Introduction to probability and modeling

Ruy Ribeiro

Workshop plan

(First 3 sessions)
0) Introduction (lecturer and students)
1) What is modeling? Types of modeling.
2) Paper discussion (identify modeling ideas and statistical issues)
3) R practice
    3a) Descriptive statistics (brief overview)
    3b) Graphing data
    3c) Mean as a statistical linear model
4) Probabilities vs. statistics (knowing the model vs. knowing the data)

(Next 3 sessions)
5) Probabilities (brief introduction)
    5a) Probability distributions (discrete vs. continuous): binomial, Poisson, negative binomial, normal, t, chi-squared (group presentation)
    5b) Working with and visualizing probabilities in R (practice)
    5c) Central limit theorem as a probability distribution
6) Discussion from practice in 5) about sampling distributions and confidence intervals (for "means")
7) Hypothesis testing (null hypothesis, alternative hypothesis, test, errors)

(Final 4 sessions))
8) Linear regression (as a statistical model): coefficients, assumptions, prediction
9) General linear model: categorical and continuous covariates, interactions, checking assumptions, confidence interval and hypothesis testing for coefficients, quality of model fit, model building
10) Practice

Software: RStudio and R

Papers:

To be discussed by the group in class:

https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0026885

May be referenced to in class as examples of modeling:

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3466592/

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2932550/

https://www.medrxiv.org/content/10.1101/2021.09.03.21263105v1.full-text