



THE UNIVERSITY OF
CHICAGO

Department of Statistics
STATISTICS COLLOQUIUM

SAMUEL WONG

Department of Statistics
Harvard University

**Efficient Sampling Methods for Protein Structure
Refinement**

MONDAY, February 4, 2013 at 4:00 PM

133 Eckhart Hall, 5734 S. University Avenue

Refreshments following the seminar in Eckhart 110

ABSTRACT

In protein folding, scientists are interested in the prediction of the three-dimensional structure, based on the amino acid sequence. Initial structures of new proteins are often built by finding templates from databases of proteins with known structure; this procedure is called homology modeling in the bioinformatics literature. The goal of refinement is to generate a structure prediction that improves upon a given homology model, especially in regions where a good template is not available.

The conformational space of a protein is too large for an exhaustive search to be computationally feasible. An efficient sampling algorithm is therefore crucial to the success of any refinement procedure. In this talk, we first focus on segments that are not well modeled by matching templates; the prediction of these segments is known as loop modeling. We present a new Monte Carlo method to sample loops, based on sequentially guiding the sampling distributions of each amino acid to obey geometric and energetic constraints. We then describe some extensions to the general refinement problem. Our method is compared with existing ones on real protein datasets, and encouraging results are obtained thus far.

For further information and inquiries about building access for persons with disabilities, please contact Dan Moreau at 773.702.8333 or send him an email at dmoreau@galton.uchicago.edu. If you wish to subscribe to our email list, please visit the following website: <https://lists.uchicago.edu/web/arc/statseminars>.