

Topics in Biomedical Data Science: Large-scale inference

[Jump to Today](#) [Edit](#)

BIODS215 Topics in Biomedical Data Science: Large-scale inference

Winter 2018

TTh 3:00PM - 4:20PM

- [Alway Building, Room M112](http://campus-map.stanford.edu/?srch=Alway+Building) (<http://campus-map.stanford.edu/?srch=Alway+Building>)
- 1/25 and 3/1: Alway Building, Room M114
- 3/6: Medical School Office Building (MSOB), Room x303

Course Instructors: Prof. Manuel A. Rivas (mrivas@stanford.edu) (<mailto:mrivas@stanford.edu>), Prof. James Zou (jamesz@stanford.edu) (<mailto:jamesz@stanford.edu>), Prof. Julia Salzman (horence@stanford.edu) (<mailto:horence@stanford.edu>)

TA: Yosuke Tanigawa (ytanigaw@stanford.edu) (<mailto:ytanigaw@stanford.edu>)

Office Hours: Yosuke Tanigawa (Tue. 1:30pm-3pm, MSOB x393); Instructors: After class or by appointment.

Course description

The recent explosion of data generated in the fields of biology and medicine has led to many analytical challenges and opportunities for understanding human health. This graduate-level course focuses on methodology for large-scale inference from biomedical data.

Pre-requisite

Students must be comfortable with basic probability and statistical principles (e.g. at the level of STATS200) and computer programming (e.g., R or Python or MATLAB).

Course requirements & grading

Two homework assignments (40%), final project (50%), class participation (10%). Most of the homework problems will require programming. You may use Python, R or MATLAB.

We encourage you to work collaboratively on homework assignments, but you **must write all your own code and written answers yourself**. Duplicating answers and/or code based on other students' work is a violation of Stanford's Honor Code. For the final project we encourage you to work in teams. In your final

report you will be asked to state precisely each person's contribution.


Late Homework Policy














Homework should be submitted electronically on the day it is due. Each student has a total of six “free” late days (a late day is 24 hours of lateness). There are no partial days, so assignments are either on time, 1 day late, 2 days late, etc. Beyond the six days, further extensions will only be granted for highly extenuating circumstances.

Recommended textbooks and resources. (Optional)

Computer Age Statistical Inference: Algorithms, Evidence, and Data Science. Bradley Efron and Trevor Hastie.

Course Summary:

Date	Details
Tue Jan 9, 2018	 Lec.1: Course introduction and Statistical Genetics (M.R.) (https://canvas.stanford.edu/calendar?event_id=62781&include_contexts=course_78313) 3pm to 4:20pm

Date	Details	
Thu Jan 11, 2018	 Lec.2: Disease Genomics (J.S.) (https://canvas.stanford.edu/calendar?event_id=62782&include_contexts=course_78313)	3pm to 4:20pm
Tue Jan 16, 2018	 Lec.3: An introduction to deep learning (J.Z.) (https://canvas.stanford.edu/calendar?event_id=63654&include_contexts=course_78313)	3pm to 4:20pm
Thu Jan 18, 2018	 Lec.4: high dimensional RNA abundance estimation (Poisson models) (J.S.) (https://canvas.stanford.edu/calendar?event_id=63655&include_contexts=course_78313)	3pm to 4:20pm
Tue Jan 23, 2018	 Lec.5: Robust methods; rank testing permutation distributions (J.S.) (https://canvas.stanford.edu/calendar?event_id=63656&include_contexts=course_78313)	3pm to 4:20pm
Thu Jan 25, 2018	 Lec.6: Monte carlo techniques-- intuition for mixing times (markov chains); martingales and permutation testing (J.S.) (https://canvas.stanford.edu/calendar?event_id=63657&include_contexts=course_78313)	3pm to 4:20pm
Tue Jan 30, 2018	 Lec.7: functional approximation: Neyman's smooth testing, splines, fourier analysis; PCA and neural nets (J.S.) (https://canvas.stanford.edu/calendar?event_id=63658&include_contexts=course_78313)	3pm to 4:20pm
Thu Feb 1, 2018	 Lec.8: Multi level modeling (M.R.) (https://canvas.stanford.edu/calendar?event_id=63659&include_contexts=course_78313)	3pm to 4:20pm
Tue Feb 6, 2018	 Lec.9: Mixture models (M.R.) (https://canvas.stanford.edu/calendar?event_id=63660&include_contexts=course_78313)	3pm to 4:20pm
Thu Feb 8, 2018	 Lec.10: High-dimensional inference methods on summary statistics (M.R.) (https://canvas.stanford.edu/calendar?event_id=63661&include_contexts=course_78313)	3pm to 4:20pm
Tue Feb 13, 2018	 Lec.11: Prediction with Bayesian models (M.R.) (https://canvas.stanford.edu/calendar?event_id=63662&include_contexts=course_78313)	3pm to 4:20pm
Thu Feb 15, 2018	 Lec.12: Distributed and Privacy-Preserving Computations (Naras) (https://canvas.stanford.edu/calendar?event_id=63663&include_contexts=course_78313)	3pm to 4:20pm
Tue Feb 20, 2018	 Lec.13: TBA (J.Z.) (https://canvas.stanford.edu/calendar?event_id=63670&include_contexts=course_78313)	3pm to 4:20pm
Thu Feb 22, 2018	 Lec.14: TBA (J.Z.) (https://canvas.stanford.edu/calendar?event_id=63671&include_contexts=course_78313)	3pm to 4:20pm

Date	Details	
Tue Feb 27, 2018	 Lec.15: Guest lecture (TBA) (https://canvas.stanford.edu/calendar?event_id=63672&include_contexts=course_78313)	3pm to 4:20pm
Thu Mar 1, 2018	 Lec.16: Guest lecture (TBA) (https://canvas.stanford.edu/calendar?event_id=63673&include_contexts=course_78313)	3pm to 4:20pm
Tue Mar 6, 2018	 Lec.17: Data Visualization (Y.T.) (https://canvas.stanford.edu/calendar?event_id=63674&include_contexts=course_78313)	3pm to 4:20pm
Thu Mar 8, 2018	 Lec.18: Frontier of Biomedical Data Science Research (https://canvas.stanford.edu/calendar?event_id=63675&include_contexts=course_78313)	3pm to 4:20pm
Tue Mar 13, 2018	 Lec.19: Project presentation (https://canvas.stanford.edu/calendar?event_id=63676&include_contexts=course_78313)	3pm to 4:20pm
Thu Mar 15, 2018	 Lec.20: Project presentation (https://canvas.stanford.edu/calendar?event_id=63677&include_contexts=course_78313)	3pm to 4:20pm
	 Reading for lecture #1 (https://canvas.stanford.edu/courses/78313/assignments/112113)	