



The University of Chicago  
Departments of Computer Science,  
Mathematics, and Statistics

## Scientific and Statistical Computing Seminar

---

**SAYAN MUKHERJEE**

Department of Statistics  
Duke University

### Probabilistic Perspective on Topological Data Analysis

**FRIDAY, October 21, 2011, at 3:30 PM**

133 Eckhart Hall, 5734 S. University Avenue (unless announced otherwise).

### ABSTRACT

In this talk we discuss the recent area of topological data analysis (TDA) from a probabilistic perspective. The talk falls under two parts. No knowledge of topology is assumed.

The first part of the talk considers a classic object in topology and geometry, a (Whitney) stratified space. This object is stated as a mixture model in the statistical sense and an algorithm for clustering observations belonging to the same mixture is stated and finite sample bounds are provided. A key idea in this procedure is relative persistence homology which was developed in computational topology.

In the second part of the talk we reconsider objects computed in computational topology as summary statistics of distributions. We focus on a particular popular object, the persistence diagram, and show that this summary (with minor restrictions) admits a valid probability space. We also show that Frechet means and variances for these summaries exist. This part of the talk will argue that these summary statistics can be used for probabilistic reasoning and classical statistical formalisms should apply. A practical implication of this work is that means and variances of persistence diagrams can be computed.

---

#### Organizers:

Lek-Heng Lim, Department of Statistics, [lekheng@galton.uchicago.edu](mailto:lekheng@galton.uchicago.edu),  
Ridgway Scott, Departments of Computer Science and Mathematics, [ridg@cs.uchicago.edu](mailto:ridg@cs.uchicago.edu),  
Jonathan Weare, Department of Mathematics. [weare@math.uchicago.edu](mailto:weare@math.uchicago.edu).  
SSC Seminar URL: <http://sites.google.com/site/uchicagoss/>

If you wish to subscribe to our email list, please visit the following website:  
<https://lists.uchicago.edu/web/arc/statseminars>.