

Bioinformatics Summer Institute - Probability/Statistics Workshop

Instructor: **Dr. Silvia Heubach**

References

Lectures:

For the first lecture, any introductory probability book will work. The one given below is just one of many choices.

- Saeed Ghahramani, *Fundamentals of Probability*, 2nd Edition, Prentice Hall, 2000

An introduction to probability including Markov Chains can be found in

- Morris H DeGroot, *Probability and Statistics*, 2nd Edition, Addison Wesley, 1986

Hidden Markov models references:

- David W. Mount, *Bioinformatics – Sequence and Genome Analysis*, Cold Spring Harbor Laboratory Press, 2001
- Peter Clote and Rolf Backofen, *Computational Molecular Biology – An Introduction*, John Wiley & Sons Ltd, 2000
- Sean R. Eddy, Hidden Markov Models, *Curr. Opin. Struct. Biol.* **6**: 361-365, 1996
- Michael S. Waterman, *Introduction to Computational Biology – Maps, sequences and genomes*, Chapman & Hall/CRC Press, 2000
- Arthur M. Lesk, *Introduction to Bioinformatics*, Oxford University Press, 2002
- L.R. Rabiner and B.H. Juang, An Introduction to Hidden Markov Models, *IEEE ASSP Magazine*, 4-16, January 1986
- Anders Krogh et. al., Hidden Markov Models in Computational Biology – Applications to Protein Modeling, *J. Mol. Biol.* (1994) **235**, 1501 – 1531
- Anna Tramontano, *Introduction to Bioinformatics*, Chapman and Haa/CRC 2007

Significance of Alignment

- David W. Mount, *Bioinformatics – Sequence and Genome Analysis*, Cold Spring Harbor Laboratory Press, 2001
- Arthur M. Lesk, *Introduction to Bioinformatics*, Oxford University Press, 2002

- <http://www.ncbi.nlm.nih.gov/BLAST/tutorial/Altschul-1.html>

Algorithms on the Web (Hidden Markov Models)

R. Hughey, K Karplus and D. Haussler (University of California Santa Cruz)

- <http://cse.ucsc.edu/research/compbio/sam.html>
- <http://cse.ucsc.edu/research/compbio/HMM-apps/HMM-applications.html>

Other References

The reference for using Bayes' formula for spam filtering is

- Gaeorge Sakkis, *Learning how to tell Ham from Spam*, Crossroads, Issue 11.2, Winter 2004

Actually, the whole issue talks about methods for filtering spam.

Assignments:

Some of the assignments were taken/adapted from the following books

- Michael S. Waterman, *Introduction to Computational Biology – Maps, sequences and genomes*, Chapman & Hall/CRC Press, 2000
- Arthur M. Lesk, *Introduction to Bioinformatics*, Oxford University Press, 2002
- Peter Clote and Rolf Backofen, *Computational Molecular Biology – An Introduction*, John Wiley & Sons Ltd, 2000
- Bernard Rosner, *Fundamentals of Biostatistics*, 6th Edition, Duxbury, 2006