U-statistics: the reverse martingale property

Wellner; 4/27/2010

Here is a replacement for the part (a) of PfS, Exercise 13.4.1, page 363:

Exercise 13.4.1.(a) (Reverse martingale property for $U-$statistics). Suppose that $X_1, X_2, \ldots$ are i.i.d. random variables with values in the measurable space $(S, \mathcal{S})$, and defined on $(\Omega, \mathcal{A}, P)$. Let $\mathcal{A}_n$ be the $\sigma-$field generated by all functions of $(X_1, X_2, \ldots)$ that are symmetric in their first n arguments. Let $h: \ S^r \to \mathbb{R}$ be a fixed kernel h of order r. Without loss of generality, we may assume that h is symmetric; if not replace h by the symmetric function

$$h_s(x_1, \ldots, x_r) \equiv \frac{1}{r!} \sum_{\pi \in \Pi_r} h(x_{\pi(1)}, \ldots, x_{\pi(r)}).$$

Prove that the sequence $\{U_n\}$ of $U-$statistics based on the kernel h, i.e.

$$U_n = \frac{1}{\binom{n}{r}} \sum_{1 \le i_1 < i_2 < \cdots < i_r \le n} h(X_{i_1}, X_{i_2}, \ldots, X_{i_r}),$$

is a reverse martingale for $n \ge r$ with respect to the filtration $\mathcal{A}_r \supset \mathcal{A}_{r+1} \supset \cdots$.

Now for a replacement of parts (b)-(c) of PfS, Exercise 13.4.1, page 363.

Exercise 13.4.1.(b)-(c) (SLLN for $U-$statistics). Suppose that the kernel h in problem 1 satisfies $E|h(X_1, \ldots, X_r)| < \infty$. Show that the corresponding sequence of U statistics $\{U_n\}_{n \ge r}$ satisfies $U_n \to_{a.s.,1} Eh(X_1, \ldots, X_r)$. (For $r > 1$ the given condition is not necessary: necessary and sufficient condition were found by Latała and Zinn (2000).) Use your result for $\{U_n\}$ to show a corresponding result for the sequence of $V-$statistics $\{V_n\}$ associated with $\{U_n\}$ under reasonable additional hypotheses.

Some history: The first mention of the reverse martingale property of $U-$statistics seems to be by Berk (1966), who actually needed a $U-$process result along the lines of those currently available in the $U-$process literature; see e.g. de la Peña and Giné (1999), pages 226 - 231. Sen (1974) gives results concerning convergence of $U-$ statistics in L_p, $p > 1$, while Lehmann (1951), Dwass (1956), Sen (1977), and Christofides and Serfling (1990) give various results for "generalized" or "k-sample" $U-$statistics. See van der Vaart (1998), chapter 12, pages 161 - 172, for a nice summary of distributional limit theorems for $U-$statistics, and see de la Peña and Giné (1999) for a treatment of $U-$processes.

Now consider the same basic situation as in Exercises 13.4.1.(a)-(c) above, but form the underlying probability space $(\Omega, \mathcal{A}, \mathbb{P})$ in the "canonical way": i.e. $(\Omega, \mathcal{A}, \mathbb{P}) = (S^\infty, \mathcal{S}^\infty, P^\infty)$ where $S^\infty = S \times S \times \cdots$, $\mathcal{S}^\infty$ is the product $\sigma-$field, and $\mathbb{P} = P^\infty$ is the product probability measure, so that $X_1, X_2, \ldots$ are the coordinate projections on $(\Omega, \mathcal{A}, \mathbb{P})$.

Now let $\mathbb{P}_n$ be the *empirical measure* of the X_i's:

$$\mathbb{P}_n \equiv n^{-1} \sum_{i=1}^n \delta_{X_i}.$$

Let $\mathcal{F}$ be a class of real-valued measurable functions $f : S \to \mathbb{R}$, $f \in L_1(S, \mathcal{S}, P)$, and write $Pf \equiv \int f dP$. For any fixed $f \in \mathcal{F}$, with $Y_i \equiv f(X_i) - Pf$ for $i = 1, \ldots, n$, we have

$$\mathbb{P}_n f - Pf = n^{-1} \sum_{i=1}^n Y_i$$

where the Y_i's are i.i.d. with mean zero, so $\{\mathbb{P}_n f - Pf, \mathcal{A}_n\}_{n \geq 1}$ forms a reverse-martingale. Thus $\{\mathbb{P}_n\}$ is a *measure - valued* reverse martingale; see e.g. Kallenberg (2005), page 74, Theorem 2.4. Since the supremum norm

$$\|\mathbb{P}_n - P\|_{\mathcal{F}}^* \equiv \left(\sup_{f \in \mathcal{F}} |\mathbb{P}_n f - Pf| \right)^*$$

is a convex function (see PfS, Exercise 13.1.1, page 350), we suspect that $\|\mathbb{P}_n - P\|_{\mathcal{F}}$ should be a reverse - sub-martingale with respect to $\{\mathcal{A}_n\}$. This is almost true. To make it true, we need to complete the filtration $\{\mathcal{A}_n\}$: with $A \triangle N$ denoting the symmetric difference $(A \setminus N) \cup (N \setminus A)$, let

$$\overline{\mathcal{A}}_n \equiv \{A \triangle N : \ A \in \mathcal{A}_n, N \in \mathcal{A}, \ \mathbb{P}(N) = 0\}.$$

The following theorem was proved by Strobl (1995); see also van der Vaart and Wellner (1996), Lemma 2.4.5, page 124, and problems 5 - 7, page 126.

Theorem. (Strobl) Suppose that $X_1, X_2, \ldots$ are the $S-$ valued coordinate maps on $(\Omega, \mathcal{A}, \mathbb{P})$ which is canonically-formed. Then every sequence of real-valued measurable cover functions of the empirical discrepancies $\|\mathbb{P}_n - P\|_{\mathcal{F}}^*$ forms a reversed sub-martingale with respect to $\{\overline{\mathcal{A}}_n\}$: that is, each $\|\mathbb{P}_n - P\|_{\mathcal{F}}^*$ is $\overline{\mathcal{A}}_n$-measurable and $\mathbb{P}-$integrable for all n, and for each m, n with $m \leq n$

$$E\left\{\|\mathbb{P}_n - P\|_{\mathcal{F}}^* | \overline{\mathcal{A}}_{n+1}\right\} \geq \|\mathbb{P}_n - P\|_{\mathcal{F}}^*.$$

References

BERK, R. H. (1966). Limiting behavior of posterior distributions when the model is incorrect. *Ann. Math. Statist. 37 (1966), 51–58; correction, ibid* **37** 745–746.

CHRISTOFIDES, T. C. and SERFLING, R. J. (1990). Maximal inequalities and convergence results for generalized U-statistics. *J. Statist. Plann. Inference* **24** 271–286.

DE LA PEÑA, V. H. and GINÉ, E. (1999). *Decoupling*. Probability and its Applications (New York), Springer-Verlag, New York. From dependence to independence, Randomly stopped processes. U-statistics and processes. Martingales and beyond.

DWASS, M. (1956). The large-sample power of rank order tests in the two-sample problem. *Ann. Math. Statist.* **27** 352–374.

KALLENBERG, O. (2005). *Probabilistic symmetries and invariance principles*. Probability and its Applications (New York), Springer, New York.

LATAŁA, R. and ZINN, J. (2000). Necessary and sufficient conditions for the strong law of large numbers for U-statistics. *Ann. Probab.* **28** 1908–1924.

LEHMANN, E. L. (1951). Consistency and unbiasedness of certain nonparametric tests. *Ann. Math. Statistics* **22** 165–179.

SEN, P. K. (1974). On L^p-convergence of U-statistics. *Ann. Inst. Statist. Math.* **26** 55–60.

SEN, P. K. (1977). Almost sure convergence of generalized U-statistics. *Ann. Probability* **5** 287–290.

STROBL, F. (1995). On the reversed sub-martingale property of empirical discrepancies in arbitrary sample spaces. *J. Theoret. Probab.* **8** 825–831.

VAN DER VAART, A. W. (1998). *Asymptotic statistics*, vol. 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge.

VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak convergence and empirical processes*. Springer Series in Statistics, Springer-Verlag, New York. With applications to statistics.