The Annals of Mathematical Statistics
1970, Vol. 41, No. 3, 829–842

CONVERGENCE IN DISTRIBUTION OF STOCHASTIC INTEGRALS[1]

BY MARK BROWN

Cornell University

0. Introduction. In this paper, convergence in distribution of sequences of quadratic mean stochastic integrals is studied by developing and extending an elegant approach introduced by J. Sethuraman [14]. Sethuraman's contribution is essentially contained in Theorem 3.1.

A type of convergence of stochastic processes, linear law convergence is introduced. It entails convergence of finite dimensional distributions and a condition on the product moment kernels of the processes. This condition has several equivalent forms and is discussed in Section 3.

Linear law convergence is well suited for deriving convergence in distribution of sequences of random variables $\{W_n, n = 1, 2, \cdots\}$, where $W_n \in L^2\{X_n(t), t \in T\}$. The convergence in distribution is derived without a sample path analysis. In fact, the random variables under consideration may not be pathwise defined. For example $\{W_n\}$ may be a sequence of quadratic mean stochastic integrals with the pathwise integrals not existing. On the other hand many important pathwise defined functionals of a process $\{X\{t\}, t \in T\}$, are not members of $L^2\{X(t), t \in T\}$, and thus not suited to linear law analysis.

Section 1 and Section 2 contain preliminary material on reproducing kernel Hilbert spaces and quadratic mean stochastic integrals. In Section 3, linear law convergence is introduced, and its basic properties derived. Section 4 contains a method by which a sequence of finite collections of random variables may be embedded into a sequence of continuous time processes satisfying the kernel condition for linear law convergence. In Section 5 linear law convergence is related to weak convergence over L^2 and reproducing kernel Hilbert spaces. In Section 6 several applications are derived.

1. Reproducing kernel Hilbert spaces. Let $[X(t)\ t \in T]$ be a complex valued second order stochastic process with product moment kernel K, so that $K(s, t) = E(X(s)\overline{X(t)})$. Let $\tilde{L}^2(X)$ be the set of all finite linear combinations $\sum_{i=1}^m a_i X(t_i)$, and let $L^2(X)$ be the closure of $\tilde{L}^2(X)$ under quadratic mean distance. $L^2(X)$ is a Hilbert space with inner product $(Z_1, Z_2) = E(Z_1\bar{Z}_2)$.

For each $Z \in L^2(X)$, let g_Z be a function over T defined by $g_Z(t) = E(Z\overline{X(t)})$. It is easy to show that the operator A defined by $A(Z) = g_Z$ is one to one. It follows that the set $\{g_Z, Z \in L^2(X)\}$ becomes a Hilbert space under the inner product $(g_{Z_1}, g_{Z_2}) = (Z_1, Z_2)_{L^2(X)}$. Call this Hilbert space H_K. It follows that $L^2(X)$ and H_K

Received September 5, 1968.

[1] Supported in part by the Office of Naval Research under Contract Nonr-225(80) (NR-042-234) at Stanford University.

are congruent (isometrically isomorphic) and that A is one to one, onto, linear and inner product preserving. We will refer to A as the congruence map from $L^2(X)$ to H_K.

Let $K(s, \cdot)$ be the function over T whose value at t is given by $K(s, t)$. For each $s \in T$, $K(s, t) = (X(s), X(t))_{L^2(X)}$, so that $K(s, \cdot) \in H_K$ with $A(X(s)) = K(s, \cdot)$. Moreover, for all $s \in T$, $g \in H_K$, $(g_Z, K(s, \cdot))_{H_K} = (Z, X(s))_{L^2(X)} = g_Z(s)$. Therefore H_K is a reproducing kernel Hilbert space (r.k.H.s.).

This association between the L^2 space spanned by the process and the r.k.H.s. spanned by the kernel of the process was first noted by Loève. Parzen has employed this and other congruences to obtain many interesting time series applications, some of which are found in [10]. A survey of the theory of r.k.H.s.'s may be found in Aronszahn [1].

Lemmas 1.1, 1.2 and 1.3 are stated without proof. Lemma 1.1 may be found in [1] page 344, and 1.2 in [1] page 383. The proof of 1.3 is obvious.

LEMMA 1.1. $\|f_n - f\|^2_{H_K} \to 0$ iff *f_n converges pointwise to f and $\|f_n\|_{H_K} \to \|f\|_{H_K}$. The sequence $\{f_n\}$ converges weakly to f in H_K* iff *f_n converges pointwise to f and* $\sup_n \|f_n\|_{H_K} < \infty$.

LEMMA 1.2. *Let H_K and H_L be* r.k.H.s.'s. *Then the following are equivalent*:

(i) *The operator T defined by $T(\sum_1^m a_i K(t_i, \cdot)) = \sum_1^m a_i L(t_i, \cdot)$ has a unique extension to a bounded linear operator from H_K to H_L.*

(ii) *$H_L \subset H_K$ in the set inclusion sense.*

(iii) $\exists\, m \ni \sum_{i,j=1}^n a_i \bar{a}_j L(t_i, t_j) \leqq m \sum_{i,j=1}^n a_i \bar{a}_j K(t_i, t_j)$ *for all n, $t_1, \cdots, t_n$, $a_1, \cdots, a_n$. We abbreviate this condition by $L \ll mK$.*

When $H_L \subset H_K$, the extension of the operator T defined above, will be referred to as the kernel operator from H_K to H_L. The operator T^* obtained by considering T as an operator from H_K to H_K, will be referred to as the starred kernel operator from H_K to H_L.

COROLLARY 1.1. *If $\{X(t)\ t \in T\}$ and $\{Y(t)\ t \in T\}$ are processes with kernels K and L respectively with $H_L \subset H_K$, then the operator S defined by $S(\sum_1^n a_i X(t_i)) = \sum_1^n a_i Y(t_i)$, has a unique extension to a bounded linear operator from $L^2(X)$ to $L^2(Y)$. Moreover* $\|S\| = \|T\|$.

PROOF. S agrees on a dense set with the bounded linear operator $A_Y^{-1} T A_X$, where A_X is the congruence map from $L^2(X)$ to H_K, A_Y is the congruence map from $L^2(Y)$ to H_L and T is the kernel map from H_K to H_L. The result now easily follows.

Define $H_L \subseteqq H_K$ to mean that $H_L \subset H_K$ and for $f, g \in H_L$, $(f, g)_{H_L} = (f, g)_{H_K}$.

LEMMA 1.3. $H_L \subseteqq H_K$ iff:

(i) $L(s, \cdot) \in H_K\ \forall\ s \in T$

(ii) $(L(s, \cdot), (L(t, \cdot))_{H_K} = L(s, t)$.

Note that if $H_L \subseteqq H_K$ then the kernel operator from H_K to H_L is the projection operator from H_K to H_L.

Let T^* be the starred kernel operator from H_K to H_L (T, considered as an operator from H_K to H_K).

LEMMA 1.4. *If* $H_L \subset H_K$ *then* T^* *is a bounded self adjoint linear operatcr and* $\|T^*\| = \|T\|^2$.

PROOF. It follows from [1] page 383, that the inclusion map from H_L to H_K is a bounded linear operator and is adjoint to T. Thus if $f_n \to f$ in H_K then $Tf_n \to Tf$ in H_L (Lemma 1.2) and $T^*f_n \to T^*f$ in H_K. Thus T^* is bounded. Now $(T^* \sum_1^n a_i K(t_i, \cdot), \sum_1^m b_j K(s_j, \cdot))_{H_K} = (\sum_1^n a_i K(t_i, \cdot), T^* \sum_1^m b_j K(s_j, \cdot))_{H_K}$ for all choices of n,m a's, b's, etc. It follows by continuity of T^* and continuity of the inner product that T^* is self adjoint. By [16] page 335, $\|T^*\| = \sup_{||f||_{H_K}=1} (Tf, f)_{H_K}$. Since T is adjoint to the inclusion map, $\|T\|^2 = \sup_{||f||_{H_K}=1} (Tf, Tf)_{H_L} = \sup_{||f||_{H_K}=1} (Tf, f)_{H_K}$. Thus $\|T^*\| = \|T\|^2$.

LEMMA 1.5. *A process* X *with kernel* K_0 *has a version with paths in the* r.k.H.s. H_K *and with* $E\,||X||_{H_K}^2 < \infty$, iff U^* *the starred kernel operator from* H_K *to* H_{K_0} *is nuclear (compact with finite trace).*

PROOF. Sufficiency follows from [10] page 487 and necessity from [15] page 67.

2. Stochastic integrals. In this paper integrals of the form:

(i) $\int_I g(t)\, dX(t)$

(ii) $\int_I X(t)\, dg(t)$

(iii) $\int_I g(t)\, X(t)\, d\mu(t)$

will be considered. In the above, I is an interval, possibly infinite, X is a random function, and g a non-random function. The notation $J(X, g, I)$ will be generic for integrals of the above form.

The integrals are of Riemann–Stieltjes (R.S.) quadratic mean (q.m.) type (see Loève [9] page 472) and extension by continuity of such integrals. That is, the set G of functions for which the R.S.q.m. integral exists is linear. The integral $J(X, \cdot, I)$ thus can be considered as a linear operator from G to $L^2(X)$. If G is embedded in a Banach space β and $J(X, \cdot, I)$ is a bounded operator relative to β, it then is uniquely extendable to a bounded linear operator from $\bar{G}$ (closure of G in β) to $L^2(X)$. Two examples of such extensions are:

(i) If $\{X(t), t \in I\}$ is an orthogonal increment process then ([4] page 99) there exists a monotone non-decreasing function F, unique up to an additive constant satisfying, $E|X(t)-X(s)|^2 = F(t)-F(s)$. The R.S.q.m. integral $\int_I g(t)\, dX(t)$, has a unique extension to $L^2(I, F)$ ([4] page 426).

(ii) If $\int_{I \times I} K^2(s, t)\, d\mu(s)\, d\mu(t) < \infty$ then the R.S.q.m. integral $\int_I g(t)\, X(t)\, d\mu(t)$, has a unique extension to the closure of G in $L^2(I, \mu)$. This follows by combining [16] page 169, Lemma 1.1 and Lemma 2.1 below.

Define $\int_I g(t)\, dK(t)$, $\int_I K(t)\, dg(t)$, etc., as R.S. integrals and extensions by continuity, where now the limit of approximating sums is taken in the norm topology of H_K. By Lemma 1.1 such integrals are pointwise, so for example, $[\int_I K(t)\, dg(t)](s) = \int_I K(t, s)\, dg(t)$.

Moreover since R.S. approximating sums to integrals involving X are mapped by A into corresponding sums involving K, it follows that $A(J(X, g, I)) = J(K, g, I)$. If the extensions of the R.S. integrals are with respect to the same norm on G (this will always be assumed), then $A(J(X, g, I)) = J(K, g, I)$ for all $g \in \bar{G}$. Thus:

LEMMA 2.1. *If* $\{X(t), t \in I\}$ *is a process with kernel* K, *then* $J(X, g, I)$ *exists iff* $J(K, g, I)$ *exists and* $A(J(X, g, I)) = J(K, g, I)$.

Now consider processes $\{X(t), t \in I\}$ and $\{Y(t), t \in I\}$, with kernels K and L respectively with $H_L \subset H_K$. By Corollary 1.1, S is bounded a linear operator. Therefore $G_X \subset G_Y$, where G_X is the class for which the R.S.q.m. integral $J(X, g, I)$ exists. If $J(X, \cdot, I)$ is extendable to $\bar{G}_X$ with respect to a given norm on G_X, then by continuity of S, $J(Y, \cdot, I)$ is also extendable to $\bar{G}_X$ with respect to this same norm, and $S(J(X, g, I)) = J(Y, g, I)$ for all $g \in \bar{G}_X$. Thus:

LEMMA 2.2. *If* X *and* Y *are processes with kernels* K *and* L *respectively, with* $H_L \subset H_K$, *then for all* $g \in \bar{G}_X$, $J(Y, g, I)$ *exists and* $J(Y, g, I) = S(J(X, g, I))$.

Note that Lemma 2.1 and Lemma 2.2 follow from [8] page 63.

Lemma 2.3 below is found in [9] page 472, and [8] page 63. It holds for R.S.q.m. integrals but not necessarily for extensions of R.S.q.m. integrals.

LEMMA 2.3. $\int_a^b X(t)\,dg(t)$ *exists as an* R.S.q.m. *integral* $\Leftrightarrow \int_a^b g(t)\,dX(t)$ *exists as an* R.S.q.m. *integral,* $\Leftrightarrow \int_{[a,b]\times[a,b]} g(s)\overline{g(t)}\, d_{12}(K(s,t))$ *exists as an* R.S. *integral* $\Leftrightarrow \int_{[a,b]\times[a,b]} K(s,t)\, d_{12}(g(s)\overline{g(t)})$ *exists as an* R.S. *integral.*

3. Linear law convergence. By the distribution of a complex random variable Z we shall mean the distribution of the random vector $\binom{\mathrm{Re}\,(Z)}{\mathrm{Im}\,(Z)}$. The notation $D(Z_n) \to D(Z)$ will mean that the sequence $\{\binom{\mathrm{Re}\,(Z_n)}{\mathrm{Im}\,(Z_n)}\}$ converges in distribution to the random vector $\binom{\mathrm{Re}\,(Z)}{\mathrm{Im}\,(Z)}$; $D(X, Y)$ will denote the distance between the random vectors $\binom{\mathrm{Re}\,(X)}{\mathrm{Im}\,(X)}$ and $\binom{\mathrm{Re}\,(Y)}{\mathrm{Im}\,(Y)}$ under Prokhorov's metric, [12] page 166, specialized to measures in Euclidean two dimensional space. It is proved in [13] page 16 that:

$$D(X, Y) \leqq (E|X-Y|^2)^{\frac{1}{3}}. \tag{1}$$

Let $\{X_n(t), t \in T\}$ $n = 0, 1, \cdots$ be a sequence of complex valued second order processes with product moment kernels $\{K_n\, n = 0, 1, \cdots\}$. Let ψ_n be the congruence map from H_{K_n} to $L^2(X_n)$, $n = 0, 1, \cdots$. Assume $H_{K_n} \subset H_{K_0}$ for all n, and let T_n be the kernel operator from H_{K_0} to H_{K_n}, and $T_n{}^*$ the starred kernel operator from H_{K_0} to H_{K_n}. Let S_n be the unique extension (see Corollary 1.1) of the operator $\tilde{S}_n$ defined on $\tilde{L}^2(X_0)$ by $\tilde{S}_n(\sum_1^m a_i X_0(t_i)) = \sum_1^m a_i X_n(t_i)$.

LEMMA 3.1. *The following four conditions are equivalent.*

(i) $\sup_n \|S_n\| < \infty$.
(ii) $\sup_n \|T_n\| < \infty$.
(iii) $\sup_n \|T_n{}^*\| < \infty$.
(iv) $\sup_n \|T_n{}^* f\|_{H_{K_0}} < \infty$ *for all* $f \in H_{K_0}$.

PROOF. It follows from Lemma 1.4 that (ii) $\Leftrightarrow$ (iii) and from Corollary 1.1 that (i) $\Leftrightarrow$ (ii). An application of the uniform boundedness principle, [16] page 204, shows that (iii) $\Leftrightarrow$ (iv).

Define $X_n \to X_0$ in f.d.d. (finite dimensional distributions) iff for all $Z \in \tilde{L}^2(X_0)$, $D(S_n Z) \to D(Z)$.

Define $X_n \to X_0$ in l.l. (linear law) iff $X_n \to X_0$ in f.d.d. and any of the four equivalent conditions in Lemma 3.1 hold.

THEOREM 3.1 (J. Sethuraman). *If* $\{X_n(t)\ t \in T\} \to X_0(t)\ t \in T\}$ *in* l.l. *then*:

(i) $\|f_n - f\| H_{K_0} \to 0 \Rightarrow D(\psi_n T_n f_n) \to D(\psi_0 f)$.

(ii) $Z_n \to Z$ *in* $L^2(X_0) \Rightarrow D(S_n Z_n) \to D(Z)$.

PROOF. We shall prove (ii), (i) and (ii) obviously being equivalent. Let $M = \sup_n \|S_n\|^2$, and let $Z_n \to Z$ in $L^2(X_0)$. Given $\varepsilon > 0\ \exists \cdot W \in \tilde{L}_2(X_0) \ni E|Z-W|^2 < \varepsilon^3$. By (1) $D(Z, W) < \varepsilon$. Also, $\exists\ N_1 \ni n > N_1 \Rightarrow E|Z_n - W|^2 < 2\varepsilon^3$. Thus $E|S_n Z_n - S_n W|^2 < 2M\varepsilon^3$, thus $D(S_n Z_n, S_n W) < (2M)^{\frac{1}{3}}\varepsilon$. It follows from f.d.d. convergence that for $n > N_2$, $D(W, S_n W) < \varepsilon$. Thus for $n > \max(N_1, N_2)$, $D(Z, S_n Z_n) \leqq D(Z, W) + D(W, S_n W) + D(S_n W, S_n Z_n) \leqq \varepsilon(2 + (2M)^{\frac{1}{3}})$.

THEOREM 3.2. *If* $X_n \to X_0$ *in* l.l. *and* $K_n(s,t) \to K(s,t)$ *for all* s, t, *then* $Z_n \to Z$ *in* $L^2(X_0) \Rightarrow E|S_n Z_n|^2 \to E|Z|^2$.

PROOF. Let $\|Z\| = (E|Z|^2)^{\frac{1}{2}}$. Note that the hypothesis implies that for $W \in \tilde{L}^2(X_0)$, $E|S_n W|^2 \to E|W|^2$. The result follows easily by approximation of Z by $W \in \tilde{L}_2(X_0)$ and use of the inequality:

$$\big|\|Z\| - \|S_n Z_n\|\big| \leqq \|Z - W\| + \big|\|W\| - \|S_n W\|\big| + \|S_n W - S_n Z_n\|.$$

For a partition π, a random function X, and a non-random function g let $R(\pi, X, g)$ be an R.S. approximating sum to $J(X_0, g)$. Define $|\pi|$ to be the length of the maximum subdivision of π.

COROLLARY 3.1. *If* $\{X_n(t),\ t \in [a,b]\} \to \{X_0(t),\ t \in [a,b]\}$ *in* l.l. *and* $J(X_0, g, [a,b])$ *exists as an* R.S.q.m. *integral, then* $|\pi_n| \to 0 \Rightarrow D(R(\pi_n, X_n, g)) \to D(J(X_0, g, [a, b]))$.

PROOF. $S_n(R(\pi_n, X_0, g)) = R(\pi_n, X_n, g)$ and $R(\pi_n, X_0, g) \to_{\text{q.m.}} J(X_0, g)$. The result follows from Theorem 3.1.

Similarly if X_0 is quadratic mean differentiable at t_0 then

$$|h_n| \to 0 \Rightarrow D\left(\frac{X_n(t_0 + h_n) - X_0(t_0)}{h_n}\right) \to D(X_0'(t_0)).$$

COROLLARY 3.2. *If* $\langle X_n(t)\ t \in I\rangle \to \langle X_0(t)\ t \in I\rangle$ *in* l.l. *then* $J(X_0, g_n, I) \to_{\text{q.m.}} J(X_0, g, I) \Rightarrow D(J(X_n, g_n, I)) \to D(J(X_0, g, I))$.

PROOF. By Lemma 2.2, $S_n(J(X_0, g_n, I)) = J(X_n, g_n, I)$.

Lemma 3.2 below provides a few sets of conditions for the q.m. convergence of $J(X, g_n, I)$ to $J(X, g, I)$. These, combined with Corollary 3.2 yield conditions for the convergence in distribution of $J(X_n, g_n, I)$ to $J(X_0, g, I)$, when $X_n \to X_0$ in l.l.

LEMMA 3.2. (i) *If K is continuous on $[a, b] \times [a, b]$ and $\{g_n\}$ is a sequence of complex functions converging weak* on $C^*[a, b]$ to $g(\int_a^b f dg_n \to \int_a^b f dg$ for all continuous complex f), then $\int_a^b X(t)\, dg_n(t) \to_{\text{q.m.}} \int_a^b X(t)\, dg(t)$.*

PROOF. [8] page 65 proves the result in the case where the $\{g_n\}$ sequence is real. The complex case follows easily.

(ii) *If X is real, K is of bounded variation on $[a, b] \times [a, b]$, $\{g_n\}$ is a sequence of complex functions converging uniformly to g on $[a, b]$, and $\int_a^b g_n(t)\, dX(t)$ exists as an* R.S.q.m. *integral for all n, then $\int_a^b g(t)\, dX(t)$ exists as an* R.S.q.m. *integral and* $\int_a^b g_n(t)\, dX(t) \to_{\text{q.m.}} \int_a^b g(t)\, dX(t)$.

PROOF. The result follows by combining [7] page 131, Theorem 9.9, and Lemma 1.1.

(iii) *If G is the set for which $J(X, g, I)$ exists as an* R.S.q.m. *integral, and $J(X, \cdot, I)$ is extendable to a continuous linear operator on $\bar{G}$, the closure of G in a Banach space β, then by definition of continuity $g_n \to g$ in β implies $J(X, g_n, I) \to_{\text{q.m.}} J(X, g, I)$.* In Section 2, two examples of such extensions were given.

THEOREM 3.3. *Assume $X_n \to X_0$ in l.l. Let $\{Z_0(\alpha), \alpha \in A\}$ be an indexed set of elements belonging to $L^2(X_0)$ and $Z_n(\alpha) = S_n(Z_0(\alpha))$, $\alpha \in A$, $n = 1, 2, \cdots$. Then $\{Z_n(\alpha)\ \alpha \in A\} \to \{Z_0(\alpha)\ \alpha \in A\}$ in l.l.*

PROOF. The operator S_n' from $L^2(Z_0)$ to $L^2(Z_n)$, defined analogously to S_n, is the restriction of S_n to $L^2(Z_0)$. Therefore $\|S_n'\| \leqq \|S_n\|$. Also by Theorem 3.1, $D(\sum_1^m a_i Z_n(\alpha_i)) = D(S_n(\sum_1^m a_i Z_0(\alpha_i))) \to D(\sum_1^m a_i Z_0(\alpha_i))$.

EXAMPLES.

(i) $Z_n(\alpha) = J(X_n, \alpha, I)\ \alpha \in A \subset \bar{G}_{X_0}$.

(ii) $Z_n(t) = X_n'(t) \quad n = 0, 1, \cdots \quad t \in A \subset I = T$

where $X_n'(t)$ is the q.m. derivative of X_n at t. Note that $H_{K_n} \subset H_{K_0}$ and the existence of $X_0'(t)$ implies the existence of $X_n'(t)$.

(iii) $Z_n(t) = X_n(t)$, $t \in A \subset T$. Thus if $\{X_n(t)\ t \in T\} \to \{X_0(t)\ t \in T\}$ in l.l. then for any $T' \subset T$, $\{X_n(t)\ t \in T'\} \to \{X_0(t)\ t \in T'\}$ in l.l.

(iv) $Z_n(t) = g(t) X_n(t)$, $t \in T$.

4. Embedding a discrete process into continuous time. Define a kernel K to be two piece linear on $[a, b]$ if whenever $0 \leqq \lambda \leqq 1$ and either $a \leqq \alpha \leqq \min(\beta, \gamma) \leqq \max(\beta, \gamma) \leqq b$ or $a \leqq \min(\beta, \gamma) \leqq \max(\beta, \gamma) \leqq \alpha \leqq b$:

$$K(\alpha, \lambda\beta + (1-\lambda)\gamma) = \lambda K(\alpha, \beta) + (1-\lambda) K(\alpha, \gamma).$$

THEOREM 4.1. *Let $\{Y(t), t \in A = \{a = t_1 < t_2 \cdots < t_n = b\}\}$ be a finite collection of random variables. Assume that for all $i, j = 1, 2, \cdots n$:*

$$E(Y(t_i)\overline{Y(t_j)}) = K(t_i, t_j)$$

where K is a kernel on $[a, b] \times [a, b]$ of the form:

$$K(s, t) = g(s)\overline{g(t)} K_0(F(s), F(t))$$

where K_0 is two piece linear, F is monotone non-decreasing and g is non-vanishing, except perhaps at points t_i for which $Y(t_i) = 0$ a.s.

Define
$$Z(t) = g(t)\left[\frac{Y(t_i)}{g(t_i)} + \frac{F(t)-F(t_i)}{F(t_{i+1})-F(t_i)}\left(\frac{Y(t_{i+1})}{g(t_{i+1})} - \frac{Y(t_i)}{g(t_i)}\right)\right]$$

for $a \leqq t \leqq b$. We interpret $0/0 = 0$ in the above.

Let L be the product moment kernel of Z. Then $H_L \subset H_K$ and the kernel map from H_K to H_L has norm 1.

PROOF. Let $g \equiv 1$. Note that:

$$L(s, t) = L(t_i, t) + \frac{F(s)-F(t_i)}{F(t_{i+1})-F(t_i)}(L(t_{i+1}, t) - L(t_i, t)) \tag{2}$$

where $t_i \leqq s < t_{i+1}$

$$L(t_i, t) = \frac{F(t_{j+1})-F(t)}{F(t_{j+1})-F(t_j)} K(F(t_i), F(t_j)) + \frac{F(t)-F(t_j)}{F(t_{j+1})-F(t_j)} K(F(t_i), F(t_{j+1})) \tag{3}$$

where $t_j \leqq t < t_{j+1}$ letting $\lambda = [F(t_{j+1})-F(t)]/[F(t_{j+1})-F(t_j)]$, $\alpha = F(t_i)$, $\beta = F(t_j)$, $\gamma = F(t_{j+1})$, and applying the definition of two piece linearity in (3) it follows that:

$$L(t_i, t) = K(t_i, t) \quad \text{and} \tag{4}$$

$$L(t_{i+1}, t) = K(t_{i+1}, t). \tag{5}$$

Substituting (4) and (5) in (2):

$$L(s, \cdot) = K(t_i, \cdot) + \frac{F(s)-F(t_i)}{F(t_{i+1})-F(t_i)}(K(t_{i+1}, \cdot) - K(t_i, \cdot))$$

Thus $L(s, \cdot) \in H_K$ for all s. Also:

$$(L(s, \cdot), L(t, \cdot))_{H_K} = L(s, t) \qquad \text{for all} \quad s, t.$$

By Lemma 1.3, $H_L \subseteqq H_K$. Therefore the norm of the kernel operator is the norm of the projection operator from H_K to H_L, and thus equals 1. The result for general g follows from the case $g \equiv 1$ by example (iv) following Theorem 3.3. □

Given a sequence of finite collections of random variables $\{Y_n(t), t \in A_n = \{a = t_1{}^n, \cdots t_{m_n}^n = b\}\}$ $n = 1, 2, \cdots$, Theorem 4.1 assures that if there exists a kernel K of the form specified in Theorem 4.1 such that:

$$E(Y_n(s)\overline{Y_n(t)}) = K(s, t) \qquad \text{for all} \quad n, (s, t) \in A_n \times A_n,$$

then the processes can be embedded into a sequence of continuous processes such that the norms of the kernel operators are uniformly bounded.

More generally we may have $E(Y_n(s)\overline{Y_n(t)}) = K_n(s, t)$, where for each n, K_n is a kernel of the form specified in Theorem 4.1. Then if we embed each process in the manner specified in Theorem 4.1, obtaining a sequence $\{H_{L_n}\}$ of r.k.H.s.'s, Theorem 4.1 assures that any kernel H_{K_0} uniformly dominating ($\sup_n \|T_n\| < \infty$), the family $\{H_{K_n}, n = 1, 2, \cdots\}$, will also dominate $\{H_{L_n}, n=1, 2, \cdots\}$.

Note that in the above embedding procedure, $Z(t) = Y(t)$ for $t \in A$, so that the initial random variables are unchanged. Also, suppose that:

$$\sup_{1 \leqq i \leqq m_n} \Pr(|Y_n(t_i^{(n)}) - Y_n(t_{i-1}^{(n)})| > \varepsilon) \to 0 \tag{6}$$

as $n \to \infty$ for all $\varepsilon > 0$. Let $Y_n^*(t) = Y_n(t_i^{(n)})$ for $t_i^{(n)} \leqq t < t_{i+1}^{(n)}$. Then if $Y_n^* \to Y_0$ in f.d.d. then for any monotone non-decreasing F, $Y_n^{**} \to Y_0$ in f.d.d., where:

$$Y_n^{**}(t) = Y_n(t_i^{(n)}) + \frac{F(t) - F(t_i^{(n)})}{F(t_{i+1}^{(n)}) - F(t_i^{(n)})} (Y_n(t_{i+1}^{(n)}) - Y_n(t_i^{(n)})).$$

Thus if (6) holds the choice of F does not affect f.d.d. convergence.

The class of kernels of the form specified in Theorem 4.1 includes many familiar kernels. $K(s,t) = \min(s,t) - st$, $s, t \in [0, 1]$, is two piece linear as is $K(s,t) = \min(s,t)$. Kernels of the form $K(s,t) = \min(F(s), F(t))$ on $[a, b] \times [a, b]$, characterize the class of orthogonal increment processes, X, on $[a, b]$, with $X(a) = 0$ a.s. Kernels of the form $g(s)\overline{g(t)} \min(F(s), F(t))$ on $[a, b] \times [a, b]$, characterize the class of L^2 Markov processes, X, on $[a, b]$, with $X(a) = 0$ a.s.

5. Relation to weak convergence. We shall say that $X_n \to X_0$ weakly on a Hilbert space H if each X_n, $n = 0, 1, \cdots$, induces a measure P_n on H and $P_n \to P_0$ weakly on H.

THEOREM 5.1. *If $X_n \to X_0$ in l.l. and K is a kernel with the property that the starred kernel operator from H_K to H_{K_0} is nuclear, then $X_n \to X_0$ weakly on H_K.*

PROOF. We will use a theorem, [12] page 171, which gives sufficient conditions for the weak convergence of a sequence of measures on a separable Hilbert space. The space H_K need not be separable, but we will show that our hypothesis implies that there exists a separable subspace of H_K on which every distribution is concentrated. The weak convergence of the measures on the subspace easily implies the weak convergence on H_K.

Let U_n^* be the starred kernel operator from H_K to H_{K_n} $n = 0, 1, \cdots$. Now $U_n^* = T_n^* U_0^*$, and the product of a nuclear operator and a bounded linear operator is nuclear, [5] page 40. Therefore each U_n^* is nuclear.

Let $\mathscr{M}_n$ be the eigenmanifold of U_n^*, $n = 0, 1, \cdots$. By Lemma 1.3, since $\mathscr{M}_0$ is a subspace of H_K, $\mathscr{M}_0$ is an r.k.H.s. with reproducing kernel L, where $L(t, \cdot) = P_{\mathscr{M}_0}(K(t, \cdot))$, the projection of $K(t, \cdot)$ on $\mathscr{M}_0 \cdot \mathscr{M}_0$ is separable and contains H_{K_0}, [16] page 336.

Let $\mathscr{N}_n$ be the null manifold of U_n^*, $n = 0, 1, \cdots$. By [16] page 339, $\mathscr{N}_n = \mathscr{M}_n^{\perp}$. Since $U_n^* = T_n^* U_0^*$, $\mathscr{N}_n \supset \mathscr{N}_0$ and thus $\mathscr{M}_n \subset \mathscr{M}_0$, $n = 1, 2, \cdots$. Thus for all $n, f \in H_K$, $U_n^* f = U_n^* P_{\mathscr{M}_0} f$. It follows that for each n, the starred kernel operator, R_n^*, from $\mathscr{M}_0$ to H_{K_n}, is the restriction of U_n^* to $\mathscr{M}_0$. Thus R_n^* has the same eigenvalues as U_n^* and is therefore a nuclear operator. Therefore all the processes have versions with paths in the separable r.k.H.s. $\mathscr{M}_0$ and $E\|X_n\|_{\mathscr{M}_0}^2 < \infty$.

Now

$$S_n[(X_0, \sum_{i=1}^m a_i L(t_i, \cdot))_{\mathcal{M}_0}] = S_n[\sum_1^m \bar{a}_i X_0(t_i)]$$
$$= \sum_1^m \bar{a}_i X_n(t_i) = (X_n, \sum_{i=1}^m a_i L(t_i, \cdot))_{\mathcal{M}_0}.$$

Moreover $(X_0, \cdot)_{\mathcal{M}_0}$ is a bounded linear operator from $\mathcal{M}_0$ to $L^2(X_0)$, since $E|(X_0, f)_{\mathcal{M}_0}|^2 \leqq E\|X_0\|_{H_K}^2 \|f\|_{\mathcal{M}_0}^2$. Therefore $S_n(X_0, f)_{\mathcal{M}_0} = (X_n, f)_{\mathcal{M}_0}$, thus $D[(X_n, f)_{\mathcal{M}_0}] \to D[(X_0, f)_{\mathcal{M}_0}]$ for all $f \in \mathcal{M}_0$. Therefore the characteristic functional of X_n converges to the characteristic functional of X_0 as $n \to \infty$.

Let $\{e_i\ i = 1, 2, \cdots\}$ be an orthonormal basis for $\mathcal{M}_0$. Let $r_N^{(n)} = \sum_N^\infty (X_n, e_i)_{\mathcal{M}_0} e_i$, $n = 0, 1, \cdots, N = 1, 2, \cdots$. Then $S_n(r_N^{(0)}) = r_N^{(n)}$. Thus,

$$\sup_n E|r_N^{(n)}|^2 \leqq (\sup_n \|S_n\|^2) E|r_N^{(0)}|^2 \to 0 \quad \text{as} \quad N \to \infty.$$

This result combined with the convergence of the characteristic functionals imply that $X_n \to X_0$ weakly on $\mathcal{M}_0$, [12] page 171. It follows immediately from the definition of weak convergence that $X_n \to X_0$ weakly on $\mathcal{M}_0$ implies $X_n \to X_0$ weakly on H_K. □

THEOREM 5.2. *If*:

(i) $\{X_n(t), t \in I\} \to \{X_0(t)\ t \in I\}$ *in* l.l.

(ii) $\int_I K_0(t, t)\, d\mu(t) < \infty$.

(iii) X_n *has a measurable version*, $n = 0, 1, \cdots$.

(iv) *The set* G_0, *for which* $\int_I g(t)\, dX_0(t)$ *exists as an* R.S.q.m. *integral, is dense in* $L^2(I, \mu)$.

Then $X_n \to X_0$ *weakly on* $L^2(I, \mu)$.

PROOF. It follows from (i) and (ii) that $\int K_n(t, t)\, d\mu(t) < \infty$ for all n. This and (iii) assure ([15] page 37) that all processes have versions with paths in $L^2(I, \mu)$ and with $E\,||X_n||_{L^2(I,\mu)}^2 < \infty$.

For each n, $g \in L^2(I, \mu)$, consider the random variable (X_n, g) defined pathwise, $(X_n, g)(w) = \int_I X_n(t, w) \overline{g(t)}\, d\mu(t)$. Since $E|(X_n, g)|^2 \leqq (\int_I K_n(t, t)\, d\mu(t))\, ||g||_{L^2(\mu)}^2$, $(X_n, \cdot)$ is a bounded linear operator on $L^2(I, \mu)$. Moreover for $g \in G_0$, $(X_n, g) = \int X_n(t) \overline{g(t)}\, d\mu(t)$, the R.S.q.m. integral, [4] page 64. Thus the R.S.q.m. integral coincides with the pathwise integral on $\bar{G}_0$, and $\bar{G}_0 = L^2(I, \mu)$ by hypothesis.

Since $S_n(X_0, g) = (X_n, g)$ (Lemma 2.2) it follows that the characteristic functionals converge. The remainder of the argument follows exactly as in Theorem 5.1. □

6. Applications.

6.1. *Wide sense stationary processes.* $\{X(t), -\infty < t < \infty\}$ is called wide sense stationary if $E(X(s)\overline{X(t)}) = R(t-s)$. The function R can be expressed as $R(v) = \int e^{iv\lambda}\, dF(\lambda)$, where F is a bounded, monotone non-decreasing function. The process X has a spectral representation $X(t) = \int e^{it\lambda}\, dZ(\lambda)$, where Z is a uniquely determined process with the class of Borel sets as its index set. Z satisfies $E(Z(B_1)\bar{Z}(B_2)) = F(B_1 \cap B_2)$ for all Borel sets B_1, B_2. There is a unique process $\{Z^*(t), -\infty < t < \infty\}$ which satisfies the requirements of being right continuous in q.m., q.m. convergent to 0 at $-\infty$, and $Z^*(b)-Z^*(a) = Z(a, b]$ for all a, b on the real line. Z is an orthogonal increment process. See Doob [4] page 527 for details of the above.

For a finite interval $[a, b]$ define $H[a, b]$ to be the r.k.H.s. of functions f on $[a, b]$ satisfying: $f(a) = 0$, $f(t) = \int_a^t f'(s)\,ds$ with $\int_a^b |f'(s)|^2\,ds < \infty$. The inner product is defined by, $(f, g) = \int_a^b f'(s)\overline{g'(s)}\,ds$. The reproducing kernel for $H[a, b]$ is $K(s, t) = \min(s, t)-a$. This is the kernel of a Wiener process on $[a, b]$ with $X(a) \equiv 0$.

If $\int \lambda^2\, dF(\lambda) < \infty$ then by [4] page 536, X is quadratic mean differentiable at all $t \in (-\infty, \infty)$, and any separable version of X has almost all its paths absolutely continuous. Moreover, the process of quadratic mean derivatives $\hat{X}$ and the process of derived sample paths X' $(X(t, w) = \int_{-\infty}^t X'(s, w)\,ds)$ are versions of one another. $\hat{X}$ and X' are q.m. continuous, wide sense stationary processes.

THEOREM 6.1. *Let* $\{X_n(t), -\infty < t < \infty\}$ $n = 0, 1, \cdots$, *be a sequence of* q.m. *continuous wide sense stationary processes with corresponding* F_n, Z_n, Z_n^*, R_n, *as described above. Suppose that* $X_n \to X_0$ *in* f.d.d., *that* $F_n \ll F_0$ *for all* n, *and* $\sup_{n,x} dF_n(x)/dF_0 < \infty$. *Then*:

(i) $X_n \to X_0$ *in* l.l.

(ii) $X_n \to X_0$ *weakly in* $L^2(I, \mu)$ *for any* μ, I, *such that* $\mu(I) < \infty$.

(iii) *If* $\int \lambda^2\, dF_0(\lambda) < \infty$ *then* $X_n' \to X_0'$ *in* l.l. *and weakly on* $L^2(I, \mu)$ *where* $\mu(I) < \infty$. *Moreover, for any finite* a, b, $\{X_n(t)-X_n(a), a \leqq t \leqq b\} \to \{X_0(t)- X_0(a), a \leqq t \leqq b\}$ *weakly on* $H[a, b]$.

(iv) $Z_n \to Z_0$ *in* l.l.; $Z_n^* \to Z_0^*$ *in* l.l. *and weakly on* $L^2(I, \mu)$, *for any* I, μ *satisfying* $\int_I F_0(s)\,d\mu(s) < \infty$.

PROOF. (i) $H_{K_n} = \{f : f(t) = \int e^{-it\lambda} g_f^{(n)}(\lambda)\, dF_n(\lambda)$, with $\int |g_f^{(n)}(\lambda)|^2\, dF_n(\lambda) < \infty\}$, with inner product $(f_1, f_2)_{H_{K^n}} = \int g_{f_1}^{(n)}(\lambda)\overline{g_{f_2}^{(n)}}(\lambda)\, dF_n(\lambda)$. This can be seen by building up from $R_n(s-t) = \int e^{-it\lambda} e^{is\lambda}\, dF_n(\lambda)$. Moreover, $(T_n^* f)(t) = \int e^{-it\lambda} g_f^{(0)}(\lambda)\, dF_n(\lambda) = \int e^{-it\lambda} g_f^{(0)}(\lambda)[dF_n(\lambda)/dF_0]\, dF_0(\lambda)$. Thus $\|T_n^* f\|_{H_{K_0}}^2 \leqq (\sup_{n,\lambda} dF_n(\lambda)/dF_0)^2 \|f\|_{H_{K_0}}^2$.

(ii) This result follows directly from Theorem 5.2. Condition (iv) of Theorem 5.2 holds because for any continuous g, $\{X_0(t)g(t), -\infty < t < \infty\}$ is q.m. continuous, and thus by [8] page 63, $\int_b^a g(t)X_0(t)\,dt$ exists as an R.S.q.m. integral. Condition (iii) of Theorem 5.2 holds because any q.m. continuous process possesses a measurable version, [15] page 34.

(iii) It follows from Example (ii) following Theorem 3.3 that $X_n' \to X_0'$ in l.l. Since the X_n' are q.m. continuous wide sense stationary processes, it follows from part (ii) of this theorem that $X_n' \to X_0'$ weakly on $L^2(I, \mu)$ when $\mu(I) < \infty$. The weak convergence of $\{X_n(t)-X_n(a)\}$ to $\{X_0(t)-X_0(a)\}$ on $H[a, b]$ follows from Lemma 1.5, Theorem 5.1, and part (i) of this theorem. Note that $E\|X_0\|_{H[a,b]}^2 = (b-a)\int \lambda^2\, dF_0(\lambda)$.

(iv) By [10] page 294, for each n there exists a congruence J_n from $L^2((-\infty, \infty), F_n)$ to $L^2(X_n)$ satisfying:

$$J_n(I_B) = Z_n(B), \qquad \text{for all indicator functions of Borel sets.} \tag{7}$$

$$J_n(e^{it(\cdot)}) = X_n(t), \quad \text{all} \quad t. \tag{8}$$

Now (8) implies that $J_n f = S_n J_0 f$ for all f in $L^2((-\infty, \infty), F_0)$. Therefore by (7), $S_n(Z_0(B)) = Z_n(B)$. By Theorem 3.3, $Z_n \to Z_0$ in l.l. Linear law convergence of $Z_n{}^*$ to $Z_0{}^*$ follows similarly. Weak convergence of $Z_n{}^*$ to $Z_0{}^*$ on $L^2(I, \mu)$ follows from Theorem 5.2.

Linear operations on wide sense stationary processes are of considerable importance in time series analysis. Every linear process derived from a wide sense stationary process X, with spectral distribution function F, and spectral representation $X(t) = \int e^{it\lambda}\, dZ(\lambda)$, can be expressed ([4] page 534) by:

$$\tilde{X}(t) = \int e^{it\lambda} G(\lambda)\, dZ(\lambda)$$

where $G(\lambda)$ (the gain of the operation) is a member of $L^2((-\infty, \infty), F)$. It follows from Corollary 3.2 and Theorem 3.3, that if $X_n \to X_0$ in l.l., then for all $G \in L^2((-\infty, \infty), F_0)$,

$$\{\tilde{X}_n(t) = \int e^{it\lambda}\, G(\lambda)\, dZ_n(\lambda),\ t \in (-\infty, \infty)\} \to \{\tilde{X}_0(t) = \int e^{it\lambda}\, G(\lambda)\, dZ_0(\lambda)\}$$

in l.l. It then follows from Theorem 6.1 that $\tilde{X}_n \to \tilde{X}_0$ weakly on $L^2(I, \mu)$ where $\mu(I) < \infty$. Also if $\int \lambda^2 |G(\lambda)|^2\, dF_0(\lambda) < \infty$, then $\{\tilde{X}_n(t) - \tilde{X}_n(a),\ a \leqq t \leqq b\} \to \{\tilde{X}_0(t) - \tilde{X}_0(a),\ a \leqq t \leqq b\}$ weakly on $H[a, b]$.

6.2. *Linear combinations of order statistics.* Let $Y_n(j/n+1) = n^{\frac{1}{2}}(U_{j,n} - j/(n+1))$, $j = 0, 1, \cdots, n+1$, where $U_{j,n}$ is the jth order statistic from a sample of size n, uniformly distributed on $[0, 1]$, $U_0 \equiv 0$, $U_{n+1} \equiv 1$.

Define $X_n(t) = Y_n(j/(n+1)) + (n+1)(t - j/(n+1))(Y_n((j+1)/(n+1)) - Y_n(j/(n+1)))$ for $j/(n+1) \leqq t < (j+1)/(n+1)$, $0 \leqq t \leqq 1$. Let $\{X_0(t)\ t \in [0, 1]\}$ be a zero mean real normal process with covariance kernel $K_0(s, t) = s(1-t)$, $0 \leqq s \leqq t \leqq 1$. By [3] page 56, $X_n \to X_0$ in f.d.d. Moreover $K_n(i/(n+1), j/(n+1)) = n/(n+2)\, K_0(i/(n+1), j/(n+1))$. It thus follows from Theorem 4.1 that $X_n \to X_0$ in l.l.

It follows from Corollary 3.1 that if $\int_{[0,1]\times[0,1]} J(s) K_0(s, t) J(t)\, ds\, dt$ exists as an R.S. integral then:

$$D(n^{-\frac{1}{2}} \textstyle\sum_1^n J(t_{j,n})(U_{j,n} - j/n+1)) \to D(\int_0^1 J(t) X_0(t)\, dt \tag{9}$$
$$= N(0, \textstyle\int_{[0,1]\times[0,1]} J(s) K_0(s, t) J(t)\, ds\, dt),$$

where $j/(n+1) \leqq t_{j,n} < (j+1)/(n+1)$.

The class of functions for which the above integral exists is larger than the class of Riemann integrable functions on [0, 1], and smaller than the class of improper Riemann integrable functions. An alternate expression for the variance is given by:

$$\int_0^1 (\tilde{J}(s) - \tilde{\tilde{J}}(1))^2\, ds = \int_0^1 \tilde{J}^2(s)\, ds - [\int_0^1 \tilde{J}(s)\, ds]^2$$

where $\tilde{J}(s) = \int_0^s J(\mu)\, d\mu$, $\tilde{\tilde{J}}(1) = \int_0^1 \tilde{J}(s)\, ds$.

The above result with $t_{j,n} = j/(n+1)$ is equivalent to that obtained in [3] in the case of finite limiting variance. However, by suitably normalizing, they obtain conditions for asymptotic normality which can hold when the limiting variance is infinite. Our approach breaks down in this case.

Another application of Corollary 3.1 is to successive differences. If

$$\int_{[0,1]\times[0,1]} J(s)J(t)\,d_{12}K_0(s,t)$$

exists as an R.S. integral, then:

$D(\sum_{j=0}^{n} J(t_{j,n})(X_n((j+1)/(n+1))-X_n(j/(n+1)))) \to D(\int_0^1 J(t)\,dX_0(t)) = N(0,\sigma^2)$, where $\sigma^2 = \int_{[0,1]\times[0,1]} J(s)J(t)\,d_{12}K_0(s,t) = \int_{\mu=0}^{1}(J(\mu)-\tilde{J}(1))^2\,d\mu$.

An application of Corollary 3.2 and Lemma 3.2 part (i) is the following: Consider $V_n = n^{-\frac{1}{2}}\sum_{j=1}^{n} c_{j,n}(U_{j,n}-j/(n+1))$. Let $f_n(t) = \sum_{j=1}^{[(n+1)t]} c_{j,n}$. Then if $f_n \to f$ weak* in C^* [0, 1], then $D(V_n) \to D(\int_0^1 X_0(t)\,df(t)) = N(0,\sigma^2)$, where $\sigma^2 = \int_{[0,1]\times[0,1]} K_0(s,t)\,df(s)\,df(t)$.

Consider the statistic $n^{-\frac{1}{2}}\sum_{j=1}^{n} J(t_{j,n})[H(U_{j,n})-H(j/(n+1))] = A_n+B_n$ where:

$$A_n = n^{-\frac{1}{2}}\sum_{j=1}^{n} J(t_{j,n})H'(j/(n+1))(U_{j,n}-j/(n+1))$$

$$B_n = n^{-\frac{1}{2}}\sum_{j=1}^{n} J(t_{j,n})[H(U_{j,n})-H(j/(n+1))-H'(j/(n+1))(U_{j,n}-j/(n+1))]$$

where $j/(n+1) \leqq t_{j,n} \leqq (j+1)/(n+1)$. We assume that H' exists at all but a finite number of points. By (9), if $\int J(s)H'(s)K_0(s,t)H'(t)J(t)\,ds\,dt$ exists as an R.S. integral, then $D(A_n) \to D(\int_0^1 J(s)H'(s)X_0(s)\,ds)$. It follows that if $B_n = op(1)$, then $D(A_n+B_n) \to D(\int_0^1 J(s)H'(s)X_0(s)\,ds)$. Conditions for $B_n = op(1)$ can be obtained from the large sample analysis in [3] page 57, for similar remainder terms.

Since G_0 contains the continuous functions on [0, 1], Condition (iv) of Theorem 5.2 holds. Condition (i) has been proved, (ii) is trivial, and (iii) follows from [15] page 34. Therefore $X_n \to X_0$ weakly on $L^2[0,1]$. Moreover, let

$$Z_n(t) = X_n(j/(n+1)), \qquad j/(n+1) \leqq t < (j+1)/(n+1), \qquad 0 \leqq t \leqq 1 \qquad \text{then}$$

$$E\|X_n - Z_n\|^2_{L^2[0,1]} \to 0 \qquad \text{as} \quad n \to \infty. \tag{10}$$

Weak convergence of X_n to X_0 and (10) imply by [2] page 25, that $Z_n \to X_0$ weakly on $L^2[0,1]$. Therefore, for example:

$$D((n+1)^{-1}\sum_{i=1}^{n} |X_n(j/(n+1))|^p) \to D(\int_0^1 |X_0(t)|^p\,dt)$$

for $0 \leqq p \leqq 2$. For $p = 2$ the associated statistic differs by $op(1)$ from the Cramér–von Mises statistic.

6.3. *Convergence of counting processes to a Poisson process.* Consider the sequence of centered counting processes: $X_n(t) = n(F_{nn}^{(t)} - F_n(t))$, $t \in [0,\infty)$, $n = 1, 2, \cdots$; where F_{nn} is the empirical cdf of a sample of size n with cdf F_n. Assume that $\lim_{n\to\infty} F_n(t) = 0$ for all t and that $\lim_{n\to\infty} nF_n(t) = F(t)$ for all t. Also assume that $F_n \ll F$ with $\sup_{n,x}(n(dF_n(X)/dF)) = C < \infty$. Let $\{X_0(t)\ t \in [0,\infty)\}$ be a centered nonhomogeneous Poisson process with $E(X_0(t)-X_0(s))^2 = F(t)-F(s)$. The f.d.d. convergence of X_n to X_0 is easy to show. The uniform boundedness of the kernel operators follows from an argument similar to that found in Section 6.1. Thus $X_n \to X_0$ in l.l. Let $Y_n(t) = X_n(t) + nF_n(t)$. Consider:

$$\sum_1^{Y_n(T)} J(X_{n,i}), \tag{11}$$

for J such that $\int_{[0,T]\times[0,T]} J(s)J(t)\, d_{12}(K_0(s,t))$ exists as an R.S. integral. Now

(12) $$\sum_1^{Y_n(T)} J(X_{n,i}) - \int_0^T J(s)\, dF_n(s) = \int_0^T J(s)\, dX_n(s) \qquad \text{a.s.}$$

By Corollary 3.2, $D(\int_0^T J(s)\, dX_n(s)) \to D(\int_0^T J(s)\, dX_0(s))$. The limiting random variable has a centered compound Poisson distribution with characteristic function, $Q(\mu) = \exp[\int_0^T [e^{i\mu J(t)} - 1 - i\mu J(t)]\, dF(t)]$.

The statistic (11) has the following interpretation. Suppose at time 0 there are n items with i.i.d. failure distribution F_n. When an item fails at time t, a cost of $J(t)$ is incurred. Then (11) represents the total cost due to failures in $[0, T]$.

Frequently in nuclear chemistry counting situations there are a large number of particles with long half-lives, so that during the counting period only a moderate number are expected to decay. In this case the distribution of the life of a particle is exponential with parameter λ. The approximating Poisson would have $F(t) = n(1-e^{-\lambda t}) \approx n\lambda t$.

Let $\{X_n(t), -\infty < t < \infty\}$ be a sequence of counting processes, with the property that with probability one, there are a finite number of events in every finite interval. For $t > 0$, $X_n(t)$ represents the number of events in $[0, t]$, and for $t < 0$, the number of events in $[t, 0)$. Suppose that $X_n \to X_0$ in l.l. Let $T_{n,i}$ $i = \pm 1, \cdots, n = 0, 1, \cdots$, be the ith arrival time of the nth process. Let w be a real function with the property that $\int_{(-\infty,\infty)\times(-\infty,\infty)} w(t-\mu)w(t-v)\, dK_0(\mu, v)$ exists as an R.S. integral for all t. Define, $Y_n(t) = \sum_{i=-\infty}^{\infty} w(t - T_{n,i})$. Then, $Y_n(t) = \int_{-\infty}^{\infty} w(t-s)\, dX_n(s)$. By Example (i) following Theorem 3.3, $Y_n \to Y_0$ in l.l.

The Y_n processes are called shot noise processes and are discussed in [11] pages 149–151. These processes arise in the theory of noise in physical devices.

6.4. *Generalized random functions.* A generalized random function, defined and discussed in [5] pages 231–254, is a stochastic process, $\{Y(\phi), \phi \in K\}$, where K is the linear topological space of infinitely differentiable functions with bounded support ([5] page 18), and Y is linear and continuous in law on K.

Convergence in f.d.d. of a sequence $\{Y_n\}$ of generalized random functions reduces to $D(Y_n(\phi)) \to D(Y_0(\phi))$ for all $\phi \in K$. If $Y_n \to Y_0$ in l.l., then $E|Y_0(\phi_n) - Y_0(\phi)|^2 = K_0(\phi - \phi_n, \phi - \phi_n) \to 0$ implies $D(Y_n(\phi_n)) \to D(Y_0(\phi))$.

An important class of generalized random processes, Y, have the property that $Y(\phi) \in L^2(X)$ for some process X. For example, the Wiener process which is neither pathwise nor q.m. differentiable, has a derivative as a generalized random process given by $X'(\phi) = -\int \phi'(t)X(t)\, dt$. Given an l.l. convergent sequence of processes $\{X_n(t), t \in T\}$ $n = 1, 2, \cdots$, to $\{X_0(t), t \in T\}$, then any corresponding sequence of generalized random functions $\{Y_n(\phi), \phi \in K\}$, with $Y_n(\phi) = S_n(Y_0(\phi))$, is l.l. convergent to $\{Y_0(\phi), \phi \in K\}$ (Theorem 3.3).

Acknowledgment. Much of the material in this paper is from the author's doctoral thesis at the Department of Statistics at Stanford. The author gratefully acknowledges the inspiring guidance of Professor Emanuel Parzen, both as an advisor and teacher.

REFERENCES

[1] ARONSZAJN, N. (1950). Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **68** 337–404.

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

[3] CHERNOFF, H., GASTWIRTH, J. L. and JOHNS, M. V. (1967). Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation. *Ann. Math. Statist.* **38** 52–73.

[4] DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

[5] GELFAND, I. M. and VILENKIN, N. Y. (1964). *Generalized Functions* **4**, English trans. Academic Press, New York.

[6] GNEDENKO, B. V. and KOLMOGOROV, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables*, English trans. Addison Wesley, Reading.

[7] HILDEBRANDT, T. H. (1963). *Introduction to the Theory of Integration*. Academic Press, New York.

[8] HILLE, E. and PHILLIPS, R. S. (1957). *Functional Analysis and Semi-Groups*. American Mathematical Society, Rhode Island.

[9] LOÈVE, M. (1963). *Probability Theory*. Van Nostrand, Princeton.

[10] PARZEN, E. (1959). Statistical inference on time series by Hilbert space methods, I. *Time Series Analysis Papers*. Holden Day, San Francisco, 251–383.

[11] PARZEN, E. (1962). *Stochastic Processes*. Holden Day, San Francisco.

[12] PROHOROV, YU., V. (1956). Convergence of random processes and limit theorems in probability theory. *Theor. Probability Appl.* **1** 157–214.

[13] ROSÉN, B. (1969). Asymptotic normality of sums of random elements with values in a real separable Hilbert space. Unpublished manuscript.

[14] SETHURAMAN, J. (1965). Limit distributions for stochastic integrals. Stanford Univ. Technical Report.

[15] SETHURAMAN, J. (1965). Limit theorems for stochastic processes. Stanford Univ. Technical Report.

[16] TAYLOR, A. E. (1958). *Introduction to Functional Analysis*. Wiley, New York.