The Annals of Statistics
1984, Vol. 12, No. 2, 483–496

INVARIANCE PRINCIPLE FOR SYMMETRIC STATISTICS

By Avi Mandelbaum and Murad S. Taqqu[1]

Cornell University

We derive invariance principles for processes associated with symmetric statistics of arbitrary order. Using a Poisson sample size, such processes can be viewed as functionals of a Poisson Point Process. Properly normalized, these functionals converge in distribution to functionals of a Gaussian random measure associated with the distribution of the observations. We thus obtain a natural description of the limiting process in terms of multiple Wiener integrals. The results are used to derive asymptotic expansions of processes arising from arbitrary square integrable U-statistics.

1. Introduction. The asymptotic behavior of symmetric statistics of finite order has been an active area of research ever since the pioneering work of Von Mises [15] and Hoeffding [4]. We refer the reader to Serfling [14] for results about limiting distributions and to Sen [13] for invariance principles. Rubin and Vitale [12] showed that the limiting distribution is equivalent to that of a linear combination of products of Hermite polynomials of independent $N(0, 1)$ random variables. Using ideas from the theory of Poisson point processes and Gaussian random measures, Dynkin and Mandelbaum [3] obtained a description of the asymptotic distribution for symmetric statistics in terms of multiple Wiener integrals. In this paper we extend the results of [3] and get invariance principles for the corresponding processes.

An alternate approach for obtaining functional limit theorems for U-statistics is due to Denker, Grillenberger and Keller [2]. They first reduce their setup to the case of observations from a uniform distribution on $[0, 1]$. Then they describe the limiting process for Von Mises statistics as a $C[0, 1]$ valued integral with respect to the Brownian bridge component of a Kiefer process. The limiting process for U-statistics is derived as a consequence.

In our paper, the observations take values in an arbitrary measurable space. Our approach enables us to deal directly with the infinite order generalization of the U-statistic. We show that the limiting process can be expressed as an infinite sum of multiple Wiener integrals. The specific representation that we use has several virtues. It clarifies the role of the time parameter, it exhibits the symmetric functions that defined the statistic and it also involves the distribution of the observations. The distribution of the observations becomes a measure associated with the underlying Gaussian random measure.

The paper treats the one-sample problem. Our technique applies to the J-sample problem considered in [2] but we do not consider that case here.

Received March 1983; revised September 1983.

[1] The second author was supported by National Science Foundation Grant ECS-80-15585.

AMS 1980 *subject classification.* Primary 60F17; secondary 62E20, 62G05, 60G99, 60K99.

Key words and phrases. Symmetric statistics, invariance principle, multiple Wiener integral, U-statistics, Hermite polynomials.

There are five sections and an appendix. The results are stated in Sections 2 and 3. The proofs are developed in Section 3 and make use of lemmas which are established in Section 5. Section 4 contains applications, in particular to U-statistics. The appendix provides a brief review on multiple Wiener integrals.

2. The main results. Let $X_1, X_2, \cdots$, be a sequence of independent and identically distributed random variables with values in an arbitrary measurable space $(\mathscr{X}, \mathscr{B})$, and let ν be their distribution: $\nu(B) = P\{X_i \in B\}$, $B \in \mathscr{B}$.

For each $T > 0$, we define the Hilbert space $\mathscr{A}_T^*$ of sequences $h = (h_0, h_1(x_1), \cdots, h_k(x_1, \cdots, x_k), \cdots)$ where h_k are symmetric measurable functions and

$$\| h \|_T^2 = \sum_{k=0}^{\infty} \frac{T^k}{k!} Eh_k^2 < \infty.$$

We use the notation

$$Eh_k^2 = Eh_k^2(X_1, \cdots, X_k) = \int_{\mathscr{X}} \cdots \int_{\mathscr{X}} h_k^2(x_1, \cdots, x_k)\nu(dx_1) \cdots \nu(dx_k).$$

Clearly $\mathscr{A}_T^*$ decreases as T increases. Set $\mathscr{A}_\infty^* = \cap_{T>0} \mathscr{A}_T^*$, then $h \in \mathscr{A}_\infty^*$ if and only if $\| h \|_T^2 < \infty$ for all $T > 0$.

To a symmetric function $h_k(x_1, \cdots, x_k)$, we associate the symmetric statistic

$$\sigma_k^n(h_k) = \sum_{1 \le s_1 < s_2 < \cdots < s_k \le n} h_k(X_{s_1}, \cdots, X_{s_k}).$$

The statistic σ_k^n is based on a sample $X_1, \cdots, X_n$ of size n, taken from a population with distribution ν. Define $\sigma_0^n(h_0) = h_0$, and for $k > n$, set $\sigma_k^n(h_k) = 0$.

Any symmetric statistic can be uniquely represented in terms of symmetric statistics associated with functions h_k satisfying the condition

$$Eh_k(x_1, \cdots, x_{k-1}, X_k) = \int_{\mathscr{X}} h_k(x_1, \cdots, x_{k-1}, y)\nu(dy) = 0. \tag{2.1}$$

(see Section 4). A symmetric function h_k which satisfies (2.1) is called *canonical.* Denote by $\mathscr{A}_T$ the closed subspace of $\mathscr{A}_T^*$ consisting of sequences $h = (h_0, h_1, h_2, \cdots)$ where $h_1, h_2, \cdots$ are canonical.

In this paper we investigate the limiting distribution, as $n \to \infty$, of the stochastic process

$$Y_n^t(h) = \sum_{k=0}^{\infty} \frac{1}{n^{k/2}} \sigma_k^{[nt]}(h_k), \quad 0 \le t \le T, \tag{2.2}$$

where $h = (h_0, h_1, h_2, \cdots) \in \mathscr{A}_T$. We also present a corresponding result for $h \in \mathscr{A}_T^*$. The limiting distribution is expressed in terms of Wiener integrals with respect to Gaussian random measures. For the convenience of the reader we give a brief description of these notions in the appendix.

Our main result is:

THEOREM 1. *Let $h = (h_0, h_1, \cdots) \in \mathscr{A}_T$, $0 < T \le \infty$. As $n \to \infty$, the process*

$Y_n^t(h)$ converges weakly in $D[0, T]$ to

$$W^t(h) = \sum_{k=0}^{\infty} \frac{1}{k!} I_k(h_k^t), \quad 0 \le t \le T. \tag{2.3}$$

Here

$$h_k^t = h_k(x_1, \cdots, x_k) 1_{[0,t]}(u_1) \cdots 1_{[0,t]}(u_k),$$

and $I_k(\cdot)$ are multiple Wiener integrals with respect to the Gaussian random measure $W(\cdot)$ on the product space

$$(\mathscr{X}, \mathscr{B}, \nu) \times ([0, T], \textit{Borel, Lebesgue}).$$

Symbolically,

$$I_k(h_k^t) = \int \cdots \int h_k(x_1, \cdots, x_k) 1_{[0,t]}(u_1) \cdots 1_{[0,t]}(u_k) \, W(dx_1, du_1) \cdots W(dx_k, du_k). \tag{2.4}$$

The main steps of the proof of Theorem 1 are given in Section 3 while auxiliary lemmas are proved in Section 5.

The space $D[0, T]$, $T \le \infty$ is the space of right continuous functions on $[0, T]$ ($[0, \infty)$ when $T = \infty$) with left limits at each $t \le T$ ($t < \infty$ when $T = \infty$). The space $D[0, T]$, $T < \infty$, is endowed with Skorohod's J_1 topology (see [1]). The topology in $D[0, \infty] = D[0, \infty)$ is the one described for example in Whitt [16]. Weak convergence in $D[0, \infty)$ means convergence in $D[0, T]$ for all fixed $0 < T < \infty$ at which the limiting process is continuous with probability one. This restriction does not play a role in Theorem 1 because

LEMMA 1. *For $h \in \mathscr{H}_T$, the process $W^t(h)$ described in* (2.3) *has a continuous version.*

Lemma 1 is proved at the end of Section 5. It follows from its proof that Theorem 1 applies also to convergence in $C[0, T]$, $T \le \infty$, if $Y_n^t(h)$ is suitably redefined using linear interpolation.

We now formulate the analogue of Theorem 1 without restricting h_k to be canonical. Let $B_1, B_2, \cdots$ be a sequence of iid real valued random variables with mean 0 and variance 1, independent of $X_1, X_2, \cdots$.

THEOREM 1*. *Let $h \in \mathscr{H}_T^*$, $0 < T \le \infty$. As $n \to \infty$, the process*

$$\sum_{k=0}^{\infty} \frac{1}{n^{k/2}} \sum_{1 \le s_1 < \cdots < s_k \le [nt]} B_{s_1} \cdots B_{s_k} h_k(X_{s_1}, \cdots, X_{s_k}) \tag{2.5}$$

converges weakly in $D[0, T]$ to the process $W^t(h)$ described in (2.3).

Theorem 1* follows from Theorem 1. Indeed, the functions

$$\tilde{h}_k(x_1, b_1; \cdots; x_k, b_k) = h_k(x_1, \cdots, x_k) b_1 \cdots b_k$$

are canonical when viewed as functions of the pairs $(X_1, B_1), \cdots, (X_k, B_k)$ and the process $W^t(\tilde{h})$ has the same distribution as $W^t(h)$ (see Section A.5 of the appendix).

3. Poissonization. To establish weak convergence in $D[0, T]$ for the process $Y_n^t(h)$ defined in (2.2) we apply two ideas found in Dynkin and Mandelbaum [3]. One idea is Poissonization of the sample size. The second involves showing that $Y_n^t(h)$ converges for special sequences $h = h^\phi$ of the form

$$h_0^\phi = 1,\ h_k^\phi(x_1, \cdots, x_k) = \phi(x_1) \cdots \phi(x_k), \quad k = 1, 2, \cdots \tag{3.1}$$

where ϕ is a fixed function. The general case follows from this particular one.

To prove Theorem 1 we introduce two identically distributed Poissonized versions of the process $Y_n^t(h)$. The first (defined in Theorem 2) has the same limiting distribution as $Y_n^t(h)$. The second (defined in Theorem 3) is used to identify the limiting distribution as that of $W^t(h)$.

Let $N_\lambda(t)$, $0 \le t \le T$ be a Poisson process ($EN_\lambda(t) = \lambda t$) which is independent of $X_1, X_2, \cdots$.

THEOREM 2. *Let $h \in \mathscr{A}_T$. As $\lambda \to \infty$, the process*

$$Z_\lambda^t(h) = \sum_{k=0}^\infty \frac{1}{\lambda^{k/2}} \sigma_k^{N_\lambda(t)}(h_k), \quad 0 \le t \le T,$$

converges weakly in $D[0, T]$ to $W^t(h)$ described in (2.3).

The processes $Y_n^t(h)$ and $Z_n^t(h)$ have the same limiting finite-dimensional distribution as $n \to \infty$. This follows from the Cramer-Wold device (see [1]) and

LEMMA 2. *Let $h \in \mathscr{A}_T$. For each $t \in [0, T]$,*

$$\lim_{n\to\infty} E[Y_n^t(h) - Z_n^t(h)]^2 = 0.$$

Lemma 2, as well as Lemmas 3, 4, 5 below are proved in Section 5.

Let now $U_1, U_2, \cdots$ be iid random variables uniformly distributed on $[0, T]$, $T < \infty$. We suppose that $U_1, U_2, \cdots$ are independent of the sequence $X_1, X_2, \cdots$ and of the random variable $N_\lambda(T)$.

THEOREM 3. *Let $h \in \mathscr{A}_T$. As $\lambda \to \infty$, the process*

$$C_\lambda^t(h) = \sum_{k=0}^\infty \frac{1}{\lambda^{k/2}} \sigma_k^{N_\lambda(T)}(h_k^t), \quad 0 \le t \le T,$$

converges weakly in $D[0, T]$ to $W^t(h)$ described in (2.3), *where*

$$\sigma_k^n(h_k^t) = \sum_{1 \le s_1 < \cdots < s_k \le n} h(X_{s_1}, \cdots, X_{s_k}) 1_{[0,t]}(U_{s_1}) \cdots 1_{[0,t]}(U_{s_k}).$$

In fact,

LEMMA 3. *For $h \in \mathscr{A}_T$ and λ fixed, the processes $C_\lambda^t(h)$ and $Z_\lambda^t(h)$, are identically distributed as random elements of $D[0, T]$.*

Combining Lemma 3 with the remark preceding Lemma 2, we see that the three processes $Y_n^t(h)$, $Z_\lambda^t(h)$, $C_\lambda^t(h)$ have the same limiting finite-dimensional distributions. The easiest to investigate is the process $C_\lambda^t(h)$. We use the facts that for a fixed t, the random variable $C_\lambda^t(h)$ can be written as

$$C_\lambda^t(h) = \sum_{k=0}^{\infty} \frac{1}{\mu^{k/2}} \sigma_k^{N_\mu}(\tilde{h}_k) \tag{3.2}$$

where $\mu = \lambda T$, $\tilde{h}_k = T^{k/2} h_k^t$, N_μ is Poisson with mean μ, and that the limiting distribution of random variables of the form (3.2) as $\mu \to \infty$ ($\lambda \to \infty$) was derived in [3]. We get

LEMMA 4. *The finite-dimensional distributions of $C_\lambda^t(h)$ converge to those of $W^t(h)$.*

To complete the proofs of Theorem 1, 2 and 3 we must verify tightness.

Let Φ stand for the set of measurable functions $\phi(x)$, $x \in \mathscr{X}$ which take a finite number of values and satisfy $E\phi = \int_{\mathscr{X}} \phi(x)\nu(dx) = 0$. For $\phi \in \Phi$, the functions h_k^ϕ are canonical, and $h^\phi \in \mathscr{A}_\infty$ since

$$\| h^\phi \|_T^2 = \sum_{k=0}^{\infty} \frac{T^k}{k!} (E\phi^2)^k = e^{TE\phi^2} < \infty, \quad 0 < T < \infty.$$

In proving their Theorem 2, Dynkin and Mandelbaum [3] showed that for $T > 0$, elements $h \in \mathscr{A}_T$ of the form

$$h = \sum_{k=1}^{m} \alpha_k h^{\phi_k}, \quad \alpha_1, \cdots, \alpha_m \text{ constants} \tag{3.3}$$

are dense in $\mathscr{A}_T$. We also note that the processes $Y_n^t(h)$, $Z_\lambda^t(h)$, $C_\lambda^t(h)$ are all square integrable martingales when $h \in \mathscr{A}_T$. Indeed, Y_n^t is a martingale with respect to the filtration $\sigma(X_1, X_2, \cdots, X_{[nt]})$, $t > 0$; Z_λ^t with respect to $\sigma(X_1, \cdots, X_{N_\lambda(t)})$; C_λ^t with respect to $\sigma(X_i, i \in S(t))$ where

$$S(t) = \{i \colon i \le N_\lambda(t),\ U_i \le t\}. \tag{3.4}$$

We use these facts to prove

LEMMA 5. *Let $T < \infty$. The processes $Y_n^t(h)$, $Z_\lambda^t(h)$, $C_\lambda^t(h)$ are tight in $D[0, T]$ for $h \in \mathscr{A}_T$ if they are tight for h of the form h^ϕ, $\phi \in \Phi$.*

When $h = h^\phi$, all the above three processes can be represented in a form from which tightness is easy to prove and the limiting distribution is easily read off.

We start with $Y_n^t(h^\phi)$ which has the form

$$Y_n^t(h^\phi) = \prod_{i=1}^{[nt]} \left(1 + \frac{\phi(X_i)}{\sqrt{n}}\right). \tag{3.5}$$

Since $\phi \in \Phi$ takes finite number of values, $Y_n^t(h^\phi)$ is positive for n large enough. By Taylor's expansion for $\log(1 + x)$, $\log Y_n^t(h^\phi)$ has the same weak limit in $D[0, T]$ as

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \phi(X_i) - \frac{1}{2n} \sum_{i=1}^{[nt]} \phi^2(X_i). \tag{3.6}$$

From Donsker's Theorem and the weak law of large numbers it follows that the limit of (3.6) is

$$I_1(\phi^t) - \tfrac{1}{2}\, tE\phi^2$$

where

$$I_1(\phi^t) = \int \phi(x) 1_{[0,t]}(u)\, W(dx, du)$$

is a Brownian motion with mean 0 and variance $E[I_1(\phi^1)]^2 = E\phi^2$. Hence, for $\phi \in \Phi$, $Y_n^t(h^\phi)$ converges weakly in $D[0, T]$ to the exponential martingale

$$\varepsilon^t(\phi) = e^{I_1(\phi^t) - (1/2)tE\phi^2}.$$

Theorem 1 is now established. Theorem 3 follows from Theorem 2 using Lemma 3. To prove Theorem 2, we first show that $Z_n^t(h^\phi)$ and $Y_n^t(h^\phi)$ have the same weak limit. The process $Z_n(t) \equiv Z_n^t(h)$ is a random time change of $Y_n(t) \equiv Y_n^t(h)$, namely

$$Z_n(t) = Y_n\left(\frac{N_n(t)}{n}\right).$$

Since, as $n \to \infty$, $N_n(t)/n$ converges to t and $\varepsilon^t(\phi)$ is continuous in t, we can apply the results of Section 17 in [1] to conclude that $Z_n^t(h^\phi)$ and $Y_n^t(h^\phi)$ have the same weak limit in $D[0, T]$. Theorem 2 now follows from Lemma 5 and the fact that $Y_n^t(h^\phi)$ converges weakly to $\varepsilon^t(\phi)$.

This concludes the proofs of Theorems 1, 2 and 3.

4. Applications. Let $X_1, X_2, \cdots$ be defined as in Section 2. To every symmetric function $h(x_1, \cdots, x_m)$ there corresponds a "U-statistic"

$$U_m^n(h) = \frac{1}{\binom{n}{m}}\, \sigma_m^n(h) \tag{4.1}$$

based on the sample $X_1, \cdots, X_n$. The U-statistic is an unbiased estimator of the parameter

$$Eh = \int_{\mathscr{X}} \cdots \int_{\mathscr{X}} h(x_1, \cdots, x_m)\nu(dx_1) \cdots \nu(dx_m).$$

We assume throughout that $Eh^2 < \infty$.

Applying Theorem 1, we shall obtain the weak limit in $D[0, \infty)$ of stochastic processes related to $U_m^{[nt]} - Eh$.

For this purpose, we use Hoeffding's orthogonal decomposition of the U statistics, that is

$$U_m^n(h) = \sum_{i=0}^m U_i^n(h_i) = Eh + \sum_{i=1}^m \frac{1}{\binom{n}{i}} \sigma_m^n(h_i) \tag{4.2}$$

where the h_i are canonical. These h_i can be obtained from h through the relation

$$h_i(X_1, \cdots, X_i) = \binom{m}{i} (I - Q_1^m) \cdots (I - Q_i^m)Q_{i+1}^m \cdots Q_m^m h(X_1, \cdots, X_m)$$

where Q_i^m is the conditional expectation with respect to $X_1, \cdots, X_{i-1}, X_{i+1}, \cdots, X_m$. The representation (4.2) was first established in [5] (see also [12], [13] and the appendix in [3]).

Now suppose that h has rank $k \geq 1$, that is

$$h_1 = \cdots = h_{k-1} = 0, \quad h_k \neq 0.$$

Then

$$n^{k/2}t^k(U_m^{[nt]} - Eh) = \sum_{i=k}^m \frac{n^{k/2}}{\binom{[nt]}{i}} t^k \sigma_i^{[nt]}(h_i). \tag{4.3}$$

Since the h_i, $i = 1, \cdots, m$ are canonical, the terms in the right-hand side of (4.3) are uncorrelated, so that for $t > 0$,

$$E[n^{k/2}t^k(U_m^{[nt]} - Eh)]^2 = \sum_{i=k}^m \frac{n^k t^{2k}}{\binom{[nt]}{i}^2} \binom{[nt]}{i} Eh_i^2 \sim k!t^k Eh_k^2$$

as $n \to \infty$. Therefore only the first term in the expansion of (4.3) contributes in the limit. The limit is identical to that of $k!\ (1/n^{k/2})\sigma_k^{[nt]}(h_k)$. Using Theorem 1 we conclude:

COROLLARY 1. *If h has rank k, then, as $n \to \infty$,*

$$n^{k/2}t^k[U_m^{[nt]}(h) - Eh]$$

converges weakly to $I_k(h_k^t)$ in $D[0, \infty)$.

Moreover, if h has rank k, then $U_m^{[nt]}(h)$ has the following asymptotic expansion

$$U_m^{[nt]}(h) \sim Eh + \frac{1}{n^{k/2}t^k} I_k(h_k^t) + \frac{1}{n^{(k+1)/2}t^{(k+1)/2}} I_{k+1}(h_{k+1}^t)$$

$$+ \cdots + \frac{1}{n^{m/2}t^m} I_m(h_m^t).$$

Note that both the rank k and the functions h_k, h_{k+1}, $\cdots$ depend on the distribution ν.

When $k=1$, $\sqrt{n}t(U_m^{[nt]}(h)-Eh)$ converges to $I_1(h_1^t)$ which is Brownian motion with variance Eh_1^2.

When $k=2$, $nt^2(U_m^{[nt]}(h)-Eh)$ converges to $I_2(h_2^t)$ which can be expressed as

$$\sum_{j=1}^{\infty} \lambda_j (B_j^2(t) - t)$$

where $B_1(t)$, $B_2(t)$, $\cdots$ are independent standard Brownian motions and $\sum_{j=1}^{\infty} \lambda_j^2 < \infty$.

The case $k = 1$, called the "non-degenerate" case, was proved in Miller and Sen [8]. The case $k=2$ follows from a special representation of a multiple Wiener integral of order 2 (see Proposition 6.18 in Neveu [11], for example). It was first proved in Neuhaus [10].

Our representation of the limiting process $I_k(k_k^t)$ is very natural. It involves explicitly the time parameter t, the function h_k, and it involves the distribution ν through the Gaussian random measure W. The representation of the limiting process simplifies further when the kernel h is a polynomial in the variables because then the multiple Wiener integrals become Hermite polynomials of Gaussian random variables. As an illustration, we can easily establish the following limit theorem stated for $D(0, 1)$ in Mori and Székely [9].

COROLLARY 2. *Let X_1, X_2, $\cdots$ be a sequence of* i.i.d. *random variables with mean* 0 *and variance* 1. *Let $p_k(x_1, \cdots, x_k) = x_1 \cdots x_k$. The process $n^{-k/2}\sigma_k^{[nt]}(p_k)$ converges in $D(0, \infty)$ to*

$$\frac{t^{k/2}}{k!} H_k\left(\frac{B(t)}{\sqrt{t}}\right), \quad t > 0$$

where $B(t)$ is the standard Brownian motion and where H_k is the Hermite polynomial of order k with leading coefficient 1.

Corollary 2 is an immediate consequence of Theorem 1 because p_k is canonical and $I_k(p_k^t) = (t^{k/2}/k!)H_k(t^{-1/2}I_1(p_1^t))$ (see Section A.1 of the appendix).

5. Proofs of lemmas. The proof of Lemma 1 is given at the end of this section. To complete the proofs of Theorems 1, 2 and 3, we now prove Lemmas 2–5.

PROOF OF LEMMA 2. The proof is similar to the one for $t = 1$ given in [3] (Section 2.4) and so we only outline the main steps.

It is easy to check (see also (2.3) in [3]) that

$$EZ_n^t(h)^2 = \| h \|_t^2, \tag{5.1}$$

hence it is sufficient to show:

$$EY_n^t(h)^2 \to \| h \|_t^2, \quad \text{as} \quad n \to \infty, \tag{5.2}$$

$$EY_n^t(h) Z_n^t(h) \to \| h \|_t^2, \quad \text{as} \quad n \to \infty. \tag{5.3}$$

But

$$EY_n^t(h)^2 = \sum_{k=0}^{\infty} a_{n,k}^t \frac{t^k}{k!} Eh_k^2, \tag{5.4}$$

$$EY_n^t(h) Z_n^t(h) = \sum_{k=0}^{\infty} b_{n,k}^t \frac{t^k}{k!} Eh_k^2, \tag{5.5}$$

where

$$a_{n,k}^t = \frac{[nt]!}{([nt]-k)!} (nt)^{-k}$$

$$b_{n,k}^t = P\{N_n(t) \le [nt] - k\} + a_{n,k}^t P\{N_n(t) \ge [nt] + 1\}.$$

Both $a_{n,k}^t$ and $b_{n,k}^t$ are bounded by 1 and $a_{n,k}^t \to 1$ as $n \to \infty$. By the central limit theorem, $P\{N_n(t) \le [nt] - k\} \to 1/2$ as $n \to \infty$, for $k = 0, 1, \cdots$. Hence (5.2) and (5.3) follow from (5.4), (5.5) and the dominated convergence theorem.

PROOF OF LEMMA 3. The processes $C_\lambda^t(h)$ and $Z_\lambda^t(h)$ both have sample paths in $D[0, T]$, hence it suffices to prove that their finite-dimensional distributions coincide. We first prove it for $h = h^\phi$, $\phi \in \Phi$.

As in (3.5), for $\phi \in \Phi$, we have

$$\log Z_\lambda^t(h^\phi) = \sum_{i=1}^{N_\lambda(t)} \log\left(1 + \frac{\phi(X_i)}{\sqrt{\lambda}}\right)$$

$$\log C_\lambda^t(h^\phi) = \sum_{i=1}^{N_\lambda(T)} \log\left(1 + \frac{\phi(X_i) 1_{[0,t]}(U_i)}{\sqrt{\lambda}}\right).$$

Letting $\alpha(X_i) = \log(1 + \phi(X_i)/\sqrt{\lambda})$, we see that

$$\log Z_\lambda^t(h^\phi) = \sum_{i=1}^{N_\lambda(t)} \alpha(X_i), \tag{5.6a}$$

$$\log C_\lambda^t(h^\phi) = \sum_{i=1}^{N_\lambda(t)} 1_{[0,t]}(U_i) \alpha(X_i), \tag{5.6b}$$

which shows that the processes in (5.6a) and (5.6b) are both constructions of the same compound Poisson process. Taking logarithms and using the Cramer-Wold device we prove similarly that the process $(Z_\lambda^t(h^{\phi_1}), \cdots, Z_\lambda^t(h^{\phi_m}))$ has the same distribution as $(C_\lambda^t(h^{\phi_1}), \cdots, C_\lambda^t(h^{\phi_m}))$.

Since both $C_\lambda^t(h)$ and $Z_\lambda^t(h)$ depend linearly on h, we conclude that Lemma 2 holds for h of the form (3.3), which are dense in $\mathscr{H}_T$. Hence to end the proof of

Lemma 3, it suffices to show that for λ fixed, $0 \le t_1 < \cdots < t_m \le T$, we have

$$(5.7) \qquad (C_\lambda^{t_1}(g), \cdots, C_\lambda^{t_m}(g)) \Rightarrow (C_\lambda^{t_1}(h), \cdots, C_\lambda^{t_m}(h))$$

and

$$(5.8) \qquad (Z_\lambda^{t_1}(g), \cdots, Z_\lambda^{t_m}(g)) \Rightarrow (Z_\lambda^{t_1}(h), \cdots, Z_\lambda^{t_m}(h))$$

as $g \to h$ in $\mathscr{A}_T$. ($\Rightarrow$ denotes convergence in distribution). The convergence in (5.7) is a consequence of the representation (3.2) and Theorem 2 in [3]. The convergence in (5.8) is implied by the relations:

$$\begin{aligned} | Ee^{i\beta\sum_{k=1}^m \alpha_k Z_\lambda^{t_k}(g)} - Ee^{i\beta\sum_{k=1}^m \alpha_k Z_\lambda^{t_k}(h)} | &\le |\beta| \textstyle\sum_{k=1}^m |\alpha_k| \{E[Z_\lambda^{t_k}(g-h)]^2\}^2 \\ &\le |\beta| \textstyle\sum_{k=1}^m |\alpha_k| \cdot \{E[Z_\lambda^T(g-h)]^2\}^{1/2} \\ &= |\beta| \textstyle\sum_{k=1}^m |\alpha_k| \, \| g - h \|_T. \end{aligned}$$

The first inequality follows from $| e^{i\sum A_k} - e^{i\sum B_k} | \le \sum |A_k - B_k|$. The second inequality holds since $Z_\lambda^t(h)$ is a square integrable martingale when $h \in \mathscr{A}_T$. The last equality is a restatement of (5.1).

REMARK. Lemma 3 is also a consequence of the fact that the following two processes with values in the configuration space of a Poisson point process are identically distributed. One process is $(X_i, 1 \le i \le N_\lambda(t))$. The second process is $(X_i, i \in S(t))$, where $S(t)$ is defined in (3.4).

PROOF OF LEMMA 4. Let $\alpha_1, \cdots, \alpha_m$ be arbitrary constants. By linearity,

$$(5.9) \qquad \textstyle\sum_{j=1}^m \alpha_j C_\lambda^{t_j}(h) = \sum_{k=0}^\infty \lambda^{-k/2} \sigma_\lambda^{N_\lambda(T)} (\sum_{j=1}^m \alpha_j h_k^{t_j})$$

where $\sum_{j=1}^m \alpha_j h_\lambda^{t_j}$ is a symmetric canonical function whose arguments are the random variables (X_i, U_i) with values in the product space

$$\mathscr{S}_T = (\mathscr{X}, \mathscr{B}, \nu) \times \left([0, T], \text{Borel}, \frac{\text{Lebesgue}}{T}\right).$$

By Theorem 2 of [3], the right-hand side of (5.9) converges to the random variable

$$L = \textstyle\sum_{k=0}^\infty \dfrac{T^{k/2}}{k!} \bar{I}_k(\sum_{j=1}^m \alpha_j h_k^{t_j}) = \sum_{j=1}^m \alpha_j \sum_{k=0}^\infty \dfrac{T^{k/2}}{k!} \bar{I}_k(h_k^{t_j})$$

where

$$\begin{aligned} T^{k/2}\bar{I}_k(h_k^{t_j}) = T^{k/2} \int \cdots \int h_k(x_1, \cdots, x_k) 1_{[0,t_j]}(u_1) \cdots 1_{[0,t_j]}(u_k) \\ W_T(dx_1, du_1) \cdots W_T(dx_k, du_k) \end{aligned}$$

is a multiple Wiener integral of order k with respect to the Gaussian random measure $W_T(\cdot)$ on the product space $\mathscr{S}_T$. Note that for $B \in \mathscr{B}$ and C Borel in

$[0, T]$,

$$EW_T^2(B \times C) = \frac{1}{T} EW^2(B \times C)$$

where $W(\cdot)$ is the Gaussian random measure defined in Theorem 1. The change of variables formula for multiple Wiener integrals given in Section A.3 of the appendix allows us to conclude that L has the same distribution as

$$\sum_{j=1}^m \alpha_j \sum_{k=0}^\infty \frac{1}{k!} I_k(h_k^{t_j}) = \sum_{j=1}^m \alpha_j W^{t_j}(h),$$

where $I_k(h_k^t)$ is described by (2.4).

Proof of Lemma 5. Fix $0 < T < \infty$ and $h \in \mathscr{H}_T$. Let $X^t(h)$, $0 \le t \le T$ stand for any of the three processes $Y_n^t(h)$, $Z_\lambda^t(h)$ or $C_\lambda^t(h)$. Then for $\varepsilon > 0$, $\delta > 0$, $g \in \mathscr{H}_T$

$$\begin{aligned}
&P\{\sup_{|s-t|<\delta} |X^t(h) - X^s(h)| > 3\varepsilon\} \\
&\quad \le P\{\sup_{|s-t|<\delta} |X^t(h-g)| > \varepsilon\} + P\{\sup_{|s-t|<\delta} |X^t(g) - X^s(g)| > \varepsilon\} \\
&\qquad + P\{\sup_{|s-t|<\delta} |X^s(h-g)| > \varepsilon\} \\
&\quad \le 2P\{\sup_{0 \le t < T} |X^t(h-g)| > \varepsilon\} + P\{\sup_{|s-t|<\delta} |X^t(g) - X^s(g)| > \varepsilon\} \\
&\quad \le \frac{2}{\varepsilon^2} E[X^T(h-g)]^2 + P\{\sup_{|s-t|<\delta} |X^t(g) - X^s(g)| > \varepsilon\}
\end{aligned}$$

where the last inequality is an application of Kolmogorov's inequality to the square integrable martingale $X^t(h-g)$. Now $E[X^T(h-g)]^2$ is equal to $\| h - g \|_T^2$ when X^t is either Z_λ^t or C_λ^t (by (2.3) of [3]) and tends to it, as $n \to \infty$, when X^t is Y_n^t (by Lemma 1). It is therefore sufficient to establish tightness of $X^t(h)$ for functions h of the form (3.3) which are dense in $\mathscr{H}_T$. For such an h

$$\begin{aligned}
&P\{\sup_{|s-t|<\delta} |X^t(h) - X^s(h)| > \varepsilon\} \\
&\quad \le \sum_{k=1}^m P\left\{\sup_{|s-t|<\delta} |X^t(h^{\phi_k}) - X^s(h^{\phi_k})| > \frac{\varepsilon}{m|\alpha_k|}\right\},
\end{aligned}$$

which proves Lemma 5.

Proof of Lemma 1. Let $T \ge 1$ be integer and let $\tilde{Y}_n^t(h)$, $0 \le t \le T$, be $Y_n^t(h)$, $0 \le t \le T$, redefined through linear interpolation. The process $\tilde{Y}_n^t(h)$ has continuous paths and its finite-dimensional distributions converge as $n \to \infty$ to those of $W^t(h)$. The sequence $\{\tilde{Y}_n^t(h^\phi), n \ge 1\}$ is tight in $C[0, T]$. Furthermore, Lemma 5 applies to $\tilde{Y}_n^t(h)$. The only change in the proof involves replacing the direct application of Kolmogorov's inequality by

$$\begin{aligned}
P\{\sup_{0 \le t \le T} |\tilde{Y}_n^t(h-g)| > \varepsilon\} &= P\{\sup_{t \in \{0,1,\dots,T\}} |\tilde{Y}_n^t(h-g)| > \varepsilon\} \\
&\le \frac{1}{\varepsilon^2} E|\tilde{Y}_n^T(h-g)|^2.
\end{aligned}$$

Therefore $\tilde{Y}_n^t(h)$ converges weakly to $W^t(h)$ in $C[0, T]$ as $n \to \infty$. The conclusion of the lemma follows.

APPENDIX: MULTIPLE WIENER INTEGRALS

A.1. Let $(\mathscr{X}, \mathscr{B}, \nu)$ be an arbitrary measure space. The Wiener integrals are defined as the Gaussian family $\{I_1(\phi), \phi \in L^2(\mathscr{X}, \mathscr{B}, \nu)\}$ with moments

$$EI_1(\phi) = 0, \quad EI_1(\phi)I_1(\psi) = \nu(\phi\psi) = \int_{\mathscr{X}} \phi(x)\psi(x)\nu(dx).$$

The subfamily $\{W(B) = I_1(1_B), B \in \mathscr{B}, \nu(B) < \infty\}$ is called the Gaussian random measure on $(\mathscr{X}, \mathscr{B}, \nu)$. The random variable $I_1(\phi)$ is called the Wiener integral of the function ϕ with respect to the Gaussian random measure $W(\cdot)$, and we write symbolically

$$I_1(\phi) = \int_{\mathscr{X}} \phi(x) W(dx).$$

A.2. Let $\mathscr{A}_k^*$ be the space of symmetric functions $h_k(x_1, \cdots, x_k)$ subject to the condition

$$\nu(h_k^2) = \int h_k^2(x_1, \cdots, x_k)\nu(dx_1) \cdots \nu(dx_k) < \infty.$$

The multiple Wiener integral of order k is a *linear* mapping $I_k(\cdot)$ from the space $\mathscr{A}_k^*$ into the space of random variables which are functionals of the Gaussian family $I_1(\phi)$. This mapping is defined uniquely by the two following conditions.

(a) For functions of the form $h_k^\phi = \phi(x_1) \cdots \phi(x_k)$,

$$I_k(h_k^\phi) = (\nu(\phi^2))^{k/2} H_k\left(\frac{I_1(\phi)}{\sqrt{\nu(\phi^2)}}\right)$$

where H_k is the Hermite polynomial of degree k with leading coefficient 1.

(b) $$EI_k(h_k)^2 = k!\nu(h_k^2).$$

We write symbolically

$$I_k(h_k) = \int_{\mathscr{X}} \cdots \int_{\mathscr{X}} h_k(x_1, \cdots, x_k) W(dx_1) \cdots W(dx_k).$$

When it is necessary to make explicit the dependence of I_k on the measure ν we write $I_k^\nu(h_k)$ instead of $I_k(h_k)$. Note that h_k^ϕ satisfies (b), and by polarization, (b) is equivalent to

$$EI_k(h_k)I_k(g_k) = k!\nu(h_k g_k)$$

for h_k and g_k in $\mathscr{A}_k^*$. Furthermore the random variables $I_k(h_k)$ and $I_\ell(h_\ell)$ are orthogonal for $k \neq \ell$ because $EH_k(X)H_\ell(Y) = 0$ for any random variables X and Y which are jointly normal with mean 0.

A.3. Since ν, through $I_1(\phi)$, determines $I_k^\nu(h_k) \equiv I_k(h_k)$, we obtain immediately the following formula for change of variables which was used in the proof of Lemma 4. If

$$\frac{d\nu}{d\mu} = (r(x))^2$$

then

$$I_k^\nu(h_k(x_1, \cdots, x_k))$$

has the same distribution as

$$I_k^\mu(h_k(x_1, \cdots, x_k)r(x_1) \cdots r(x_k)).$$

A.4. Randomization involves expanding the measure space on which the Gaussian random measure is defined without modifying the distributions of the corresponding multiple Wiener integrals.

More precisely, let $(\mathscr{X}, \mathscr{B}, \nu)$ and I_k^ν be as before. Introduce a second measure space $(\mathscr{U}, \mathscr{L}, \mu)$ and let $I_k^{\nu\times\mu}$ be the multiple Wiener integral associated with the product space $(\mathscr{X} \times \mathscr{U}, \mathscr{B} \times \mathscr{L}, \nu \times \mu)$. Then $I_k^{\nu\times\mu}(\tilde{h}_k)$ and $I_k^\nu(h_k)$ have the same distribution if

$$\tilde{h}_k(x_1, u_1; \cdots; x_k, u_k) = h_k(x_1, \cdots, x_k)\alpha(u_1) \cdots \alpha(u_k)$$

where $\nu(h_k^2) < \infty$ and $\mu(\alpha^2) = 1$. Indeed, h_k is approximated by linear combination of the form h_k^ϕ and the random variables $I_1^{\nu\times\mu}(\phi\alpha)$ and $I_1^\nu(\phi)$ are both normal with mean 0 and variance $\nu \times \mu(\phi^2\alpha^2) = \nu(\phi^2)$.

A.5. A functional is a random variable which is measurable with respect to the σ-algebra generated by the Wiener integrals $\{I_1(\phi), \phi \in L^2(\mathscr{X}, \mathscr{B}, \nu)\}$. Every functional X satisfying $EX^2 < \infty$ has a unique representation of the form

$$X = \sum_{k=0}^\infty \frac{1}{k!} I_k(h_k)$$

where $h_k \in \mathscr{H}_k^*$ and $\sum_{k=0}^\infty (1/k!)\nu(h_k^2) < \infty$. Indeed, the expression for the moment generating function of the Hermite polynomials yields

$$e^{I_1(\phi)-(1/2)\nu(\phi^2)} = \sum_{k=0}^\infty \frac{1}{k!} I_k(h_k^\phi)$$

where $h_k^\phi = \phi_1(x_1) \cdots \phi(x_k)$. This is a representation for the functional $X^\phi = e^{I_1(\phi)-(1/2)\nu(\phi^2)}$. The representation for an arbitrary functional X satisfying $EX^2 < \infty$, follows from the fact that the span of the family $\{X^\phi, \phi \in L^2(\mathscr{X}, \mathscr{B}, \nu)\}$ is dense (in the L^2 sense) in the space of square integrable functionals.

A.6. There are several approaches for constructing multiple Wiener integrals. Ours is closest to Neveu [11] and it follows that of [3]. In Itô [6], it is assumed that ν is nonatomic. Itô's approach can be used when ν is atomic by introducing randomization (see for example [7], page 35).

It is also possible to limit the definition of multiple Wiener integrals to $\mathscr{A}_k$—the subspace of $\mathscr{A}_k^*$ which contains only canonical functions—by imposing the condition $\nu(\phi) = 0$ in (a) of A.2. Then, to integrate non-canonical functions one can randomize, as we have done in the proof of Theorem 1*.

REFERENCES

[1] BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[2] DENKER, M., GRILLENBERGER, C. and KELLER, G. (1982). A note on invariance principles for von Mises' statistics. Preprint.

[3] DYNKIN, E. B. and MANDELBAUM, A. (1983). Symmetric statistics, Poisson point processes and multiple Wiener integrals. *Ann. Statist.* **11** 739–745.

[4] HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 239–325.

[5] HOEFFDING, W. (1961). The strong law of large number for U-statistics. Institute of Statistics, Univ. of N. Carolina, Mimeograph Series No. 302.

[6] ITÔ, K. (1951). Multiple Wiener integral. *J. Math. Soc. Japan* **3** 157–169.

[7] MAJOR, P. (1980). Multiple Wiener-Itô integrals. *Lecture Notes in Math* **849**. Springer, Berlin.

[8] MILLER, R. G., JR. and SEN, P. K. (1972). Weak convergence of U-statistics and von Mises' differentiable statistical functions. *Ann. Math. Statist.* **43** 31–41.

[9] MÓRI, T. F. and SZÉKELY, G. J. (1982). Asymptotic behaviour of symmetric polynomial statistics. *Ann. Probab.* **10** 124–131.

[10] NEUHAUS, G. (1977). Functional limit theorems for U-statistics in the degenerate case. *J. Multivariate Anal.* **7** 424–439.

[11] NEVEU, J. (1968). *Processus Aléatoires Gaussiens.* Les Presses de l'Université de Montreal, Montreal.

[12] RUBIN, H. AND VITALE, R. A. (1980). Asymptotic distribution of symmetric statistics. *Ann. Statist.* **8** 165–170.

[13] SEN, P. K. (1982). *Sequential Non-Parametrics: Invariance Principles and Statistical Inference.* Wiley, New York.

[14] SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

[15] VON MISES, R. (1947). On the asymptotic distribution of differentiable statistical functions. *Ann. Math. Statist.* **18** 309–348.

[16] WHITT, W. (1980). Some useful functions for functional limit theorems. *Math. of Op. Research* **5** 67–85.

SCHOOL OF OPERATIONS RESEARCH
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853