[4] E. V. KRICHAGINA, R. SH. LIPTSER, AND A. A. PUKHAL'SKII, *Diffusion approximation of systems with arrival depending on queue, and with arbitrary servicing*, Theory Probab. Appl., 33 (1988), pp. 114–124.

[5] M. LOÈVE, *Probability Theory*, 3rd ed., Van Nostrand, Princeton, NJ, 1963.

[6] A. PLUCIŃSKA, *On stochastic process determined by the conditional expectation and the conditional variance*, Stochastics, 2 (1983), pp. 115–129.

[7] ———, *A characterization of Gaussian distribution based on the Lyapunov condition*, Theory Probab. Appl., 29 (1984), pp. 591–595.

ON HILBERT-SPACE-VALUED U-STATISTICS*

M. L. PURI† AND V. V. SAZONOV

(*Submitted by Yu. V. Prokhorov*)

In this note we show that an elementary argument similar to the one used in [1] for real-valued U-statistics can be applied to Hilbert-space-valued U-statistics to obtain the rate of convergence of their distributions to their Gaussian limits.

Let $\{X_i,\ i \geqq 1\}$ be a sequence of independent and identically distributed random variables with values in a measurable space $(S, \mathcal{S})$, where $\mathcal{S}$ is a σ-algebra. Let $h(s,t)$ be a function on $S \times S$ with values in a separable Hilbert space H; $h(s,t)$ is symmetric, i.e., $h(s,t) = h(t,s)$, $s, t \in S$, and measurable in the sense that $h^{-1}(E) \in \mathcal{S} \times \mathcal{S}$ for any Borel set $E \subset H$. We consider the U-statistic of degree 2, viz.,

$$(1) \qquad U_n = \binom{n}{2}^{-1} \sum_{1 \leqq i < j < n} h(X_i, X_j), \qquad n = 1, 2, \cdots$$

(the U-statistics of degree higher that 2 can be treated similarly).

Without loss of generality, let $\mathbf{E}h(X_1, X_2) = 0$, and assume that $\sigma^2 = \mathbf{E}|h(X_1, X_2)|^2 < \infty$. Let $g(x) = \mathbf{E}(h(X_1, X_2) \mid X_1 = x)$, denote by V the covariance operator corresponding to $g(X_1)$, i.e., the operator defined by $(Vx, y) = \mathbf{E}(g(X_1), x)(g(X_1), y)$, and let $\sigma_1^2 \geqq \sigma_2^2 \geqq \cdots$ be the eigenvalues of V. Furthermore, let Y be a Gaussian random variable with zero mean and covariance operator V, and let $B_r(a) = \{x \in H\colon\ |x - a| < r\}$.

Then, we have the following theorem.

THEOREM. *For all* $n \geqq 1$, $a \in H$, $r \geqq 0$,

$$(2) \qquad \begin{aligned} \Delta_n(a,r) &= \left|\mathbf{P}\big((n^{1/2}/2)U_n \in B_r(a)\big) - \mathbf{P}\big(Y \in B_r(a)\big)\right| \\ &\leqq c\Big(\Big(\big(\sigma_1^{-1}\sigma_2^{-1}(\sigma + |a|^2\big)\big(\mathbf{E}\big|h(X_1,X_2)\big|^2 - 2\sigma^2\big)n^{-1}\Big)^{1/3} \\ &\quad + \sigma^3\Big(\prod_{i=1}^{6}\sigma_i^{-1}\mathbf{E}\big|g(X_1)\big|^3\big(1 + |a|^3\big)n^{-1/2}\Big), \end{aligned}$$

where c is an absolute constant.

If, in addition, $\mathbf{E}|h(X_1, X_2)|^{2p} < \infty$ *for a positive integer* $p \geqq 2$, *then*

$$(3) \qquad \Delta_n(a,r) = O\big(n^{-p/(2p+1)}\big).$$

Proof. Without loss of generality, we may assume that $\sigma_6 \neq 0$ and $\mathbf{E}|g(X_1)^3| < \infty$. Put

$$(4) \qquad \widehat{U}_n = \binom{n}{2}^{-1} \sum_{1 \leqq i < j \leqq n} \big(g(X_i) + g(X_j)\big) = \frac{2}{n}\sum_{i=1}^{n} g(X_i),$$

*Received by the editors December 18, 1990.

†This research was supported by the Office of Naval Research under contract N00014-91-J-1020.

the projection of U_n. Letting $V_n = n^{1/2}U_n/2$ and $\widehat{V}_n = n^{1/2}\widehat{U}_n/2$, we can write

$$\Delta_n(a,r) \leqq \left|\mathbf{P}\big(V_n \in B_r(a)\big) - \mathbf{P}\big(\widehat{V}_n \in B_r(a)\big)\right| + \left|\mathbf{P}\big(\widehat{V}_n \in B_r(a)\big) - \mathbf{P}\big(Y \in B_r(a)\big)\right| = \Delta_{n1}(a,r) + \Delta_{n2}(a,r).$$

Observe now that, for any $\varepsilon > 0$, $q > 0$,

$$\begin{aligned}\Delta_{n1}(a,r) &= \left|\mathbf{P}\big(V_n \in B_r(a),\ \widehat{V}_n \notin B_r(a)\big) - \mathbf{P}\big(\widehat{V}_n \in B_r(a),\ V_n \notin B_r(a)\big)\right| \\ &\leqq \mathbf{P}\big(\widehat{V}_n \in B_{r+\varepsilon}(a)\backslash B_r(a)\big) + \mathbf{P}\big(\widehat{V}_n \in B_r(a)\backslash B_{r-\varepsilon}(a)\big) + 2\mathbf{P}\big(|V_n - \widehat{V}_n| \geqq \varepsilon\big) \\ &\leqq \Delta_{n2}(a, r-\varepsilon) + \Delta_{n2}(a, r+\varepsilon) + \mathbf{P}\big(Y \in B_{r+\varepsilon}(a)\backslash B_{r-\varepsilon}(a)\big) + 2\varepsilon^{-q}\mathbf{E}|V_n - \widehat{V}_n|^q,\end{aligned}$$

where $B_{r-\varepsilon}(a) = \varnothing$ if $\varepsilon > r$. Thus

$$\Delta_n(a,r) \leqq \sum_{i=1}^{3} \Delta_{n2}(a, r_i) + \mathbf{P}\big(Y \in B_{r+\varepsilon}(a)\backslash B_{r-\varepsilon}(a)\big) + 2\varepsilon^{-q}\mathbf{E}|V_n - \widehat{V}_n|^q, \tag{5}$$

where $r_1 = r - \varepsilon$, $r_2 = r$, $r_3 = r + \varepsilon$. By [2], we have

$$\Delta_{n2}(a, r_i) \leqq c\sigma^3\left(\prod_{i=1}^{6}\sigma_i^{-1}\right)\mathbf{E}\left|g(X_1)\right|^3\big(1 + |a|^3\big)n^{-1/2}, \qquad i = 1,2,3, \tag{6}$$

and, by [3, p. 84],

$$\mathbf{P}\big(Y \in B_{r+\varepsilon}(a)\backslash B_{r-\varepsilon}(a)\big) \leqq c\sigma_1^{-1}\sigma_2^{-1}\big(\sigma + |a|\big)\varepsilon. \tag{7}$$

Furthermore, it is easy to check that

$$\mathbf{E}\big(h(X_i, X_j), g(X_i)\big)^2 = \sigma^2 \tag{8}$$

and, for any, i, j, k, $i \neq j \neq k$,

$$\mathbf{E}\big(h(X_i, X_j), h(X_i, X_k)\big) = \mathbf{E}\big[\mathbf{E}\big(h(X_i, X_j), h(X_i, X_k)\big) \mid X_i\big] = \sigma^2. \tag{9}$$

A simple straightforward calculation using (8) and (9) gives

$$\mathbf{E}|V_n - \widehat{V}_n|^2 = 2^{-1}(n-1)^{-1}\big(\mathbf{E}\left|h(X_1, X_2)\right|^2 - 2\sigma^2\big). \tag{10}$$

Combining (5)–(7), (10) and putting

$$\varepsilon = \big(n^{-1}\big(\mathbf{E}\left|h(X_1, X_2)\right|^2 - 2\sigma^2\big)\sigma_1\sigma_2(\sigma + |a|)^{-1}\big)^{1/3},$$

we obtain (2).

For an integer $p \geqq 2$, the same reasoning as in Theorem 2.1 in [1] applied to

$$\mathbf{E}|V_n - \widehat{V}_n|^{2p} = n^p 2^{-2p}\binom{n}{2}^{-2p}\left(\sum_{1 \leqq i < j \leqq n} H(X_i, X_j),\ \sum_{1 \leqq i < j \leqq n} H(X_i, X_j)\right)^p,$$

where $H(X_i, X_j) = h(X_i, X_j) - g(X_i) - g(X_j)$, gives

$$\mathbf{E}|V_n - \widehat{V}_n|^{2p} = O(n^{-p}). \tag{11}$$

Combining (5)–(7), (11) and putting $\varepsilon = n^{-p/(2p+1)}$, we obtain (3).

An obvious consequence of the above theorem is that if $|g(X_1)|$ has moments of all orders, then

$$\Delta_n(a,r) = O\big(n^{-1/2+\delta}\big)$$

for any $\delta > 0$.

REFERENCES

[1] W. F. GRAMS AND R. J. SERFLING, *Convergence rates for U-statistics*, Ann. Statist., 1 (1973), pp. 153–160.

[2] V. V. SAZONOV, V. V. UL'YANOV, AND B. A. ZALESSKII, *A regular estimate of the convergence rate in the central limit theorem in Hilbert space*, Math. USSR Sbornik, 180 (1989), pp. 1587–1613. (In Russian.)

[3] V. I. PAULAUSKAS AND A. YU. RACHKAUSKAS, *Accuracy of Normal Approximation in the Central Limit Theorem in Banach Spaces*, Mokslas, Vilnius, 1987. (In Russian.)

INTEGRALS WITH RESPECT TO RANDOM MEASURES AND RANDOM LINEAR FUNCTIONALS*

V. N. RADCHENKO

(*Translated by V. Paulauskas*)

Let X be a compact Hausdorff topological space, $\mathfrak{B}$ a Borel σ-algebra of X, and $\mathfrak{B}_0$ some algebra of subsets of X generating $\mathfrak{B}$. Denote by $B(X)$ the space of real measurable and bounded functions and by $C(X)$ the space of real continuous functions on X. For f in $B(X)$ or $C(X)$ let

$$\|f\| = \sup_{x\in X} \big|f(x)\big|.$$

DEFINITION 1. By a random measure μ defined on $\mathfrak{B}$ (or on $\mathfrak{B}_0$) we mean a collection of random variables $\{\mu(A),\ A \in \mathfrak{B}\}$ ($A \in \mathfrak{B}_0$, respectively) such that

1) $\mu(A \cup B) = \mu(A) + \mu(B)$ a.s. for any A, B in $\mathfrak{B}$ (or $\mathfrak{B}_0$), $A \cap B = \varnothing$;

2) $\mu(A_n) \xrightarrow{\mathbf{P}} 0$ for A_n, $n \geqq 1$, in $\mathfrak{B}$ (or $\mathfrak{B}_0$), $A_n \downarrow \varnothing$.

We shall prove the following fact about the extension of a random measure.

THEOREM 1. *A random measure μ defined on $\mathfrak{B}_0$ can be extended to a random measure defined on $\mathfrak{B}$ if and only if the following relation holds for any sequence of pairwise disjoint sets $A_n \in \mathfrak{B}_0$:*

$$\lim_{c\to\infty} \sup_n \mathbf{P}\big\{\big|\mu(A_n)\big| \geqq c\big\} = 0. \tag{1}$$

Proof. The necessity of (1) is obvious. Let us proof the sufficiency.

It is well known that the space of random variables with the topology of convergence in probability is complete and metrizable. Therefore according to [1, Thm. 9.2] it is sufficient to show that $\mu(A_n) \xrightarrow{\mathbf{P}} 0$ for any disjoint $A_n \in \mathfrak{B}_0$. Consider such A_n.

From [2, Chap. 5, Lemma 4.3 c)] it follows that, for any c_k, $|c_k| \leqq 1$, and $l \in \mathbf{N}$, $c \in \mathbf{R}$,

$$\mathbf{P}\left\{\left|\sum_{k=1}^{l} c_k\mu(A_k)\right| \geqq c\right\} \leqq 8 \max_{\lambda_k=\pm1} \mathbf{P}\left\{\left|\sum_{k=1}^{l} \lambda_k\mu(A_k)\right| \geqq c/8\right\}.$$

Taking the sets $\bigcup_{k:\ \lambda_k=1} A_k$ and $\bigcup_{k:\ \lambda_k=-1} A_k$ we can ascertain that

$$\mathbf{P}\left\{\left|\sum_{k=1}^{l} c_k\mu(A_k)\right| \geqq c\right\} \leqq 8 \sup_{A,B\in\mathfrak{B}_0} \mathbf{P}\big\{\big|\mu(A) - \mu(B)\big| \geqq c/8\big\}.$$

---

*Received by the editors January 25, 1988.