2

Properties of $\boldsymbol{P}$-statistics

2.1 Preliminaries: Martingales

In order to discuss properties of P-statistics we first need to review briefly the basic results on martingales. The theory of martingales is rich and profound and reaches far beyond the basic results stated here. Interested reader is referred to one of many excellent texts available, like for instance Billingsley (1999), Chow and Teicher (1978) or Ethier and Kurtz (1986).

For $T \subset \mathbf{R}$ consider an integrable stochastic process $(X_t)_{t\in T}$ defined on a probability space $(\Omega, \mathcal{F}, P)$.

Definition 2.1.1 (Martingale). *A family of σ-fields $(\mathcal{F}_t)_{t\in T}$ is called a filtration if $\mathcal{F}_t \subset \mathcal{F}$ for any $t \in T$ and $\mathcal{F}_s \subset \mathcal{F}_t$ for any $s \le t$, $s,t \in T$. The process $(X_t,\, \mathcal{F}_t)_{t\in T}$ is called a martingale if*

(a) X_t is $\mathcal{F}_t$-measurable for any $t \in T$
(b) $E(X_t|\mathcal{F}_s) = X_s$ for any $s \le t$, $s,t \in T$.

It turns out that in the theory of P-statistics a related concept of a backward (or reversed) martingale is also useful.

Definition 2.1.2 (Backward martingale). *Consider a family of σ-fields $(\mathcal{G}_t)_{t\in T}$ such that $\mathcal{G}_t \subset \mathcal{F}$ for any $t \in T$ and $\mathcal{G}_t \subset \mathcal{G}_s$ for any $s \le t$, $s,t \in T$. The process $(X_t,\, \mathcal{G}_t)_{t\in T}$ is called a backward martingale if*

(a) X_t is $\mathcal{G}_t$-measurable for any $t \in T$
(b) $E(X_s|\mathcal{G}_t) = X_t$ for any $s \le t$, $s,t \in T$.

Let us prove the following result known as the Doob or maximal inequality (see, Chow and Teicher 1978, chapter 7 or Ethier and Kurtz 1986, chapter 2).

Theorem 2.1.1 (Maximal inequality for martingales). *Let n be an arbitrary positive integer and $(X_t, \mathcal{F}_t)_{t\in T}$ be a martingale with a countable index set T satisfying $\{0,n\} \subset T \subset [0,\infty)$. Then for any $\varepsilon > 0$*

$$P\left(\sup_{t\in T\cap[0,n]}|X_t|>\varepsilon\right)\le\varepsilon^{-2}EX_n^2.$$

Proof. Assume first that $T\cap[0,n]=\{0=t_0,t_1,\dots,t_M=n\}$ is finite, $t_i<t_{i+1}$, $i=0,1,\dots,M-1$. For an arbitrary $\varepsilon>0$ define

$$A_k=\cap_{i=0}^{k-1}\{|X_{t_i}|\le\varepsilon\}\cap\{|X_{t_k}|>\varepsilon\},\quad k=0,\dots,M.$$

Note that A_k, $k=0,\dots,M$, are disjoint, $A_k\in\mathcal{F}_k$ for each k and

$$\left\{\sup_{t\in T\cap[0,n]}|X_t|>\varepsilon\right\}=\left\{\max_{t\in T\cap[0,n]}|X_t|>\varepsilon\right\}=\bigcup_{k=0}^{M}A_k.$$

Moreover, due to the martingale property,

$$E[(X_n-X_{t_k})X_{t_k}I(A_k)]=E[X_{t_k}I(A_k)E[X_n-X_{t_k}|\mathcal{F}_{t_k}]]=0.$$

In view of the above, we may write

$$\begin{aligned}EX_n^2&=\sum_{k=0}^{M}EX_n^2I(A_k)\\&=\sum_{k=0}^{M}E[(X_n-X_{t_k})^2I(A_k)]+E[X_{t_k}^2I(A_k)]\\&\ge\sum_{k=0}^{M}E[X_{t_k}^2I(A_k)]\\&\ge\sum_{k=0}^{M}\varepsilon^2P(A_k)=\varepsilon^2\,P\left(\bigcup_{k=0}^{M}A_k\right)\\&=\varepsilon^2\,P(\sup_{t\in T\cap[0,n]}|X_t|>\varepsilon).\end{aligned}$$

In order to prove the result for countable T consider a sequence $T_0\subset T_1\subset T_2\subset\dots$ of finite subsets of T such that $\{0,n\}\in T_0$ and $\bigcup_k T_k=T$. Then by continuity from below of the probability it follows that for any $\varepsilon>0$

$$\begin{aligned}P\left(\sup_{t\in T\cap[0,n]}|X_t|>\varepsilon\right)&=P\left(\bigcup_k\left\{\max_{t\in T_k\cap[0,n]}|X_t|>\varepsilon\right\}\right)\\&=\lim_{k\to\infty}P\left(\max_{t\in T_k\cap[0,n]}|X_t|>\varepsilon\right)\le\varepsilon^{-2}EX_n^2.\end{aligned}$$ □

The result above may be easily extended to the case of an uncountable index set, say, $T=[0,1]$, if we assume some regularity conditions on the trajectories of $(X_t)_{t\in T}$.

Proposition 2.1.1. *Assume that $T = [0,1]$ and the martingale $(X_t)_{t\in T}$ has trajectories which are a.s. cadlag, i.e., are right continuous and have left-hand limits a.s. Then*

$$P(\sup_{t\in[0,1]} |X_t| > \varepsilon) \le \varepsilon^{-2} E X_1^2.$$

□

In a similar way one may prove an analogous result for backward martingales. Let us state it for the record as

Theorem 2.1.2 (Maximal inequality for backward martingales). *Let n be an arbitrary positive integer and $(X_t, \mathcal{G}_t)_{t\in T}$ be a backward martingale with $\{n\} \subset T \subset [n, \infty)$ and T being a countable set. Then for any $\varepsilon > 0$*

$$P(\sup_{t\in T\cap[n,\infty)} |X_t| > \varepsilon) \le \varepsilon^{-2} E X_n^2.$$

Proof. The proof is similar to that of Theorem 2.1.1 and thus omitted. □

2.2 *H*-decomposition of a *P*-statistic

We first give the result extending the H-decomposition (1.15) to an arbitrary P-statistic.

Consider an infinite random matrix $\mathbb{X}^{(\infty)}$ satisfying the assumptions (A1-A2) as it was introduced in the previous chapter. We denote by $\mathbb{X}(i_1,\ldots,i_p|j_1,\ldots,j_q)$ a sub-matrix constructed out of $\mathbb{X}^{(m,n)}$ by taking only the entries at the intersections of the selected p rows $\{i_1,\ldots,i_p\} \subseteq \{1,\ldots,m\}$ and q columns $\{j_1,\ldots,j_q\} \subseteq \{1,\ldots,n\}$.

Note that $Per_h\,\mathbb{X}(i_1,\ldots,i_p|j_1,\ldots,j_q)$ may be regarded as either row- or column- symmetric function. Thus we may apply to it, in both settings, the general symmetrization operation defined by (1.2) with $\mathcal{I} = \mathbf{R}^q$, $k = p, l = m$ when regarded as a row function and with $\mathcal{I} = \mathbf{R}^p$, $k = q, l = n$ when regarded as a column function. Note that under the assumptions (A1-A2) the exchangeability requirement for the random row or column vectors is satisfied.

Denote the row symmetrization of a generalized permanent by

$$\begin{aligned} Per_h\,\mathbb{X}(p|j_1,\ldots,j_q) &= \sum_{1\le i_1<\ldots<i_p\le m} Per_h\,\mathbb{X}(i_1,\ldots,i_p|j_1,\ldots,j_q) \\ &= \pi_p^m Per_h\,\mathbb{X}(\cdot|j_1,\ldots,j_q) \end{aligned}$$

for any $\{j_1,\ldots,j_q\} \subset \{1,\ldots,n\}$ and the similar column symmetrization by

$$\begin{aligned} Per_h\,\mathbb{X}(i_1,\ldots,i_p|q) &= \sum_{1\le j_1<\ldots<j_q\le n} Per_h\,\mathbb{X}(i_1,\ldots,i_p|j_1,\ldots,j_q) \\ &= \pi_q^n Per_h\,\mathbb{X}(i_1,\ldots,i_p|\cdot) \end{aligned}$$

for any $\{i_1, \ldots, i_p\} \subset \{1, \ldots, m\}$.

Similarly,

$$\begin{aligned} Per_h\, \mathbb{X}\,(p|q) &= \sum_{1 \le j_1 < \ldots < j_q \le n} Per_h\, \mathbb{X}\,(p|j_1, \ldots, j_q) \\ &= \sum_{1 \le i_1 < \ldots < i_p \le m} Per_h\, \mathbb{X}\,(i_1, \ldots, i_p|q) \\ &= \sum_{1 \le j_1 < \ldots < j_q \le n} \sum_{1 \le i_1 < \ldots < i_p \le m} Per_h\, \mathbb{X}\,(i_1, \ldots, i_p|j_1, \ldots, j_q). \end{aligned}$$

Using this notation, we may now easily extend the Laplace expansion formula for permanents given by (1.1). Let us state it in the form of the following lemma.

Lemma 2.2.1. *Let $\mathbb{X}^{(m,n)}$ be an $m \times n$-matrix with $1 \le m \le n$. Then*

$$Per_h\, \mathbb{X}^{(m,n)} = Per_h\, \mathbb{X}\,(1, 2, \ldots, m|m). \tag{2.1}$$

□

The above formula suggests that in order to decompose $Per_h\, \mathbb{X}$, it suffices to decompose a generalized permanent of a square submatrix with columns $j_1 \ldots, j_m$

$$Per_h\, \mathbb{X}\,(1, 2, \ldots, m|j_1, j_2, \ldots, j_m).$$

Assume for convenience that $Eh = 0$ (otherwise taking $\tilde{h} = h - Eh$ in what follows).

Denote by S_q a set of permutations of $\{1, \ldots, q\}$. In view of the above (cf. (1.14)) the decomposition of $Per_h\, \mathbb{X}\,(1, 2, \ldots, m|j_1, j_2, \ldots, j_m)$ can be obtained as follows

$$\begin{aligned} &Per_h\, \mathbb{X}\,(1, 2, \ldots, m|j_1 \ldots, j_m) \\ &= \sum_{\sigma \in S_m} h(X_{1,j_{\sigma(1)}}, X_{2,j_{\sigma(2)}}, \ldots, X_{m,j_{\sigma(m)}}) \\ &= \sum_{\sigma \in S_m} \sum_{c=1}^{m} (\pi_c^m g_c)(X_{1,j_{\sigma(1)}}, X_{2,j_{\sigma(2)}}, \ldots, X_{m,j_{\sigma(m)}}) \\ &= \sum_{c=1}^{m} \sum_{1 \le i_1 < \ldots < i_c \le m} \sum_{\sigma \in S_m} g_c(X_{i_1,j_{\sigma(i_1)}}, \ldots, X_{i_c,j_{\sigma(i_c)}}). \end{aligned}$$

But by the definition of a generalized permanent and in view of (2.1)

$$
\begin{aligned}
&\sum_{\sigma \in S_m} g_c(X_{i_1, j_{\sigma(i_1)}}, \ldots, X_{i_c, j_{\sigma(i_c)}}) \\
&= \sum_{1 \le l_1 < \ldots < l_c \le m} \sum_{\tau \in S_c} g_c(X_{i_{\tau(1)}, j_{l_1}}, \ldots, X_{i_{\tau(c)}, j_{l_c}}) \\
&= \sum_{1 \le l_1 < \ldots < l_c \le m} Per_{g_c} \mathbb{X}\,(i_1, \ldots, i_c | j_{l_1}, \ldots, j_{l_c}) \\
&= Per_{g_c} \mathbb{X}\,(i_1, \ldots, i_c | j_1, \ldots, j_m).
\end{aligned}
$$

Hence

$$
\begin{aligned}
&Per_h \, \mathbb{X}\,(1, 2, \ldots, m | j_1 \ldots, j_m) \\
&= \sum_{c=1}^{m} \sum_{1 \le i_1 < \ldots < i_c \le m} Per_{g_c} \mathbb{X}\,(i_1, \ldots, i_c | j_1, \ldots, j_m) \\
&= \sum_{c=1}^{m} Per_{g_c} \mathbb{X}\,(c | j_1, \ldots, j_m).
\end{aligned}
$$

Thus, using again (2.1), we may write

$$
\begin{aligned}
Per_h \, \mathbb{X}^{(m,n)} &= Per_h \, \mathbb{X}\,(1, 2, \ldots, m | m) \\
&= \sum_{1 \le j_1 < \ldots < j_m \le n} Per_h \, \mathbb{X}\,(1, 2, \ldots, m | j_1, \ldots, j_m) \\
&= \sum_{1 \le j_1 < \ldots < j_m \le n} \sum_{c=1}^{m} Per_{g_c} \mathbb{X}\,(c | j_1, \ldots, j_m) \\
&= \sum_{c=1}^{m} \binom{n-c}{m-c} (m-c)! \sum_{1 \le j_1 < \ldots < j_c \le n} Per_{g_c} \mathbb{X}\,(c | j_1, \ldots, j_c) \\
&= \sum_{c=1}^{m} \binom{n-c}{m-c} (m-c)! \, Per_{g_c} \mathbb{X}\,(c | c).
\end{aligned}
$$

Rewriting the above as

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

and denoting

$$
U_{g_c}^{(m,n)} = \binom{m}{c}^{-1} \binom{n}{c}^{-1} Per_{g_c} \mathbb{X}\,(c | c)/c! \tag{2.2}
$$

we may summarize the above considerations as follows.

Theorem 2.2.1. *Under assumptions* (A1-A2) *on the entries of a matrix* $\mathbb{X}^{(m,n)}$*, suppose also that the kernel* h *satisfies* $E|h| < \infty$*. Then*

$$\binom{n}{m}^{-1}\left(Per_h\,\mathbb{X}^{(m,n)} - E Per_h\,\mathbb{X}^{(m,n)}\right)/m! = \sum_{c=1}^{m}\binom{m}{c}U_{g_c}^{(m,n)},$$

where the components $U_{g_c}^{(m,n)}$ *are defined by* (2.2). □

If $\mathbb{X}^{(m,n)} = \underline{\mathrm{X}}$ that is, $\mathbb{X}^{(m,n)}$ is a one dimensional projection matrix then by (1.22) a P-statistic is a U-statistic and the Theorem 2.2.1 reduces to the H-decomposition (1.15). Let us also note that for any fixed $c \geq 1$ the component $U_{g_c}^{(m,n)}$ is simply a P-statistic obtained by symmetrization (with respect to rows and columns of $\mathbb{X}^{(m,n)}$) of a generalized permanent with a kernel function g_c based on a square $c \times c$ submatrix of $\mathbb{X}^{(m,n)}$.

For the purpose of illustration let us provide a simple example of an application of Theorem 2.2.1.

Example 2.2.1 (Random permanent decomposition). Let $E\,X_{i,j} = \mu \neq 0$ and consider the kernel $h(y_1,\ldots,y_m) = \prod_{i=1}^{m}(y_i/\mu)$. Then

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

and by Theorem 2.2.1 we have

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

where

$$U_c^{(m,n)} = \binom{n}{c}^{-1}\binom{m}{c}^{-1}(\mu^c c!)^{-1}Per\,\tilde{\mathbb{X}}\,(c|c) \tag{2.3}$$

$$= \binom{n}{c}^{-1}\binom{m}{c}^{-1}(\mu^c c!)^{-1}\sum_{1\leq i_1<\ldots<i_c\leq m}\;\sum_{1\leq j_1<\ldots<j_c\leq n} Per\,\tilde{\mathbb{X}}\,(i_1,\ldots,i_c|j_1,\ldots,j_c), \tag{2.4}$$

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

The fact that the generalized H-decomposition given in Theorem 2.2.1 retains the uncorrelated martingale structure of the components $U_{g_c}^{(m,n)}$'s is verified in the following result.

Theorem 2.2.2. *Assume $m = m_n$ is a non-decreasing sequence in n. Under assumptions (A1-A2) on the entries of matrix $\mathbb{X}^{(\infty)}$, for a fixed $c \geq 1$*

(i) if $\mathcal{F}_c^{(n)} = \sigma\{U_{g_c}^{(m_n,n)}, U_{g_c}^{(m_{n+1},n+1)},\ldots\}$, *then for* $n_0(c) = \min\{n : m_n \geq c\}$

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

that is, $\left(U_{g_c}^{(m_n,n)}, \mathcal{F}_c^{(n)}\right)_{n=n_0(c),n_0(c)+1,\ldots}$ *is a backward martingale,* *(ii) the elements $(U^{(m,n)}_{g_c})_c$ of the H-decomposition of a P-statistic are orthogonal, i.e.*

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

Proof. (i) Let us denote by $U^{(m_n,n)}_{g_c}(l_1,\ldots,l_{m_{n+1}-m_n};k)$ an element of the Hoeffding decomposition (as given in Theorem 2.2.1) of an $m_n \times n$ matrix, which is obtained by deleting the rows numbered $l_1,\ldots,l_{m_{n+1}-m_n}$ and the k-th column from the $m_{n+1}\times(n+1)$ matrix, i.e.,

$$U^{(m_n,n)}_{g_c}(l_1,\ldots,l_{m_{n+1}-m_n};k) = \binom{m_n}{c}^{-1}\binom{n}{c}^{-1}(c!)^{-1}$$
$$\times \sum_{\substack{1\le i_1<\ldots<i_c\le m_{n+1}\\ \{i_1,\ldots,i_c\}\cap\{l_1,\ldots,l_{m_{n+1}-m_n}\}=\emptyset}} \; \sum_{\substack{1\le j_1<\ldots j_c\le n+1\\ k\notin\{j_1,\ldots,j_c\}}} Per_{g_c}\,\mathbb{X}\,(i_1,\ldots,i_c|j_1,\ldots,j_c).$$

Observe that

$$\sum_{k=1}^{n+1}\sum_{1\le l_1<\ldots<l_{m_{n+1}-m_n}\le m_{n+1}} U^{(m_n,n)}_{g_c}(l_1,\ldots,l_{m_{n+1}-m_n};k)$$
$$= \binom{m_{n+1}-c}{m_n-c}\frac{n+1-c}{\binom{m_n}{c}\binom{n}{c}}\; Per_{g_c}\,\mathbb{X}^{(m_{n+1},n+1)}(c|c)/c!,$$

where $\mathbb{X}^{(m_{n+1},n+1)}$ is the $m_{n+1}\times(n+1)$ dimensional matrix created from the left upper corner of the infinite matrix $\mathbb{X}^{(\infty)}$. 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

$$\sum_{k=1}^{n+1}\sum_{1\le l_1<\ldots<l_{m_{n+1}-m_n}\le m_{n+1}} U^{(m_n,n)}_{g_c}(l_1,\ldots,l_{m_{n+1}-m_n};k)$$
$$= (n+1)\binom{m_{n+1}}{m_n}U^{(m_{n+1},n+1)}_{g_c}. \tag{2.5}$$

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

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

yielding

$$E\left[\sum_{k=1}^{n+1}\sum_{1\le l_1<\ldots<l_{m_{n+1}-m_n}\le m_{n+1}} U_{g_c}^{(m_n,n)}(l_1,\ldots,l_{m_{n+1}-m_n};k)\middle|\mathcal{F}_c^{(n+1)}\right] \tag{2.6}$$

$$= (n+1)\binom{m_{n+1}}{m_n}E(U_{g_c}^{(m_n,n)}|\mathcal{F}_c^{(n+1)}). \tag{2.7}$$

On the other hand, in view of the identity (2.5), we have that (2.6) is equal to

$$(n+1)\binom{m_{n+1}}{m_n}E(U_{g_c}^{(m_{n+1},n+1)}|\mathcal{F}_c^{(n+1)}) = (n+1)\binom{m_{n+1}}{m_n}U_{g_c}^{(m_{n+1},n+1)}. \tag{2.8}$$

Comparing the expression (2.7) with that at the right-hand side of (2.8), we arrive at *(i)*.

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

$$Cov\left(Per_{g_{c_1}}\mathbb{X}(c_1|c_1), Per_{g_{c_2}}\mathbb{X}(c_2|c_2)\right) = 0.$$

To this end we will show that for any pairs of fixed sets of rows $\{i_1,\ldots,i_{c_1}\}$, $\{k_1,\ldots,k_{c_2}\}$ and columns $\{j_1,\ldots,j_{c_1}\}$, $\{l_1,\ldots,l_{c_2}\}$

$$Cov\left[Per_{g_{c_1}}\mathbb{X}(i_1,\ldots,i_{c_1}|j_1,\ldots,j_{c_1}), Per_{g_{c_2}}\mathbb{X}(k_1,\ldots,k_{c_2}|l_1,\ldots,l_{c_2})\right] = 0. \tag{2.9}$$

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

$$\begin{aligned}&Cov\left[Per_{g_{c_1}}\mathbb{X}(i_1,\ldots,i_{c_1}|j_1,\ldots,j_{c_1}), Per_{g_{c_2}}\mathbb{X}(k_1,\ldots,k_{c_2}|l_1,\ldots,l_{c_2})\right]\\ &= \sum_{\sigma\in S_{c_1},\ \tau\in S_{c_2}} E\left(g_{c_1}(X_{i_1,j_{\sigma(1)}},\ldots,X_{i_{c_1},j_{\sigma(c_1)}})\ g_{c_2}(X_{k_1,l_{\tau(1)}},\ldots,X_{k_{c_2},l_{\tau(c_2)}})\right).\end{aligned}$$

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 $j_s \in \{j_1,\ldots,j_{c_1}\}$ such that $j_s \notin \{l_1,\ldots,l_{c_2}\}$. But by the assumption of the independence of columns and the property of canonical functions (1.12) the standard conditioning argument implies now that all the summands above are zero and thus (2.9) follows. The proof of the result is complete. □

2.3 Variance Formula for a *P*-statistic

Using the representation of Theorem 2.2.1 it is possible to derive a general formula for the variance of a P-statistic which proves very useful later on.

Theorem 2.3.1 (The variance formula for a *P*-statistic). *Under the assumptions* (A1-A2) *on* $\mathbb{X}^{(\infty)}$ *let us suppose* $Eh^2 < \infty$. *Then for* $c = 1, \ldots, m$

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

where

$$D(c,d) = \sum_{s=0}^{d} \binom{d}{s} (-1)^{d-s} \rho_{c,s}$$

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

$$\begin{aligned}\rho_{c,s} = E\, \big[& g_c(X_{11}, \ldots, X_{ss}, X_{i_{s+1},s+1} \ldots, X_{i_c,c}) \times \\ & g_c(X_{11}, \ldots, X_{ss}, X_{j_{s+1},s+1} \ldots, X_{j_c,c}) \big]\end{aligned}$$

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

Proof. First, let us note that by Theorem 2.2.2 it follows that

$$\begin{aligned}Var\, \frac{Per_h\, \mathbb{X}}{\binom{n}{m} m!} = \sum_{c=1}^{m} \binom{m}{c}^2 Var\, U_{g_c}^{(m,n)} + \sum_{1 \le c_1 \neq c_2 \le m} \binom{m}{c_1} \binom{m}{c_2} \\ Cov\, \left[U_{g_{c_1}}^{(m,n)}, U_{g_{c_2}}^{(m,n)} \right] = \sum_{c=1}^{m} \binom{m}{c}^2 Var\, U_{g_c}^{(m,n)}. \qquad (2.10)\end{aligned}$$

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

$$\begin{aligned}\binom{n}{c}^2 \binom{m}{c}^2 c!^2\, Var\, U_{g_c}^{(m,n)} &= Var\ Per_{g_c}\, \mathbb{X}\, (c|c) \\ &= \sum_{1 \le j_1 < \ldots < j_c \le n} Var\ (Per_{g_c}\, \mathbb{X}\, (c|j_1, \ldots, j_c))\end{aligned}$$

since, by independence of columns of $\mathbb{X}^{(\infty)}$, we have (cf. also the proof of part *(ii)* of the Theorem 2.2.2)

$$Cov\ [Per_{g_c}\, \mathbb{X}\, (c|j_1, \ldots, j_c), Per_{g_c}\, \mathbb{X}\, (c|l_1, \ldots, l_c)] = 0$$

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

Since the columns of $\mathbb{X}^{(\infty)}$ are identically distributed, we obtain

$$\binom{n}{c} \binom{m}{c}^2 c!^2\, Var\, U_{g_c}^{(m,n)} = Var\ (Per_{g_c}\, \mathbb{X}\, (c|1, \ldots, c)).$$

Let us note that the number of pairs of $c \times c$ submatrices of $\mathbb{X}^{(\infty)}$ 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 $\mathbb{X}^{(\infty)}$ are identically distributed). Hence, for given c, the above right-hand side can be written as

$$\binom{m}{c} \sum_{k=\max(0,2c-m)}^{c} \binom{c}{k}\binom{m-c}{c-k}$$
$$\times\, Cov\, [Per_{g_c}\, \mathbb{X}\,(1,\ldots,k,i_{k+1},\ldots,i_c|1,\ldots,c),$$
$$Per_{g_c}\, \mathbb{X}\,(1,\ldots,k,l_{k+1},\ldots,l_c|1,\ldots,c)]$$

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

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

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 $\mathbb{X}^{(\infty)}$ it follows that the partial covariances having exactly l $(0\le l\le k)$ elements in common are the same and equal to

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

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 $Per_{g_c}\,\mathbb{Y}[k,c]$, where $\mathbb{Y}[k,c]$ is a $k\times c$ matrix, having exactly l elements in common with the diagonal entries $y_{1,1},\ldots,y_{k,k}$. First, we fix l factors in $\binom{k}{l}$ ways. If we assume that $y_{1,1},\ldots,y_{l,l}$ are fixed, then the remaining factors, in the c-tuples we are looking for, have to be of the form $y_{l+1,j_{l+1}},\ldots,y_{k,j_k}$, where $j_d\ne d$, $d=l+1,\ldots,k$. 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,

$$\mathcal{R}_l(k,c) = \sum_{j=0}^{k-l}(-1)^{j}\binom{k-l}{j}\binom{c-l-j}{k-l-j}(k-l-j)!. \tag{2.11}$$

Consequently, the number of pairs of c-tuples in $Per\,\mathbb{Y}[k,c]$ 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

$$\begin{aligned}&Cov\,[Per_{g_c}\,\mathbb{X}\,(1,\ldots,k,i_{k+1},\ldots,i_c|1,\ldots,c),\\ &\quad Per_{g_c}\,\mathbb{X}\,(1,\ldots,k,l_{k+1},\ldots,l_c|1,\ldots,c)]\\ &= (c-k)!^2\sum_{l=0}^{k}\binom{c}{k}k!\binom{k}{l}\rho_{c,l}\mathcal{R}_l(k,c)\ .\end{aligned}$$

Now, returning to the formula for the variance of $U_{g_c}^{(m,n)}$ we obtain by (2.11)

$$\begin{aligned}&\binom{n}{c}\binom{m}{c}Var\,U_{g_c}^{(m,n)}\\ &= \frac{1}{c!^2}\sum_{k=\max(0,2c-m)}^{c}\binom{c}{k}^2\binom{m-c}{c-k}k!(c-k)!^2\sum_{l=0}^{k}\binom{k}{l}\rho_{c,l}\,\mathcal{R}_l(k,c)\\ &= \sum_{k=\max(0,2c-m)}^{c}\binom{m-c}{c-k}\sum_{d=0}^{k}\binom{c-d}{k-d}\frac{1}{d!}D(c,d),\end{aligned} \tag{2.12}$$

since

$$\begin{aligned}\sum_{l=0}^{k}\binom{k}{l}\rho_{c,l}\mathcal{R}_l(k,c) &= \sum_{l=0}^{k}\binom{k}{l}\rho_{c,l}\sum_{d=l}^{k}(-1)^{d-l}\binom{k-l}{d-l}\binom{c-d}{k-d}(k-d)!\\ &= \sum_{d=0}^{k}\binom{c-d}{k-d}(k-d)!\binom{k}{d}\sum_{l=0}^{d}\binom{d}{l}(-1)^{d-l}\rho_{c,l}\end{aligned}$$

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

Observe that we can further simplify the expression (2.12), since

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

where the last equality follows by applying the hypergeometric summation rule for the inner sum. Thus, we can rewrite (2.12) as

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

which along with (2.10) completes the proof. □

Example 2.3.1 (Variance of a random permanent). Consider again the kernel function as in Example 2.2.1. Let as before $\mu \neq 0$ be the common mean of $X_{i,j}$ and additionally let $0 < \sigma^2 < \infty$ be the common variance. By $\gamma = \sigma/\mu$ we denote the coefficient of variation and by $\rho > 0$ the correlation between any two components from the same column. Then,

$$\rho_{c,s} = \gamma^{2c} \rho^{c-s}$$

and thus

$$D(c,d) = \gamma^{2c} \rho^{c-d} (1-\rho)^d.$$

As shown before in Example 2.2.1 the canonical function $U_{g_c}^{(m,n)}$ of the H-decomposition is given by (2.3). From Theorem 2.3.1 we have now that

$$Var\, U_c^{(m,n)} = \binom{n}{c}^{-1} \binom{m}{c}^{-1} \gamma^{2c} \sum_{d=0}^{c} \binom{m-d}{c-d} \frac{\rho^{c-d}(1-\rho)^d}{d!},$$

and hence by the orthogonality of $U_{g_c}^{(m,n)}$ for $c = 1, 2, \ldots, m$, it follows that

$$Var \left(\frac{Per\, \mathbb{X}^{(m,n)}}{\binom{n}{m} m!\, \mu^m} \right) = \sum_{c=1}^{m} \frac{\binom{m}{c} \gamma^{2c}}{\binom{n}{c}} \sum_{d=0}^{c} \frac{1}{d!} \binom{m-d}{c-d} (1-\rho)^d \rho^{c-d}.$$

An important special case of Theorem 2.3.1 occurs when the rows of $\mathbb{X}^{(\infty)}$ are also independent. Then the matrix $\mathbb{X}^{(\infty)}$ has all independent and identically distributed entries and

$$D(c,d) = 0 \quad \text{for} \quad d < c.$$

Since $\rho(c,c) = Var\, g_c$, we see that for each canonical component $U^{(m,n)}_{g_c}$ we must have

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

which entails (by orthogonality) that the formula for the variance of a P-statistic degenerate of degree $r-1$ in case when $\mathbb{X}^{(\infty)}$ has all independent and identically distributed entries is

$$Var\left(\frac{Per_h\,\mathbb{X}^{(m,n)}}{\binom{n}{m}\,m!}\right) = \sum_{c=r}^{m}\frac{\binom{m}{c}}{\binom{n}{c}}\frac{Var\, g_c}{c!}. \tag{2.13}$$

Another important special case is that of a U-statistic which corresponds to $\mathbb{X}^{(\infty)} = \underline{\mathrm{X}}$ i.e., the one dimensional projection matrix. Then in Theorem 2.3.1 we have

$$D(c,s) = 0 \quad \text{for} \quad s > 0$$

and obtain the following variance formula for a U-statistic of degree of degeneration $r-1$ as defined in Definition 1.4.1 with $l = n$ and $k = m$

$$Var\left(\frac{Per_h\,\underline{\mathrm{X}}}{\binom{n}{m}\,m!}\right) = Var\left(U^{(m)}_n(h)\right) = \sum_{c=r}^{m}\frac{\binom{m}{c}^2}{\binom{n}{c}}\,Var\, g_c. \tag{2.14}$$

We conclude this chapter with the following example illustrating one possible use of Theorem 2.3.1. The example could be considered as a warm-up before the material of the next chapter. Recall that for a sequence of random variables X_n and a sequence of real numbers a_n the notation $X_n = o_p(a_n)$ stands for $\forall_{\varepsilon>0} P(|X_n| > a_n\varepsilon) \to 0$ as $n \to \infty$. In this notation we may state the following simple approximation theorem for non-degenerate P-statistics.

Example 2.3.2 (Approximation theorem for P-statistics). Assume (A1-A2) for matrix $\mathbb{X}^{(\infty)}$. Suppose $m/\sqrt{n} \to 0$ as $n \to \infty$, $Eh^2 < \infty$ and $Var\,(g_1) \neq 0$ where g_1 does not depend on m, n, as well as for some non-random positive constants M_1 and M_2 which do not depend upon m, n, we have for $c = 1, 2, 3, \ldots$

$$\forall_{d\le c}\; D(c,d) \le M_1^c M_2^d < \infty.$$

Then

$$\frac{Per_h\,\mathbb{X}^{(m,n)}}{\binom{n}{m}\,m!} - Eh = \frac{1}{n}\sum_{ij} g_1(X_{ij}) + o_p\left(\frac{m}{\sqrt{n}}\right). \tag{2.15}$$

The formula (2.15) is a simple consequence of the properties of P-statistics summarized by Theorems 2.2.1, 2.2.2, and 2.3.1 in this chapter. Note that for non-degenerate P-statistics $(r = 1)$

$$U^{(m,n)}_{g_1} = \binom{m}{1}^{-1}\binom{n}{1}^{-1} Per_{g_1}\mathbb{X}\,(1|1) = \frac{1}{mn}\sum_{ij} g_1(x_{ij})$$

and thus by Theorem 2.2.1

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

where the two summands are orthogonal by Theorem 2.2.2. 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 $m^2/n \to 0$. Under the assumptions of the example we have

$$Var\,U_{g_c}^{(m,n)} = \binom{n}{c}^{-1}\binom{m}{c}^{-1}\sum_{d=0}^{c}\binom{m-d}{c-d}\frac{D(c,d)}{d!} \le \binom{n}{c}^{-1}\exp(M_2)\,M_1^c$$

since $\binom{m-d}{c-d} \le \binom{m}{c}$ for $1 \le d \le c \le m$. Therefore by Theorem 2.3.1

$$\begin{aligned}\frac{n}{m^2} Var\,R_{m,n} &= \frac{n}{m^2}\sum_{c=2}^{m}\binom{m}{c}^2 Var\,U_{g_c}^{(m,n)} \le \frac{n}{m^2}\exp(M_2)\sum_{c=2}^{m}\binom{m}{c}^2\binom{n}{c}^{-1}M_1^c \\ &\le \frac{n}{m^2}\exp(M_2)\sum_{c=2}^{m}\left(M_1\frac{m^2}{n}\right)^c\frac{1}{c!}\,,\end{aligned}$$

where the last inequality follows by

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

Consequently,

$$\frac{n}{m^2} Var\,R_{m,n} \le M_1^2\exp(M_2)\,\frac{m^2}{n}\sum_{c=0}^{m}\left(\frac{M_1 m^2}{n}\right)^c\frac{1}{c!} \to 0,$$

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

In particular, continuing the case of random permanent (see Examples 2.2.1 and 2.3.1), we take $h(y_1,\ldots,y_m) = \prod_{i=1}^m (y_i/\mu)$ such that $Eh = \mu \neq 0$. Then $[g_1(X_{i,j})] = \mu^{-1}[X_{i,j} - \mu]$. Thus the above formula provides the approximation for the classical permanent function since then we may take (see the formula for $D(c,d)$ in Example 2.3.1) $M_1 = \rho\gamma^2$ and $M_2 = (1-\rho)/\rho$.

2.4 Bibliographic Details

The standard general references on the martingale theory are Chow and Teicher (1978); Billingsley (1999) and for continuous time processes e.g., Ethier and Kurtz (1986). The general martingale decomposition for a U-statistic, which became latter known as the H-decomposition, was presented for the first time in Hoeffding (1961). The notion of a P-statistic and its general decomposition comes from the papers of Rempała and Wesołowski (2003) and Rempała (2001). The idea is related to a general canonical decomposition of a symmetric functional as described in Dynkin and Mandelbaum (1983) or in Lee (1990, chapter 4). Most of the material of this chapter discussed in Sections 2 and 3 comes from the work of Rempała (2001), which, though appeared earlier, extended to P-statistics the ideas developed for random permanents in Rempała and Wesołowski (2002b) and Rempała and Wesołowski (2002c).

http://www.springer.com/978-0-387-75145-0

Symmetric Functionals on Random Matrices and
Random Matchings Problems
Rempala, G.; Wesolowski, J.
2008, XIV, 184 p., Hardcover
ISBN: 978-0-387-75145-0