

BINF 3121 Statistics for Bioinformatics

Introduction to the class

Learning Goals

- Understand core ideas of statistical inference
 - basic statistical concepts (like you'd learn in any statistics class)
 - ***plus*** how we use these ideas in genomics
- Learn to do data analysis
 - using R, statistical programming language
 - produce convincing & reproducible data analysis reports

How the class works

- Do reading, tutorials before class on Monday (prep work)
 - Quizzes on the reading are due before class
- In-class activities reinforce the reading, tutorials
 - Try to finish in class; in-class work due by Friday
 - If you need help, go to TA office hours
- Learn by doing
 - Solving problems
 - Explaining, teaching

Don't skip class - you will get very lost

- Bring a laptop to *every* class session
 - Can check out laptops from library
 - Also can share with a class-mate
 - Will use RStudio for in-class exercises
- You must attend every class
 - Much material presented only in class
 - If you skip class, there is no way to make it up
 - Class engagement score
 - An assignment in Canvas; Can earn up to 10 points per class



Tips on how to get a top score in classroom engagement

- Voluntarily join class discussions
 - Offer questions or comments during class discussions
- Shows respect; listens to others' contributions
 - Don't interrupt
 - Refer to what others say and build on it: "yes, and"
- Actively engage in class exercises
- Use TA hours to clarify ideas
 - Attend TA office hours
 - Make appointments if office hours conflict with your schedule
- Prepare for class by completing all pre-work
- Arrive on time and stay for entire class
- Submit assignments on time

Grading

- Each assignment earns points, final grade depends on % points earned
 - See syllabus for details
 - You can calculate your grade to determine how you're doing
- Pre-work assignments that are due before the next class starts **must** be completed by then
 - No late assignments accepted

Assignments

- Stanford OLI Probst module reading + Canvas quizzes
 - Due by **start of class on Mondays**
- R tutorials and quizzes
 - Also due by the **start of class on Mondays**
- In-class activities where you learn how to use R functions to solve problems in probability & statistics
 - Typically due on **Friday** after class
 - Must bring laptop to class to practice, get help as needed

One midterm

- March 26
- Covers the Stanford OLI Probstat assigned modules
 - Questions taken from Probstat examples, the quizzes, some new
- Open book, open computer exam
 - You will use R to answer questions.
 - This will test your knowledge and your ability to use computational tools (R & RStudio) to answer questions.

Assignments, con't

- After you have a good foundation in R, probability & statistics, we'll move on to the data analysis part of the class
- Learn to use R Markdown to write data analyses reports
 - Your coding skills will get a workout
 - But if you work hard and master the material, you'll have a very advanced knowledge of data analysis (great preparation for graduate school, medical school, or a job in bioinformatics)

Canvas

- Most assignments are already available in Canvas
- Use the Calendar view, Assignments view, etc to keep track of what's due and when
- You will do a lot!
 - Keep up.
 - Don't wait until the day before something is due to get started.

Before next meeting

- Install R, RStudio on a laptop
 - Make sure that it works; go to TA office hours for help if you need it
 - We may not be able to help you trouble-shoot during next class
- Bring laptop to class for in-class tutorial and introduction to R
- Complete Stanford on-line course - Part I Exploratory Data Analysis
 - Take first quiz before coming to class