OPNS 525: Emerging Areas in Operations Management:

Statistical Learning in Sequential Decision-Making

Time: Spring Quarter, Tuesday/ Thursday 2:00-3:30

Location: TBD (likely the new Kellogg building)

Professor: Daniel Russo

Course Description: This course offers an advanced introduction to topics at the intersection of statistical (machine) learning and sequential decision-making. A tentative course plan is as follows. We will begin by covering classic work on optimal hypothesis testing when data can be gathered sequentially and interactively. The second part of the class focuses on bandit learning and the design and analysis of algorithms that balance exploration/exploitation. The last part of the course introduces reinforcement learning, including methods for value function approximation and algorithms for efficient exploration.

Students should have experience with mathematical proofs, coding for numerical computation, and the basics of statistics, optimization, dynamic programming, and stochastic processes.

Tentative outline:

1. Sequential and adaptive hypothesis testing
    - Wald's sequential probability ratio test and optimal stopping
    - Chernoff's optimal sequential design of experiments for hypothesis testing
2. Bandit Learning
    - Upper-confidence bound algorithms and the optimistic principle
    - Thompson sampling
    - Regret analysis
    - Applications to the online shortest path problem and dynamic pricing
3. Reinforcement learning (RL)
    - Value function learning: least-squares value iteration, temporal differences, and Q-learning
    - Parametric approximations to the value function
    - The exploration problem in RL