Donsker results for the smoothed empirical process

Eric Beutner[1], **Henryk Zähle**[2,*]

[1] *Department of Quantitative Economics, Maastricht University, Maastricht, The Netherlands*
[2] *Department of Mathematics, Saarland University, Saarbrücken, Germany*
[*] Email: zaehle@math.uni-sb.de

The empirical probability measure $\widehat{\mu}_n$ of identically distributed real-valued random variables $X_1, \ldots, X_n$ with distribution μ is the random measure that uniformly allocates total mass one to the random atoms $X_1, \ldots, X_n$. The corresponding empirical process with index set $\mathcal{G}$ consisting of measurable functions is given by

$$\sqrt{n}\Big(\int g \, d\widehat{\mu}_n(\omega) - \int g \, d\mu \Big), \qquad g \in \mathcal{G}, \, \omega \in \Omega$$

and plays a central role in the field of nonparametric statistics. Under suitable conditions this process converges in distribution to a non-degenerate limit process as $n \to \infty$, and much is already known about it.

The smoothed empirical process is defined analogously where $\widehat{\mu}_n$ is replaced by a smoothed version based on a kernel density estimator. In this talk I present new results on convergence in distribution of the smoothed empirical process for large index sets $\mathcal{G}$ under weak assumptions. The results cover both a MISE optimal choice of the bandwidth and short-range dependence of $X_1, \ldots, X_n$ $(, X_{n+1}, \ldots)$. The results continue to hold under long-range dependence when $\sqrt{n}$ is replaced by a suitable "non-central" rate.