Peter Grünwald: the Catch-Up Phenomenon in Model Selection and Model Averaging

We resolve a long-standing debate in statistics, known as the AIC-BIC dilemma: model selection/averaging methods like BIC, the Bayes factor, and MDL are consistent (they eventually infer the correct model) but, when used for prediction, the rate at which predictions improve can be suboptimal. Methods like AIC and leave-one-out cross-validation are inconsistent but typically converge at the optimal rate. We give a novel analysis of the slow convergence of the Bayesian-type methods. Based on this analysis, we propose the switching method, a modification of Bayesian model averaging that achieves both consistency and minimax optimal convergence rates. The method is related to expert-tracking algorithms developed in the COLT literature, and has time complexity comparable to Bayes. Experiments with nonparametric density estimation confirm that our large-sample theoretical results also hold in practice in small samples. We also discuss how our results can coexist with seemingly incompatible results by Yang (2005), who proved that the strengths of AIC and BIC cannot always be shared.

Joint work with T. van Erven and S. de Rooij.

Peter Grünwald (Ph.D. 1998) heads the information-theoretic learning group at the CWI, the national research institute for mathematics and computer science in the Netherlands, located in Amsterdam. He is also part-time professor at Leiden University. His research interests lie where statistics, computer science and information theory meet: theories of learning from data. He is author of the book *The Minimum Description Length Principle* (MIT Press, June 2007). Recently, he has been involved in exposing the flawed statistics that were used in the trial against Lucia de B., nurse and alleged serial killer.