Data Science in Bioinformatics 2023

Week	Week	Person	Lecture topic	Lecture goals + details	Exercises
1	35	Palle	Introduction	**Lecture goals** • Course introduction • A small quiz of your present knowledge/skills • Padlet intro + super small discussion **Video: A Gentle Introduction to Tidy Statistics in R** https://www.rstudio.com/resources/webinars/a-gentle-introduction-to-tidy-statistics-in-r/	**R for Data Science** Ch.1. Welcome - Introduction Ch.2. Explore - Introduction Ch.3. Explore - Data visualization Exercise goals • R + rstudio intro • Rmarkdown • Simple figure exercise
2	36	Palle	Refresher and discussion of the basic parts of statistics	**Analysis of Biological Data** Ch.1. Statistics and samples Ch.2. Displaying data Ch.3. Describing data Ch.4. Estimating with uncertainty	**R for Data Science** Ch.4. Explore – Workflow basics Ch.5. Explore – Data transformation Download, unzip etc. the birth data + new statsDK dataset
3	37	Palle	**Introduction to Assignment 1** **Deadline: Assignment 1 (end of week)** Probability	**Analysis of Biological Data** 5. Probability Introduction to assignment 1 A guided discussion about: What is probability? Conditional probability (wasps) Bayes (cannabis)	**R for Data Science** Ch.6. Explore - Workflow: scripts Ch.7. Explore - Exploratory Data Analysis (these are mostly for the assignment – not really for the probability exercise) Probability exercise
4	38	Thomas	(students peer review assignment) Hypothesis testing and proportions	**Analysis of Biological Data** 6. Hypothesis testing 7. Analyzing proportions Multiple testing	**R for Data Science** Ch. 4: Learn about the concept of "tidy" data as a standardized data input and output format for all packages in the tidyverse. Proportion of lowly expressed genes found on autosome vs. X (Mammals dataset) Similar to example 7.2 in the ABD book
5	39	Thomas	At class: Present 2-3 submissions Fitting models	**Analysis of Biological Data** 8. Fitting probability models to frequency data (Poisson + binomial)	Case study: A meta-analysis of p-values Practicing the Poisson distribution and GOF tests with the Poisson Case study: planning a sequencing experiment
6	40	Thomas	Analysis of count **Introduction to Assignment 2**	**Analysis of Biological Data** 9. Contingency analysis: associations between categorical variables	Book exercises + chisq.test() on mammals dataset
7	41	Thomas (Palle gone)	Comparing groups of numbers **Deadline: Assignment 2 "A data story" (end of week)**	**Analysis of Biological Data** 10. The normal distribution 11. Inference for a normal population 12. Comparing two means 13. Handling violations of assumptions	Analyse GC of chromosome X genes vs. autosomal genes for each species in the Mammals dataset. t.test() wilcox.test() Permutation test
	42	(Palle gone)	(students peer review assignment)	**Fall break - No lectures or exercises**	

8	43	Palle	At class: Present 2-3 submissions Comparing groups of numbers Anova	**Analysis of Biological Data** 14. Designing experiments 15. Comparing means of more than two groups	ANOVA walkthrough + simulating power.
9	44	Palle	Comparing two numeric variables **Introduction to Assignment 3**	**Analysis of Biological Data** 16. Correlation between numerical variables 17. Regression	Simple regression and correlation (Mammals dataset) There should be extra time for assignment 3
10	45	Palle	**Deadline: Assignment 3 "Our small analysis" (end of week)**	**Analysis of Biological Data** 18. Multiple explanatory variables	A final look at the Fast X effect: a linear model to explain variation in rates of evolution.
11	46	Thomas	(students peer review assignment) At class: Present 2-3 submissions	13. Permutation 19. Computer intensive methods Publish Exam topics	Resampling Drosophila dataset Bootstrap correlation from week 44 Permutation test regression from week 44 Permutation testing (t.test) – redo book example
12	47	Thomas	Likelihood Introduction to Assignment 4 **Deadline: Assignment 4 "Our big analysis" (Sunday)**	20. Likelihood	Likelihood concept by worked examples: Estimating and comparing frequencies Poisson distributed events (mutations, terrorist attacks, etc) Likelihood ratio tests
13	48	Thomas	Peer review assignment 4	Likelihood continued	Likelihood for normally distributed data, a taste for Bayesian inference
14	49	All		Xmas cookies and feedback on course Exam information	FEEDBACK on the course