



Central limit theorem for multiple integrals with respect to the empirical process

Hélène Boistard*, Eustasio del Barrio¹

Universidad de Valladolid, Spain

ARTICLE INFO

Article history:

Received 9 July 2007

Received in revised form 20 May 2008

Accepted 23 July 2008

Available online 6 August 2008

ABSTRACT

This paper deals with weak convergence of multiple integrals with respect to the empirical process. A stochastic integral with respect to the Brownian bridge is introduced to express the limit in a unified way for the degenerate and non-degenerate cases.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

In this paper we consider the limiting distribution of multiple stochastic integrals with respect to the empirical process, namely, of objects of type

$$J_{n,m}(h) = \int' h(x_1, \dots, x_m) d\mathbb{G}_n(x_1) \cdots d\mathbb{G}_n(x_m). \tag{1}$$

Here $\mathbb{G}_n = \sqrt{n}(\mathbb{P}_n - P)$ denotes the normalized empirical measure based on the i.i.d. random variables X_1, \dots, X_n with common distribution P on some measurable space (X, \mathcal{X}) and h is a real valued, symmetric function of m variables in $L_2(P^m)$. From now on, any such h will be called a kernel. \int' denotes an integral in which diagonals $x_r = x_s, 1 \leq r < s \leq m$ are omitted from the domain of integration.

These multiple stochastic integrals appear naturally, for instance, as higher order terms in the delta method for some non-parametric maximum likelihood problems, see, e.g., Major (2006, 2007) for further motivation and examples. They are closely connected to U -statistics, namely, statistics of type

$$U_n(h) = \frac{1}{\binom{n}{m}} \sum_{1 \leq i_1 < \dots < i_m \leq n} h(X_{i_1}, \dots, X_{i_m}). \tag{2}$$

In fact, if the kernel h is completely degenerate then $J_{n,m}(h) = \frac{n!}{(n-m)!n^{m/2}} U_n(h)$. Hence, the classical central limit theorem (CLT) for completely degenerate U -statistics (see, e.g., Arcones and Giné (1992)) provides the limiting distribution of $J_{n,m}(h)$. This limit is the multiple Wiener integral, that we denote by $I_m(h)$.

In the case of a non-degenerate kernel h , the U -statistic $U_n(h)$ and the multiple integral $J_{n,m}(h)$ are not equal and do not have the same asymptotic behaviour. On the one hand, the CLT for $U_n(h)$ can be obtained via the Hoeffding decomposition from the CLT for completely degenerate U -statistics. A note in Arcones and Giné (1992) adds that the exact form of the limit can be quite complicated when the degeneracy order of the kernel is greater than 1.

* Corresponding address: Laboratoire GREMAQ, Université de Toulouse 1, Manufacture des Tabacs, 21 allée de Brienne, 31000 Toulouse, France.
E-mail addresses: helene@boistard.fr (H. Boistard), tasio@eio.uva.es (E. del Barrio).

¹ Departamento E.I.O., Facultad de Ciencias, C/Prado de la Magdalena s/n, 47005 Valladolid, Spain.

On the other hand, the convergence in distribution of $J_{n,m}(h)$ can be deduced from recent work by Major (2007, 2006). However, the limit distribution is not described with precision. The results there (Theorem 9.4 in Major (2007) or Proposition 3 in Major (2006)) express $J_{n,m}(h)$ as the sum of empirical integrals of completely degenerate kernels: the conclusion is

$$J_{n,m}(h) = \sum_{k=0}^m C(n, k) J_{n,m-k}(\pi_{m-k}h), \tag{3}$$

where the completely degenerate kernels $\pi_{m-k}h$ are defined in (9) and with coefficients $C(n, k)$ such that for all k ,

$$C(n, k) \rightarrow C_k, n \rightarrow \infty.$$

From this it is deduced that $J_{n,m}(h)$ converges in distribution to

$$\sum_{k=0}^m C_k I_{m-k}(\pi_{m-k}h).$$

We need to know with precision the coefficients C_k to be able to know the limit distribution. Our Theorem 3.2 provides a decomposition of the same type as (3), for symmetric kernels, and proves the convergence of coefficients to explicit values that can be expressed in terms of the Hermite polynomials. The general CLT for multiple integrals (Theorem 3.3) is deduced from this convergence.

The limiting distribution of the multiple integral $J_{n,m}(h)$ is a multiple Wiener integral in the case of a completely degenerate kernel. When studying the non-degenerate case, we find that the limit could be expressed in a simple way if we could define formally an integral with respect to the Brownian bridge. The definition is given in Definition 2.2. The main property of that integral is Proposition 2.3, which is an analogue to the Hoeffding decomposition.

This article is organized as follows. Section 2 defines multiple integrals with respect to the Brownian bridge. In Section 3, we prove the central limit theorem for the multiple integral with respect to the empirical process, giving an explicit form for the limit.

2. Multiple Wiener integrals

When studying the asymptotic behaviour of multiple integrals with respect to the empirical measure, we face the difficulty caused by the fact that the limit of the empirical measure is not the white noise with respect to which the multiple Wiener integrals are defined. Let us pay attention to the case in which P is the uniform distribution on $(0, 1)$. Then $\{\mathbb{G}_n(x)\}_{x \in (0,1)}$ converges in distribution (in different topologies) to a Brownian bridge, $\{B(x)\}_{x \in (0,1)}$. This Brownian bridge can be represented as $B(x) = W(x) - xW(1)$, where W is a standard Brownian motion. The convergence of \mathbb{G}_n to B seems to suggest that for suitable functions $h(x_1, \dots, x_m)$ the following may hold:

$$\int_{[0,1]^m} h(x_1, \dots, x_m) d\mathbb{G}_n(x_1) \cdots d\mathbb{G}_n(x_m) \xrightarrow{w} \int_{[0,1]^m} h(x_1, \dots, x_m) dB(x_1) \cdots dB(x_m). \tag{4}$$

Before we give a meaning to the multiple integral with respect to the Brownian bridge, let us recall some facts about Hermite polynomials and multiple Wiener integral. We call $H_n(x)$ the normalized Hermite polynomial of degree n . The first two of these polynomials are $H_0(x) = 1, H_1(x) = x$ and the others can be calculated from the following recursion

$$(n + 1)H_{n+1}(x) = xH_n(x) - H_{n-1}(x). \tag{5}$$

Moreover, these polynomials satisfy

$$H'_n(x) = H_{n-1}(x). \tag{6}$$

A proof of these facts can be found e.g. in Nualart (1995).

Multiple Wiener integrals for square integrable kernels h of m variables will be denoted $I_m(h)$. We refer to Nualart (1995) for the definition and properties. We include here a technical result about multiple integrals which is relevant to this paper.

Lemma 2.1. *If $f \in L_2(P^p)$ and is completely degenerate, then*

$$I_{p+q}(f \otimes 1^{\otimes q}) = I_p(f)q!H_q(I_1(1)).$$

Proof. It follows from Proposition 1.1.3 and Theorem 1.1.2 in Nualart (1995). We need to introduce the following notation as in Nualart (1995), p. 10. For two symmetric functions $f \in L_2(P^p), g \in L_2(P^q)$, and $1 \leq r \leq \min(p, q)$, the contraction of r indices of f and g is the function defined by

$$f \otimes_r g(t_1, \dots, t_{p+q-2r}) = \int_{[0,1]^r} f(t_1, \dots, t_{p-r}, s)g(t_{p+1}, \dots, t_{p+q-r}, s)P^r(ds).$$

The contraction $f \otimes_0 g$ is defined as $f \otimes g$.

From Theorem 1.1.2 in Nualart (1995), it follows that

$$I_q(1^{\otimes q}) = q!H_q(I_1(1)). \tag{7}$$

On the other hand, from Proposition 1.1.3, we obtain

$$I_p(f)I_q(1^{\otimes q}) = \sum_{r=0}^{p \wedge q} r! \binom{p}{r} \binom{q}{r} I_{p+q-2r}(f \otimes_r 1^{\otimes q}). \tag{8}$$

But since f is completely degenerate, $f \otimes_r 1^{\otimes q} = 0$ for all $r \geq 1$. Indeed,

$$\begin{aligned} f \otimes_r 1^{\otimes q}(t_1, \dots, t_{p+q-2r}) &= \int_{[0,1]^r} f(t_1, \dots, t_{p-r}, s) 1^{\otimes q}(t_{p+1}, \dots, t_{p+q-r}, s) P^r(ds) \\ &= 0 \quad \text{for } r \geq 1. \end{aligned}$$

Hence, the right-hand side in (8) equals $I_{p+q}(f \otimes 1^{\otimes q})$, which, combined to (7), yields the result. \square

We can now give a meaning to the right-hand term of equality (4). We can use the representation $B(x) = W(x) - xW(1)$ and write (assuming that h is symmetric in its m variables for the sake of simplicity) formally $dB(x) = dW(x) - W(1)dx$ and hence

$$\begin{aligned} \int_{[0,1]^m} h(x_1, \dots, x_m) dB(x_1) \cdots dB(x_m) &= \int_{[0,1]^m} h(x_1, \dots, x_m) dW(x_1) \cdots dW(x_m) \\ &\quad - m \int_{[0,1]^{m-1}} \rho_{m-1} h(x_1, \dots, x_{m-1}) dW(x_1) \cdots dW(x_{m-1}) W(1) \\ &\quad + \binom{m}{2} \int_{[0,1]^{m-2}} \rho_{m-2} h(x_1, \dots, x_{m-2}) dW(x_1) \cdots dW(x_{m-2}) W(1)^2 \\ &\quad + \cdots + (-1)^m \rho_0 h W(1)^m, \end{aligned}$$

where $\rho_j h$ is the function of j variables defined as $\rho_j h(x_1, \dots, x_j) = \delta_{x_1} \times \cdots \times \delta_{x_j} \times P^{m-j} h$.

This formal calculus motivates our definition of *multiple integrals with respect to the Brownian bridge*.

Definition 2.2. Given a kernel $h \in L_2(P^m)$ we define

$$J_m(h) = \sum_{k=0}^m (-1)^k \binom{m}{k} (I_{m-k}(\rho_{m-k} h)) I_1(1)^k.$$

Recall that h is completely degenerate if $\rho_k h = 0$ for $k < m$. In that case, $J_m(h) = I_m(h)$.

We will use the notation $\pi_j h$ for the Hoeffding projection of order j of h , i.e.,

$$\pi_j h(x_1, \dots, x_j) = (\delta_{x_1} - P) \times \cdots \times (\delta_{x_j} - P) \times P^{m-j} h, \quad j = 0, \dots, m. \tag{9}$$

Then $\pi_0 h = P^m h$ and $\pi_j h$ is completely degenerate if $j \geq 1$. With this notation we can prove the following result, which is an analogue to the Hoeffding decomposition for integrals with respect to the Brownian bridge.

Proposition 2.3. For any kernel $h \in L_2(P^m)$

$$J_m(h) = \sum_{k=0}^m \frac{m!}{(m-k)!} H_k(0) I_{m-k}(\pi_{m-k} h),$$

where H_k is the k th Hermite polynomial.

Proof. Writing $h(x_1, \dots, x_m) = \delta_{x_1} \times \cdots \times \delta_{x_m} h = (\delta_{x_1} - P + P) \times \cdots \times (\delta_{x_m} - P + P) h$ we obtain

$$\begin{aligned} h(x_1, \dots, x_m) &= \pi_m h(x_1, \dots, x_m) + \sum_{\{i_1, \dots, i_{m-1}\} \subset \{1, \dots, m\}} \pi_{m-1} h(x_{i_1}, \dots, x_{i_{m-1}}) \\ &\quad + \cdots + \sum_{\{i_1, \dots, i_{m-k}\} \subset \{1, \dots, m\}} \pi_{m-k} h(x_{i_1}, \dots, x_{i_{m-k}}) + \cdots + \pi_0 h. \end{aligned}$$

Let us introduce the functions $\phi_j^{m-k} h$ defined for $1 \leq j \leq m$ as

$$\phi_j^{m-k} h : (x_1, \dots, x_j) \mapsto \begin{cases} \sum_{\{i_1, \dots, i_{m-k}\} \subset \{1, \dots, j\}} \pi_{m-k} h(x_{i_1}, \dots, x_{i_{m-k}}) & \text{for } 0 \leq k \leq m-1, \\ \pi_0 h & \text{for } k = m. \end{cases}$$

With this notation, we have: $h = \sum_{k=0}^m \phi_m^{m-k} h$. Now, we observe that

$$\rho_j \phi_m^{m-k} h = \begin{cases} 0 & \text{if } j = 0, \dots, m - k - 1, \\ \phi_j^{m-k} h & \text{if } j \geq m - k. \end{cases}$$

From this and from Definition 2.2 we obtain

$$J_m(\phi_m^{m-k} h) = \sum_{j=0}^k (-1)^j \binom{m}{j} I_{m-j}(\phi_{m-j}^{m-k} h) I_1(1)^j. \tag{10}$$

By linearity of the stochastic integral and by symmetry we have

$$I_{m-j}(\phi_{m-j}^{m-k} h) = \binom{m-j}{m-k} I_{m-j}(\pi_{m-k} h \otimes 1^{\otimes k-j}). \tag{11}$$

The binomial formula, the degeneracy of $\pi_{m-k} h$ and the properties of the Hermite polynomials (cf. Lemma 2.1) yield $I_{m-j}(\pi_{m-k} h \otimes 1^{\otimes k-j}) = I_{m-k}(\pi_{m-k} h)(k-j)! H_{k-j}(I_1(1))$. Combining this equality with (10) and (11) we conclude that

$$J_m(\phi_m^{m-k} h) = \binom{m}{k} I_{m-k}(\pi_{m-k} h) \times \left(\sum_{j=0}^k \binom{k}{j} (-1)^j (k-j)! H_{k-j}(I_1(1)) I_1(1)^j \right).$$

The proof can be completed with the following simple lemma, which is a particular case of Lemma 4.1 in Arcones and Giné (1993). □

Lemma 2.4.

$$\sum_{j=0}^k \binom{k}{j} (-1)^j (k-j)! H_{k-j}(x) x^j = k! H_k(0).$$

3. Central limit theorem for multiple empirical integrals

The multiple stochastic integrals introduced in Section 2 will allow us to write a unified CLT for multiple stochastic integrals with respect to the empirical process. The key point here is that we find a precise Hoeffding type decomposition, with convenient explicit expressions for the coefficients. Then, we give the general CLT, valid for both degenerate and non-degenerate kernels. The limit has exactly the same form in both cases. Here, we can see the advantage of considering multiple empirical integrals instead of U -statistics: the homogeneity of results without having to take into account the degeneracy of the kernel.

Our result will follow from the CLT for completely degenerate U -statistics and from Theorem 3.2. First, we write in our notation the existing result for degenerate kernels. A proof can be found e.g. in Rubin and Vitale (1980) and Arcones and Giné (1992) or de la Peña and Giné (1999). In what follows, \xrightarrow{w} denotes the convergence in distribution.

Theorem 3.1. *Let (X_1, \dots, X_n) be an i.i.d. sample with non-atomic distribution P and h_1, \dots, h_k completely degenerate kernels in m_1, \dots, m_k variables. Then*

$$(J_{n,m_1}(h_1), \dots, J_{n,m_k}(h_k)) \xrightarrow{w} (I_{m_1}(h_1), \dots, I_{m_k}(h_k)).$$

In order to deduce the limit distribution in the general case from the degenerate case, we need the decomposition below, which is an analogue of the Hoeffding decomposition. This result improves Proposition 3 of Major (2006) in the sense that it provides an explicit expression for the coefficients in decomposition (3).

Theorem 3.2. *Let (X_1, \dots, X_n) be an i.i.d. sample with distribution P and h a kernel of m variables. We assume that P does not have atoms. We have the following decomposition of $J_{n,m}(h)$ as a sum of multiple integrals based on degenerate kernels:*

$$J_{n,m}(h) = \sum_{j=0}^m K_{n,j,m} J_{n,j}(\tau_j h) \tag{12}$$

where the coefficients are, for $j \leq m$,

$$K_{n,j,m} = n^{-(m-j)/2} \sum_{k=0}^{m-j} (-1)^k n^k \binom{m}{k} \binom{m-k}{m-k-j} \frac{(n-j)!}{(n-m+k)!}.$$

Moreover, these coefficients satisfy the following convergence:

$$\lim_{n \rightarrow \infty} K_{n,j,m} = \frac{m!}{j!} H_{m-j}(0),$$

where for all $k \geq 0$, H_k is the Hermite polynomial of order k .

Before we prove the theorem, let us introduce a notation to be used in the following: we will denote by \tilde{h} the kernel

$$\tilde{h}(x_1, \dots, x_m) = \begin{cases} h(x_1, \dots, x_m) & \text{if } x_i \neq x_j, \text{ for } 1 \leq i < j \leq m \\ 0 & \text{otherwise.} \end{cases}$$

As a consequence, $J_{n,m}(h)$ is equal to $J_{n,m}(\tilde{h})$ and we can integrate \tilde{h} also on the diagonal, since it vanishes there.

Proof. We first prove the decomposition (12). We are going to use the following notation: for coefficients c_{i_1, \dots, i_p} , we define

$$\begin{aligned} \sum_{\Delta_p} c_{i_1, \dots, i_p} &= \sum_{1 \leq i_1, \dots, i_p \leq n \text{ all different}} c_{i_1, \dots, i_p} \\ J_{n,m}(h) &= n^{-m/2} \sum_{1 \leq i_1, \dots, i_m \leq n} \int \tilde{h}(y_1, \dots, y_m) d(\delta_{X_{i_1}} - P)(y_1) \dots d(\delta_{X_{i_m}} - P)(y_m) \\ &= n^{-m/2} \sum_{k=0}^m (-1)^k \binom{m}{k} \sum_{1 \leq i_1, \dots, i_{m-k} \leq n} \int \tilde{h}(X_{i_1}, \dots, X_{i_{m-k}}, y_1, \dots, y_k) dP(y_1) \dots dP(y_k) \\ &= n^{-m/2} \sum_{k=0}^m (-1)^k n^k \binom{m}{k} \sum_{\Delta_{m-k}} \int \tilde{h}(X_{i_1}, \dots, X_{i_{m-k}}, y_1, \dots, y_k) dP(y_1) \dots dP(y_k) \\ &= n^{-m/2} \sum_{k=0}^m (-1)^k n^k \binom{m}{k} \sum_{\Delta_{m-k}} \rho_{m-k} h(X_{i_1}, \dots, X_{i_{m-k}}) \end{aligned}$$

because the measure P does not have atoms. The variable

$$\frac{1}{(m-k)! \binom{n}{m-k}} \sum_{\Delta_{m-k}} \rho_{m-k} h(X_{i_1}, \dots, X_{i_{m-k}})$$

is a U -statistic and the Hoeffding decomposition (e.g., (3.5.1), p. 137 in de la Peña and Giné (1999)) gives (with the notation $U_n^{(l)}$ as the average of the values of a function of l variables at the points $(X_{i_1}, \dots, X_{i_l})$, $(i_1, \dots, i_l) \in \Delta_l$),

$$\begin{aligned} \frac{1}{(m-k)! \binom{n}{m-k}} \sum_{\Delta_{m-k}} \rho_{m-k} h(X_{i_1}, \dots, X_{i_{m-k}}) &= \sum_{j=0}^{m-k} \binom{m-k}{j} U_n^{(j)}(\pi_j \rho_{m-k} h) \\ &= \sum_{j=0}^{m-k} \binom{m-k}{j} U_n^{(j)}(\pi_j h) \end{aligned}$$

that is,

$$\sum_{\Delta_{m-k}} \rho_{m-k} h(X_{i_1}, \dots, X_{i_{m-k}}) = \sum_{j=0}^{m-k} \binom{m-k}{j} \frac{(n-j)!}{(n-m+k)!} \sum_{\Delta_j} \pi_j h(X_{i_1}, \dots, X_{i_j}).$$

It then follows from this, the previous set of identities and the definition of $K_{n,j,m}$, that

$$J_{n,m}(h) = \sum_{j=0}^m K_{n,j,m} J_{n,j}(\pi_j h).$$

Now, it remains to prove that for all $j \leq m$, $K_{n,j,m} \rightarrow \frac{m!}{j!} H_{m-j}(0)$ when $n \rightarrow \infty$. We are going to prove first this fact for $j = 0$ and whatever m , and then by induction on j we will prove the formula for any j and m . The convergence of $K_{n,0,m}$ can be proved by induction on m . For $m = 0$, it is easy because $K_{n,0,0} = 1 = H_0(0)$. Now, we assume that the convergence of $K_{n,0,p}$ to $p!H_p(0)$ is satisfied for all $p \leq m$. In order to prove that it holds also for $p = m + 1$, we use the following expression:

$$K_{n,0,m+1} = -n^{-1/2} m K_{n,0,m} - m K_{n,0,m-1}. \tag{13}$$

The proof, which is obtained via some technical computations, is postponed to the Appendix. Hence, by the induction hypothesis,

$$K_{n,0,m+1} \rightarrow -m(m-1)!H_{m-1}(0) = -m!H_{m-1}(0).$$

But this is the expected limit, since by property (5) of the Hermite polynomials,

$$(m + 1)!H_{m+1}(0) = m!(m + 1)H_{m+1}(0) = m!(0H_m(0) - H_{m-1}(0)) = -m!H_{m-1}(0).$$

Hence the convergence is proved for $j = 0$ and any m .

Let us now prove by induction on j that $K_{n,j,m}$ converges to the expected limit. We assume that the following property at rank j is satisfied: for any m such that $j \leq m$, $K_{n,j,m} \rightarrow \frac{m!}{j!}H_{m-j}(0)$. This property is obviously satisfied for $j = 0$ by what we have just proved. Now let us prove the property at rank $j + 1$. So, let m be such that $j + 1 \leq m$. Some technical but elementary computations lead to the following recursive expression for $K_{n,j+1,m}$:

$$K_{n,j+1,m} = n^{1/2} \frac{m-j}{n-j} K_{n,j,m} + \frac{n}{n-j} \frac{m}{j+1} K_{n,j,m-1}. \tag{14}$$

The proof of this formula is given in the Appendix. Now, we use the induction hypothesis. $n^{1/2} \frac{m-j}{n-j}$ tends to 0 and $K_{n,j,m}$ converges by the induction hypothesis, so that $n^{1/2} \frac{m-j}{n-j} K_{n,j,m} \rightarrow 0$. On the other hand, $j \leq m - 1$, so that, employing the induction hypothesis:

$$\frac{n}{n-j} \frac{m}{j+1} K_{n,j,m-1} \rightarrow \frac{m}{j+1} \frac{(m-1)!}{j!} H_{m-1-j}(0) = \frac{m!}{(j+1)!} H_{m-1-j}(0)$$

and this proves that $K_{n,j+1,m} \rightarrow \frac{m!}{(j+1)!} H_{m-j-1}(0)$, as $n \rightarrow \infty$. \square

Theorem 3.3. Let (X_1, \dots, X_n) be an i.i.d. sample with non-atomic distribution P and let h_1, \dots, h_k be square integrable kernels in m_1, \dots, m_k variables. Then

$$(J_{n,m_1}(h_1), \dots, J_{n,m_k}(h_k)) \xrightarrow{w} (J_{m_1}(h_1), \dots, J_{m_k}(h_k)).$$

Proof. Without loss of generality, we take $k = 1$. By Theorems 3.1 and 3.2 which give the CLT for the multiple integral of degenerate kernels with respect to the empirical process, we have

$$J_{n,m}(h) = \sum_{j=0}^m K_{n,j,m} J_{n,j}(\pi_j h) \xrightarrow{w} \sum_{j=0}^m \frac{m!}{j!} H_{m-j}(0) I_j(\pi_j h) = J_m(h),$$

where we use Proposition 2.3 in the last identity. \square

Remark 3.4. This theorem gives a neat version of Theorem 10.4' in Major (2007), providing an explicit expression for the constants in the limit. The reader should also compare to the CLT for general U -statistics (see Section 2A in Arcones and Giné (1992) or Remark 4.2.5 in de la Peña and Giné (1999)) as well as to Theorem 1 in Dynkin and Mandelbaum (1983) where limit theorems for symmetric statistics of possibly infinite order are derived. The picture is much clearer and easier for stochastic integrals with respect to the empirical measure.

Acknowledgements

The authors want to thank an anonymous referee for her/his careful reading of the manuscript and suggestions that have helped to improve the final version of this manuscript. The authors have been partially supported by the Spanish Ministerio de Educación y Ciencia and FEDER, grant MTM2005-08519-C02-01,02 and the Consejería de Educación y Cultura de la Junta de Castilla y León, grant PAPIJCL VA102A06.

Appendix. Technical proofs

Proof of (13). Isolating the first and last terms that compose $K_{n,0,m+1}$, we have:

$$\begin{aligned} K_{n,0,m+1} &= n^{-(m+1)/2} \sum_{k=0}^{m+1} (-1)^k n^k \binom{m+1}{k} \frac{n!}{(n-m+k-1)!} \\ &= n^{-(m+1)/2} \frac{n!}{(n-m-1)!} + n^{-(m+1)/2} (-1)^{m+1} n^{m+1} \end{aligned} \tag{A.1}$$

$$+ n^{-(m+1)/2} \sum_{k=1}^m (-1)^k n^k \left(\binom{m}{k} + \binom{m}{k-1} \right) \frac{n!}{(n-m+k-1)!}, \tag{A.2}$$

by the binomial formula. The sum (A.2) over k is decomposed into two sums. The first one is $n^{-(m+1)/2} \sum_{k=1}^m (-1)^k n^k \binom{m}{k} \frac{n!}{(n-m+k-1)!}$, which is equal to

$$\begin{aligned} & n^{-(m+1)/2} \sum_{k=1}^m (-1)^k n^k \binom{m}{k} \left((n-m) \frac{n!}{(n-m+k)!} + k \frac{n!}{(n-m+k)!} \right) \\ &= n^{-1/2} (n-m) n^{-m/2} \sum_{k=0}^m (-1)^k n^k \binom{m}{k} \frac{n!}{(n-m+k)!} - (n-m) n^{-(m+1)/2} \frac{n!}{(n-m)!} \\ & \quad + n^{-(m+1)/2} \sum_{j=0}^{m-1} (-1)^{j+1} n^{j+1} \frac{m!}{j!(m-1-j)!} \frac{n!}{(n-m+j+1)!} \\ &= n^{-1/2} (n-m) K_{n,0,m} - n^{-(m+1)/2} \frac{n!}{(n-m-1)!} - n^{-(m+1)/2} n m \sum_{j=0}^{m-1} (-1)^j n^j \binom{m-1}{j} \frac{n!}{(n-(m-1)+j)!} \\ &= n^{-1/2} (n-m) K_{n,0,m} - n^{-(m+1)/2} \frac{n!}{(n-m-1)!} - m K_{n,0,m-1}. \end{aligned} \tag{A.3}$$

The second sum is $n^{-(m+1)/2} \sum_{k=1}^m (-1)^k n^k \binom{m}{k-1} \frac{n!}{(n-m+k-1)!}$, which is equal to

$$- n^{1/2} K_{n,0,m} - n^{-(m+1)/2} n^{m+1} (-1)^{m+1}. \tag{A.4}$$

Now, we add (A.1) and the terms (A.3) and (A.4) which sum up (A.2). The terms (A.1) cancel with terms appearing in (A.2) and we get

$$\begin{aligned} K_{n,0,m+1} &= n^{-1/2} (n-m) K_{n,0,m} - m K_{n,0,m-1} - n^{1/2} K_{n,0,m} \\ &= -n^{-1/2} m K_{n,0,m} - m K_{n,0,m-1}. \quad \square \end{aligned}$$

Proof of the induction formula (14). It will be useful to employ the following, more compact expression for $K_{n,j,m}$:

$$K_{n,j,m} = n^{-(m-j)/2} \frac{m!}{j!} \sum_{k=0}^{m-j} (-1)^k \frac{n^k}{k!} \binom{n-j}{m-j-k}.$$

Using the last expression,

$$\begin{aligned} K_{n,j+1,m} &= n^{-(m-j-1)/2} \frac{m!}{(j+1)!} \sum_{k=0}^{m-j-1} (-1)^k \frac{n^k}{k!} \binom{n-j-1}{m-j-1-k} \\ &= n^{-(m-j-1)/2} \frac{m!}{(j+1)!} \sum_{k=0}^{m-j-1} (-1)^k \frac{n^k}{k!} \binom{n-j}{m-j-k} \end{aligned} \tag{A.5}$$

$$- n^{-(m-j-1)/2} \frac{m!}{(j+1)!} \sum_{k=0}^{m-j-1} (-1)^k \frac{n^k}{k!} \binom{n-j-1}{m-j-k}. \tag{A.6}$$

We can treat separately these two terms. (A.5) is equal to

$$\frac{n^{1/2}}{j+1} K_{n,j,m} - n^{-(m-j-1)/2} \frac{m!}{(j+1)!} (-1)^{m-j} \frac{n^{m-j}}{(m-j)!}.$$

(A.6) can be handled using

$$\binom{n-j-1}{m-j-k} = \frac{1}{n-j} \binom{n-j}{m-j-k} ((n-m)+k).$$

Therefore, (A.6) is equal to the sum of the two terms

$$-\frac{n-m}{n-j} n^{-(m-j-1)/2} \frac{m!}{(j+1)!} \sum_{k=0}^{m-j-1} (-1)^k \frac{n^k}{k!} \binom{n-j}{m-j-k} \tag{A.7}$$

$$-\frac{1}{n-j} n^{-(m-j-1)/2} \frac{m!}{(j+1)!} \sum_{k=0}^{m-j-1} (-1)^k \frac{n^k}{k!} \binom{n-j}{m-j-k}. \tag{A.8}$$

(A.7) is equal to

$$-\frac{n-m}{(n-j)(j+1)}n^{1/2}K_{n,j,m} + \frac{n-m}{(n-j)}n^{-(m-j-1)/2}\frac{m!}{(j+1)!}(-1)^{m-j}\frac{n^{m-j}}{(m-j)!}.$$

The term (A.8) is

$$\frac{n}{n-j}\frac{m}{j+1}K_{n,j,m-1} + \frac{1}{n-j}n^{-(m-j-1)/2}\frac{m!}{(j+1)!}(-1)^{m-j}\frac{n^{m-j}}{(m-j-1)!}$$

Therefore, after simplification, $K_{n,j+1,m}$ is equal to the following:

$$n^{1/2}\frac{m-j}{n-j}K_{n,j,m} + \frac{n}{n-j}\frac{m}{j+1}K_{n,j,m-1}. \quad \square$$

References

- Arcones, M.A., Giné, E., 1992. On the bootstrap of U and V statistics. *Ann. Statist.* 20 (2), 655–674.
- Arcones, M.A., Giné, E., 1993. On decoupling, series expansions, and tail behavior of chaos processes. *J. Theoret. Probab.* 6 (1), 101–122.
- de la Peña, V.H., Giné, E., 1999. Decoupling. In: *Probability and its Applications* (New York). Springer-Verlag, New York.
- Dynkin, E.B., Mandelbaum, A., 1983. Symmetric statistics, Poisson point processes, and multiple Wiener integrals. *Ann. Statist.* 11 (3), 739–745.
- Major, P., 2006. An estimate on the supremum of a nice class of stochastic integrals and U -statistics. *Probab. Theory Related Fields* 134 (3), 489–537.
- Major, P., 2007. On the tail behaviour of multiple random integrals and degenerate U -statistics. Preprint.
- Nualart, D., 1995. Malliavin calculus and related topics. In: *Probability and its Applications* (New York). Springer-Verlag, New York.
- Rubin, H., Vitale, R.A., 1980. Asymptotic distribution of symmetric statistics. *Ann. Statist.* 8 (1), 165–170.