$S_t \circ c(h)(a_s)_{ti} c(s) + \mathfrak{S}_e m^i n(a_r)$

Department of Mathematics, University of Utah

Non-central limit theorems for random selections

Johannes Schauer

Graz University of Technology

Time and Place: Thursday September 11, 2008; 9:30–10:30 p.m.; JWB 335

Selection from a finite population is a basic procedure of statistics and large sample properties of many classical tests and estimators are closely related to the asymptotic behavior of sampling variables. Consider an array $\{x_{i,n},\ 1 \leq i \leq n,\ n \geq 1\}$ of real numbers and let $X_1^{(n)}, \ldots, X_m^{(n)}$ be the random variables obtained by drawing, with or without replacement, $m = m(n)$ elements from the set $\{x_{1,n}, \ldots, x_{n,n}\}$. Under the uniform asymptotic negligibility condition $\max_{1 \leq i \leq n} |x_{i,n}| \to 0$, the process $Z_{n,m}(t) = \sum_{k \leq mt} X_k^{(n)}$ $(0 \leq t \leq 1)$, suitably centered and normalized, is known to converge weakly to Brownian motion, respectively Brownian bridge, depending on whether we draw with or without replacement. Allowing the $x_{i,n}$ themselves to be random leads to classical results for bootstrap and permutation statistics. The purpose of our paper is to study the behavior of $Z_{n,m}(t)$ in the case when the condition $\max_{1 \leq i \leq n} |x_{i,n}| \to 0$ does not hold. We will prove that under mild regularity conditions $Z_{n,m}(t)$ still converges weakly, but its limit will depend on the "non negligible" elements of the array $\{x_{i,n}\}$. We give a series representation of the limiting processes as sums of independent jump processes. In the case of random $\{x_{i,n}\}$, our results lead to new and unusual limit theorems for bootstrap and permutation CUSUM statistics.

We discuss several statistical applications of our results. The simplest situation violating the asymptotic negligibility condition is resampling from a (normalized) sample of i.i.d. random variables in the domain of attraction of an α-stable law, $0 < \alpha < 2$. In this case the bootstrap and permutation CUSUM statistics converge in distribution to a random and nondegenerate limit, extending results of Athreya on the bootstrap (Ann. Statist. 15, 1987, p. 724-731). We consider some other examples with heavy tailed distributions and some oscillating situations as well. For example, in the case of semistable variables the bootstrap and permutation statistics do not converge, but they exhibit a remarkable logarithmic periodicity: instead of obtaining (random) limits as $n \to \infty$, we find different limits for all subsequences of the form $n_k = \lfloor c\, 2^k \rfloor$ as $k \to \infty$ for $1 \le c < 2$.

This is joint work with István Berkes and Lajos Horváth.