Stat/Biostat Project Classes: Expectations and Tips

Vladimir Minin

Department of Statistics
University of Washington, Seattle

March 29, 2012
## Course goals

### Understanding your paper
- This is a good time to spend hours on this one formula
- Read other papers and appropriate reference books
- What does it mean for you to understand something?

### Implementing your paper
- Think about the data and how they were generated
- Try to reproduce all data analysis from your paper
- Design a set of simulation experiments
- Can you break the paper’s method?
Course goals

Understanding your paper

- This is a good time to spend hours on this one formula
- Read other papers and appropriate reference books
- What does it mean for you to understand something?

Implementing your paper

- Think about the data and how they were generated
- Try to reproduce all data analysis from your paper
- Design a set of simulation experiments
- Can you break the paper’s method?
Your participation

- Keep explaining your paper to whoever will listen. We’ll organize group discussions to help you. Explaining what part of the paper you don’t understand often solves the problem.

- Try to understand what everyone else is doing. There are many connections between your projects to be discovered.

- Bug the instructor during and outside of the classroom.

- Talk to other students during and outside of the classroom.

- Be critical of any advice by your peers and the instructors.

- Don’t trust derivations in your paper.
Tips: writing

- Don’t put off writing, start recording your thoughts from day 1. Think by writing

- You can write code, math, and documentation at the same time using Sweave

- Create an outline of your report, make a plan of what you want to accomplish

- Start a day with two hours of writing. It does not matter how much you actually write as long as you keep writing every day
Tips: coding

- Don’t put off coding, start numerical experimentations from day 1
- Start small. Reduce your problem to a ridiculously simple example and beat this example to death
- If math is hard, simulate
- Premature optimization is the root of all evil
More on this later

We’ll be posting links on the discussion board(s)
Side effects of taking Stat/Biostat Project classes may include:

1. Increased scientific curiosity
2. Involuntary wikipedia browsing
3. Clusterphobia
4. Markov dependency
5. Professorial self-confidence
6. Coding blues
7. Flashes of academic cynicism