

HARVARD UNIVERSITY
DEPARTMENT OF STATISTICS



Tel. 617-495-5496
Fax. 617-496-8057
statdept@stat.harvard.edu
<http://www.stat.harvard.edu/>

Science Center 7th Fl.
One Oxford Street
Cambridge, MA 02138

STATISTICS COLLOQUIUM

MONDAY, APRIL 27, 2015

TALK: 4:15 PM — SCIENCE CENTER RM. 705

RECEPTION: 3:50 PM — SCIENCE CENTER, 7TH FLOOR

**“A Consistent and Scalable Bayesian Model Selection Method for
Logistic Regression”**

**Xuming He
Department of Statistics
University of Michigan**

ABSTRACT

Bayesian model selection faces challenges both in theory and in computation when the number of potential covariates p is large. We propose a pseudo-Bayesian variable selection method for logistic regression that adapts to both the sample size n and the number of potential covariates p with two important features. First, it has strong model selection consistency even when p is large. Second, we propose a new Gibbs sampler that does not require p -squared operations in each of its iterations. In contrast with the standard Gibbs sampler which requires sampling from a p dimensional multivariate normal distribution with a non-sparse covariance matrix, our new algorithm is much more scalable to high dimensional problems, both in memory and in computational efficiency. We compare our proposed method with several leading variable selection methods through a simulation study to show that our proposed approach selects the correct model, when there is one, with higher probabilities than most competitors. The talk is based on ongoing work with Naveen Narisetty and Juan Shen..