



THE UNIVERSITY OF  
**CHICAGO**

Department of Statistics

STATISTICS COLLOQUIUM

---

STEFAN WAGER

Statistics Department  
Stanford University

**Statistical Estimation with Random Forests**

TUESDAY, January 12, 2016, at 12:00 PM

Math-Stat Building (Stevanovich Center), Room 112, 5727 S. University Ave.

ABSTRACT

Random forests, introduced by Breiman (2001), are among the most widely used machine learning algorithms today, with applications in fields as varied as ecology, genetics, and remote sensing. Random forests have been found empirically to fit complex interactions in high dimensions, all while remaining strikingly resilient to overfitting. In principle, these qualities ought to also make random forests good statistical estimators. However, our current understanding of the statistics of random forest predictions is not good enough to make random forests usable as a part of a standard applied statistics pipeline: in particular, we lack robust consistency guarantees and asymptotic inferential tools. In this talk, I will present some recent results that seek to overcome these limitations. The first half of the talk develops a Gaussian theory for random forests in low dimensions that allows for valid asymptotic inference, and applies the resulting methodology to the problem of heterogeneous treatment effect estimation. The second half of the talk then considers high-dimensional properties of regression trees and forests in a setting motivated by the work of Berk et al. (2013) on valid post-selection inference; at a high level, we find that the amount by which a random forest can overfit to training data scales only logarithmically in the ambient dimension of the problem.

This talk is based on joint work with Susan Athey, Brad Efron, Trevor Hastie, and Guenther Walther

---

For further information and about building access for persons with disabilities, please contact Courtney Tillman at 773.702.8333 or send email ([cmtillman@galton.uchicago.edu](mailto:cmtillman@galton.uchicago.edu)). If you wish to subscribe to our email list, please visit the following website: <https://lists.uchicago.edu/web/arc/statseminars>.