac{\binom{mn}{r}}{\binom{m}{r}^2}\frac{1}{r!} = \frac{\binom{mn}{r}}{r!\binom{n}{r}\binom{m}{r}}$$

$$= \frac{mn(mn-1)\cdots(mn-r+1)}{m(m-1)\cdots(m-r+1)n(n-1)\cdots(n-r+1)}$$

$$= \frac{(1-1/mn)\cdots(1-(r-1)/mn)}{(1-1/m)\cdots(1-(r-1)/m)(1-1/n)\cdots(1-(r-1)/n)} \to 1 \tag{18}$$

as $m, n \to \infty$. Thus, (17a) $\to 0$ as $m, n \to \infty$ and we only need to argue that so does (17b). To this end, let us note that by (18), for sufficiently large m, n,

$$\sum_{v=r+1}^{m}\left[\frac{\binom{m}{v}}{\binom{n}{v}}\frac{1}{v!}\frac{\binom{mn}{r}}{\binom{m}{r}^2} - \frac{\binom{m}{v}^2}{\binom{mn}{v}}\frac{\binom{mn}{r}}{\binom{m}{r}^2}\right]\operatorname{Var} g_v$$

$$\le \sum_{v=r+1}^{m}\frac{\binom{m}{v}}{\binom{n}{v}}\frac{1}{v!}\frac{\binom{mn}{r}}{\binom{m}{r}^2}c_v \le 2\sum_{v=r+1}^{m}\frac{\binom{m}{v}}{\binom{n}{v}}\frac{r!}{v!}\frac{\binom{n}{r}}{\binom{m}{r}}c_v$$

$$= 2\sum_{v=r+1}^{m}\frac{\binom{m-r}{v-r}}{\binom{n-r}{v-r}}\frac{r!}{v!}c_v \le 2r!\frac{m-r}{n-r}\sum_{v=r+1}^{m}\frac{1}{v!}c_v \to 0,$$

in view of (14) and $m/n \to 0$. The latter inequality follows since for $1 \le v \le m \le n$ the expression $\binom{m}{v}\Big/\binom{n}{v}$ is a non-increasing function of v. The assertion follows now via (15).

Case 3 $r = 1$ and $m = m_1 \geq 1$ is a constant.

If $m_1 = 1$ the result is obvious. If $m_1 > 1$ then proceeding similarly as above we see that for $r = 1$ the expression (17a) equals zero and we only need to show that (17b) tends to zero as n increases. This is obvious upon noticing that for $v = 2, \ldots, m_1$ the expressions in the square brackets are of order $O(n^{-v+1})$. The theorem is proved. ∎

Remark 3 Since for any two sequences $\{a_k\}$ and $\{b_k\}$ of nonnegative real numbers satisfying $a_k = \sum_{c=1}^{k} \binom{k}{c} b_c$ for $k \geq 1$, the condition $\sum_{k=1}^{\infty} b_k/k! < \infty$ is equivalent to $\sum_{k=1}^{\infty} a_k/k! < \infty$, in view of relation (5) it is easily seen that a small modification of the above proof allows us for an alternative formulation of Theorem 4 where the $\operatorname{Var} g_v$'s are replaced with the $\operatorname{Var} h_v$'s. In this alternative form the assumptions of the theorem are perhaps easier to verify and, in particular, it is readily seen that they are satisfied whenever $\liminf_n \operatorname{Var} h_r > 0$ and $\sup_n Eh^2 < \infty$.

As an immediate consequence of Theorem 4 we obtain the following

THEOREM 5 *Suppose that the non-degenerate* ($r = 1$) *U-statistic* (9) *satisfies the assumptions of Theorem* 4.

(i) *If* $n - m \to \infty$ *with* $m/n \to 0$, *then*

$$\frac{U_{\mathcal{D}_{\text{per}}}^{(m)}}{(\operatorname{Var} U_{\mathcal{D}_{\text{per}}}^{(m)})^{1/2}} - \frac{U_{mn}^{(m)}}{(\operatorname{Var} U_{mn}^{(m)})^{1/2}} \xrightarrow{\text{Pr}} 0.$$

(ii) *If* $n, m \to \infty$ *with* $m/n \to \lambda > 0$, *then*

$$U_{\mathcal{D}_{\text{per}}}^{(m)} - U_{mn}^{(m)} \xrightarrow{\text{Pr}} 0.$$

Proof Let us note that for any U-statistic $U_l^{(k)}$ and its incomplete version $U_{\mathcal{D}}^{(k)}$ given by (6), due to the fact that the random variables $h(S)$, $S \in D$ are equidistributed, we have

$$\begin{aligned}\operatorname{Cov}(U_l^{(k)}, U_{\mathcal{D}}^{(k)}) &= \operatorname{Cov}\left(U_l^{(k)}, \frac{1}{|\mathcal{D}|}\sum_{S\in\mathcal{D}} h(S)\right) = \frac{1}{|\mathcal{D}|}\sum_{S\in\mathcal{D}} \operatorname{Cov}(U_l^{(k)}, h(S)) \\ &= \operatorname{Cov}(U_l^{(k)}, h(S_1)) \quad \text{for any } S_1 \in \mathcal{D}.\end{aligned}$$

Repeating the above argument with $\mathcal{D} = S_{l,k}$ entails

$$\operatorname{Cov}(U_l^{(k)}, U_{\mathcal{D}}^{(k)}) = \operatorname{Var} U_l^{(k)}.$$

Applying the last identity to $U_{mn}^{(m)}$ and $U_{\mathcal{D}_{\text{per}}}^{(m)}$ under the assumptions of (i) we obtain

$$\operatorname{Var}\left(\frac{U_{\mathcal{D}_{\text{per}}}^{(m)}}{(\operatorname{Var} U_{\mathcal{D}_{\text{per}}}^{(m)})^{1/2}} - \frac{U_{mn}^{(m)}}{(\operatorname{Var} U_{mn}^{(m)})^{1/2}}\right) = 2 - 2\left(\frac{\operatorname{Var} U_{mn}^{(m)}}{\operatorname{Var} U_{\mathcal{D}_{\text{per}}}^{(m)}}\right)^{1/2} \to 0$$

in view of Theorem 4, and the result holds via Chebychev's Inequality. The second part of the theorem follows similarly, since under the assumptions of Theorem 4 and (ii) we have that $0 < \liminf \operatorname{Var} U_{mn}^{(m)}$ and $\limsup \operatorname{Var} U_{\mathcal{D}_{\text{per}}}^{(m)} < \infty$. ■

4 APPLICATIONS

In this section we would like to provide some examples of possible applications of our results. First, we consider and obvious efficiency comparison in the case of a finite order U-statistic. For the sake of simplicity we take $m = 2$.

4.1 Relative Efficiency of the Incomplete Sample Variance

Consider $m = 2$, $\{X_{ij}\}$ i.i.d. real valued random variables with variance $0 < \sigma^2$ and central fourth moment $\mu_4 < \infty$. If we take $h(x, y) = (x - y)^2/2$ then obviously $U_{2n}^{(2)} = S_{(2n)}^2$ is the usual sample variance estimator

$$\begin{aligned} S_{(2n)}^2 &= \sum_{1 \le i \le 2, 1 \le j \le n} \frac{(X_{ij} - \bar{X})^2}{2n - 1} \\ &= \frac{1}{2n(2n-1)} \sum_{\substack{\{(i,j) \neq (k,l): \\ 1 \le i,k \le 2; 1 \le j,l \le n\}}} \frac{1}{2}(X_{ij} - X_{kl})^2 \end{aligned} \tag{19}$$

where

$$\bar{X} = \sum_{1 \le i \le 2, 1 \le j \le n} \left(\frac{X_{ij}}{2n} \right).$$

Since $r = 1$ in this case, our Theorem 4 applies for USPD

$$S_{\mathcal{D}_{\text{per}}}^2 = \frac{1}{n(n-1)} \sum_{1 \le j \ne 1 \le n} \frac{1}{2}(X_{1j} - X_{2l})^2$$

which for any given n contains less than a half of the terms present in (19).

Let us note that in this case a permanent design is also a minimum variance design since for a U-statistic of degree $m = 2$ (with square integrable kernel) any balanced design $\mathcal{D} \subseteq S_{2n,2}$ satisfying $|\mathcal{D}| = 2\binom{2n}{2}$ is meager. In particular, it follows that in this case a permanent design is more efficient than a random design of the same size.

A direct comparison of the variances can be also performed, since

$$\operatorname{Var} S_{(2n)}^2 = \frac{\mu_4 - \sigma^4}{2n} + \frac{\sigma^4}{\binom{2n}{2}}$$

by applying (4) with $l = 2n$ and $k = 2$, and

$$\operatorname{Var} S^2_{\mathcal{D}_{\text{per}}} = \frac{\mu_4 - \sigma^4}{2n} + \frac{\sigma^4}{2\binom{n}{2}}$$

by (12) with $m = 2$. As we can see the difference in the above expressions is only in the term of order $O(n^{-2})$. For the sake of example, taking $\sigma^2 = 1$ and $\mu_4 = 3$, we have tabulated below the ratio of $\operatorname{Var} S^2_{(2n)}/\operatorname{Var} S^2_{\mathcal{D}_{\text{per}}}$ for several different values of n.

n	$Var\, S^2_{(2n)}/Var\, S^2_{\mathcal{D}_{per}}$	$\binom{2n}{2}$	% used by PD
5	0.89	45	44%
10	0.95	190	47%
20	0.97	780	49%
50	0.99	4950	49.5%

As we can see from the above table, the efficiency of the permanent design appears reasonable, even with a relatively small sample size. However, the permanent design is still fairly large and up to half of the size of the complete design. In our next example we will thus consider the case when the size of our incomplete design, relative to the complete one, tends to 0.

4.2 Relative Efficiency of USPD's for the Kernels of Increasing Order

Let $\{X_{ij}\}$ $(1 \le i \le m,\ 1 \le j \le n)$ be i.i.d. Bernoulli random variables with mean $1/2$. Let us consider the U-statistic based on the kernel $h(x_1, \ldots, x_m) = x_1 \cdots x_m$ and suppose that $m \to \infty$ as $n \to \infty$.

Then $U^{(m)}_{mn} = T^{(m)}_{mn}$ where

$$T^{(m)}_{mn} = \frac{\binom{\sum_{ij} X_{ij}}{m}}{\binom{mn}{m}}$$

and the summation above is taken over all the indices i, j. In this case

$$\operatorname{Var} T^{(m)}_{mn} = \frac{1}{2^m} \sum_{v=1}^{m} \frac{\binom{m}{v}^2}{\binom{mn}{v}}$$

and

$$\operatorname{Var} T^{(m)}_{\mathcal{D}_{\text{per}}} = \frac{1}{2^m} \sum_{v=1}^{m} \frac{\binom{m}{v}}{\binom{n}{v} v!}$$

Below we have tabulated several values of the ratio $\mathrm{Var}\, T_{mn}^{(m)}/\mathrm{Var}\, T_{\mathcal{D}_{\mathrm{per}}}^{(m)}$ for three different sequences $m = m(n)$ and different total sample sizes mn. All non-integer values have been rounded to the nearest integer.

m(n)	*Approx. sample size* ($mn \approx$)	*Var* $T_{mn}^{(m)}$ / *Var* $T_{\mathcal{D}_{per}}^{(m)}$	*% used by PD*
$\ln(n)$	10	0.938	44.4%
	100	0.989	21%
	1000	0.998	3.69%
$\sqrt{n}$	10	0.938	44.4%
	100	0.976	2.5%
	1000	0.995	0.2%
$n/2$	10	0.938	44.4%
	100	0.950	0.13%
	1000	0.981	$<10^{-10}$%

From the above table we can clearly see that whereas the efficiency of the incomplete statistic is higher for slowly growing m, the biggest gains in design reduction are achieved for m growing at a faster rate. This is consistent with the observation made earlier in Remark 1.

4.3 Limiting Laws for Generalized Permanents

Expanding somewhat on the setting of the last example, let us derive with the help of Theorem 5, the limiting laws of USPD's with product kernels.

Let $[X_{ij}]$ be an $m \times n$ ($m \leq n$) matrix of i.i.d. random variables and let ϕ be a function $\phi\colon \mathcal{I} \to \mathrm{R}$ such that $\mu_\phi = E\phi(X_{11}) \neq 0$ and $0 < \sigma_\phi^2 = \mathrm{Var}\, \phi(X_{11}) < \infty$. Let us also consider the product kernel function $h(x_1, \ldots, x_m) = \phi(x_1) \cdots \phi(x_m)$ where without loss of generality we may assume that $\mu_\phi = 1$. For this particular kernel choice the U-statistic (9) is known as a (normalized) elementary symmetric polynomial of degree m. In the sequel we shall denote it by $S_{mn}^{(m)}(\phi)$ and its USPD counterpart by $S_{\mathcal{D}_{\mathrm{per}}}^{(m)}(\phi)$. If we define *a generalized permanent* of the matrix $[X_{ij}]$ by

$$\mathrm{Per}_\phi[X_{ij}] = \sum_{1 \leq i_1 \neq i_2 \neq \cdots \neq i_m \leq n} \phi(X_{1,i1}) \cdots \phi(X_{m,i_m})$$

then obviously

$$\mathrm{Per}_\phi[X_{ij}] = n^{[m]} S_{\mathcal{D}_{\mathrm{per}}}^{(m)}(\phi)$$

where $n^{[m]} = \binom{n}{m} m!$.

THEOREM 6 *Let $\mathcal{N}$ denote a standard normal random variable.*

(i) *If $n - m \to \infty$ with $m/n \to 0$ then*

$$\sqrt{\frac{n}{m}} \frac{\mathrm{Per}_\phi[X_{ij}] - n^{[m]}}{\sigma_\phi n^{[m]}} \xrightarrow{D} \mathcal{N}.$$

(ii) *If* $n, m \to \infty$ *with* $m/n \to \lambda > 0$ *then*

$$\frac{\mathrm{Per}_\phi[X_{ij}]}{n^{[m]}} \xrightarrow{D} \exp\left(\sqrt{\lambda}\sigma_\phi \mathcal{N} - \frac{\lambda\sigma_\phi^2}{2}\right).$$

Proof It is readily seen that for $S_{mn}^{(m)}(\phi)$ we have $r = 1$, $ES_{mn}^{(m)}(\phi) = 1$, Var $g_v = \sigma_\phi^{2v}, v = 1, \ldots, m$. It is also well known (cf. *e.g.*, van Es and Helmers, 1988) that as $k, l \to \infty$ with $k^2/l \to 0$ the asymptotic distribution of $(S_l^{(k)}(\phi) - ES_l^{(k)}(\phi))/(\mathrm{Var}\ S_l^{(k)}(\phi))^{1/2}$ is standard normal. Since $S_{mn}^{(m)}$ is a subsequence of $S_l^{(k)}(\phi)$ satisfying the assumptions of Theorem 4 as well as that on the rate of growth of m and n of part (i) of Theorem 5, the limiting distributions of $(S_{mn}^{(m)}(\phi) - ES_{mn}^{(m)}(\phi))/(\mathrm{Var}\, S_{mn}^{(m)}(\phi))^{1/2}$ and $(S_{\mathcal{D}_{\mathrm{per}}\phi}^{(m)} - E(S_{\mathcal{D}_{\mathrm{per}}}^{(m)}(\phi)))/\ (\mathrm{Var}\ (S_{\mathcal{D}_{\mathrm{per}}}^{(m)}(\phi)))^{1/2}$ must coincide, and the assertion of part (i) follows in view of the fact that as $n - m \to \infty$ and $m/n \to 0$

$$\frac{\mathrm{Var}\ S_{mn}^{(m)}(\phi)}{m\sigma_\phi^2/n} \to 1.$$

The second part of the theorem follows similarly, except that here we use the assertion of part (ii) of Theorem 5 and the result of Koroljuk and Borovskikh (1992) which states that if $l, k \to \infty$ and $k/l \to \lambda > 0$ then $S_l^{(k)}(\phi)$ converges in distribution to a lognormal random variable at the right hand side of (ii). ■

Some important special cases of the above theorem for $\{X_{ij}\}$ real i.i.d. random variables are, for instance, the following.

Example 1 Limiting law for random permanents (Rempala and Wesolowski, 1999) Taking $\phi = Id$ we obtain from Theorem 6 as its special case the limit theorem for random permanents of Rempala and Wesolowski (1999). In particular, for $m = 1$ the result reduces to the usual CLT for i.i.d. random variables. On the other hand, for $n = m$ and the X_{ij}'s are i.i.d. Bernoulli it is well known that $\mathrm{Per}[X_{ij}]$ equals the number of perfect matchings in a bipartite graph given by the reduced adjacency matrix X_{ij} and hence in this case we obtain from Theorem 6 an alternative form of the limit theorem on the asymptotic law for the number of perfect matchings of Janson (1994).

Example 2 Limiting law for the estimated probability of containment Let F be the law of the X_{ij}'s and let x be any real number for which $F(x) > 0$. Taking $\phi(X) = \phi_x(X) = I[X \le x]$ (where $I[\cdot]$ is an indicator function) we obtain the limiting laws for the USPD

$$M_{\mathcal{D}_{\mathrm{per}}}^{(m)}(x) = \sum_{1 \le i_1 \ne i_2 \ne \cdots \ne i_m \le n} \frac{I[\max(X_{1,i_1}, \ldots, X_{m,i_m}) \le x]}{n^{[m]}}$$

which again may be viewed as a generalization of the CLT for the discrete empirical process $M^{(1)}(x)$ when $m = 1$. Under some regularity conditions the extension of this result to the stationary random variables is also possible (cf. Györfi *et al.*, 1989).

4.4 Non-parametric Subsampling Scheme

As a natural application of Theorem 5 in the case of USIO let us consider the "permanent-design" version of the subsampling method of Politis and Romano (1994). The original method of Politis and Romano (which in case of an i.i.d. sample is closely related to d-jackknife) can be described as follows. For $\{X_{ij}\}$ i.i.d. random sample of size $l = mn$ let T_l be a real valued functional on $\mathcal{I}^l$ and suppose that for some sequence τ_l and some parameter $\theta(P)$

$$J_l(P) - \tau_l(T_l - \theta(P)) \xrightarrow{D} J(P) \quad \text{as } l \to \infty$$

In order to obtain an asymptotic confidence interval for $\theta(P)$ when the law $J(P)$ is not known Politis and Romano considered the empirical quantiles of the statistic

$$L_{mn}^{(m)}(x) = \binom{mn}{m}^{-1} \sum_{S \in \mathcal{S}_{mn,m}} I[\tau_m(T_m(S) - T_{mn}) \leq x] \tag{20}$$

where $T_m(S)$ is a functional T_m evaluated at the subset of data S. (Actually, the above is a special case of the method of Politis and Romano who, more generally, considered $L_k(x)$ based on the sample of size k and of order b where $b/k \to 0$ as $k, b \to \infty$.) As it was shown by Politis and Romano the method yields asymptotically correct confidence level as long as

$$\frac{\tau_m}{\tau_{mn}} \to 0 \tag{21}$$

as $m, n \to \infty$. The key observation leading to that conclusion is that for the U-statistic $UL_{mn}^{(m)}(x)$ given by (20) with T_{mn} replaced by $\theta(P)$ we have

$$UL_{mn}^{(m)}(x) \xrightarrow{\Pr} J(x, P) \tag{22}$$

at every continuity point x of $J(\cdot, P)$ and that in view of (21) the limiting laws of $L_{mn}^{(m)}(\cdot)$ and $UL_{mn}^{(m)}(\cdot)$ must coincide.

Since for large values of m,n the computation of $L_{mn}^{(m)}(\cdot)$ becomes quite difficult, it seems to be of interest to consider some incomplete version of (20) which would nevertheless share its asymptotic properties.

Under an additional assumption that $m/n \to \lambda > 0$ we could employ our method here and consider a permanent-design based version of (20). The validity of this approach follows from the fact that since the U-statistic $UL_{mn}^{(m)}(x)$ has a bounded kernel, it obviously satisfies the condition (15) and thus by (ii) of Theorem 5 the limiting laws of $UL_{mn}^{(m)}(x)$ and its permanent-design counterpart must coincide. The fact that interchanging T_{mn} and $\theta(P)$ is permissible follows, similarly as in Politis and Romano (1994), in view of condition (21).

References

Blom, G. (1976). Some properties of incomplete U-statistics. *Biometrika*, **63**(3), 573–580.

Brown, B. M. and Kildea, D. G. (1978). Reduced U-statistics and the Hodges–Lehmann estimator. *Ann. Statist.*, **6**(4), 828–835.

Enqvist, E. (1985). A note on incomplete U-statistics for stationary absolutely regular processes. *Contributions to Probability and Statistics in Honour of Gunnar Blom*. University of Lund, Lund, Germany, pp. 97–103.

Frees, E. W. (1989). Infinite order U-statistics. *Scand. J. Statist.*, **16**(1), 29–45.

Giné, E., Kwapień, S., Latala, R. and Ziun, J. (2001). The LIL for canonical U-statistics of order 2. *Ann. Probability*, **29**(1), 520–557.

Giné, E. and Zinn, J. (1994). A remark on convergence in distribution of *U*-statistics. *Ann. Probab.*, **22**(l), 117–125.

Györfi, L., Härdle, W., Sarda, P. and Vieu, P. (1989). *Nonparametric Curve Estimation from Time Series.* Springer-Verlag, Berlin.

Herrndorf, N. (1986). An invariance principle for reduced *U*-statistics. *Metrika*, **33**(3–4), 179–188.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics*, **19**, 293–325.

Janson, S. (1984). The asymptotic distributions of incomplete *U*-statistics. *Z. Wahrsch. Verw. Gebiete*, **66**(4), 495–505.

Janson, S. (1994). The numbers of spanning trees, Hamilton cycles and perfect matchings in a random graph. *Combin. Probab. Comput.*, **3**(l), 97–126.

Korolyuk, V. S. and Borovskikh, Y. V. (1992). Random permanents and symmetric statistics. *Probability Theory and Mathematical Statistics (Kiev, 1991).* World Sci. Publishing, River Edge, NJ, pp. 176–187.

Latala, R. and Zinn, J. (1999). Necessary and sufficient conditions for SLLN for *U*-statistics. *Ann. Probability*, **28**(4), 1908–1924.

Lee, A. J. (1982). On incomplete *U*-statistics having minimum variance. *Austral. J. Statist.*, **24**(3), 275–282.

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

Nowicki, K. and Wierman, J. C. (1988). Subgraph counts in random graphs using incomplete *U*-statistics methods. *Proceedings of the First Japan Conference on Graph Theory and Applications (Hakone, 1986)*, Vol. 72, pp. 299–310.

Politis, D. N. and Romano, J. P. (1994). Large sample confidence regions based on subsamples under minimal assumptions. *Ann. Statist.*, **22**(4), 2031–2050.

Rempala, G. (1998). Strong law of large numbers for *U*-statistics of varying order. *Statist. Probab. Lett.*, **39**(3), 263–270.

Rempala, G. and Wesobwski, J. (1999). Asymptotic behavior of random permanents. *Statist. Probab. Lett.*, **45**, 149–158.

Székely, G. J. (1982). A limit theorem for elementary symmetric polynomials of independent random variables. *Z. Wahrsch. Verw. Gebiete*, **59**(3), 355–359.

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