Probability Pretest

Consider the following probability distribution:

\[
\begin{align*}
P(A=\text{true}) &= 2/3 \\
P(B=\text{true} \mid A=\text{true}) &= 2/5 \\
P(B=\text{true} \mid A=\text{false}) &= 4/5 \\
P(A=\text{false}) &= 1/3 \\
P(B=\text{false} \mid A=\text{true}) &= 3/5 \\
P(B=\text{false} \mid A=\text{false}) &= 1/5
\end{align*}
\]

What is \(P(B=\text{true})\)?

a) 2/5  

b) 3/5  

c) 7/15  

d) 8/15  

e) None of the above.  

f) No idea.

What is \(P(A=\text{true} \mid B=\text{true})\)?

a) 2/3  

b) 3/7  

c) 4/7  

d) 4/15  

e) None of the above.  

f) No idea.
Logistics

**Instructor:** Daniel Lowd  
Email: lowd@cs.uoregon.edu  
Office: 262 Deschutes

**Teaching Assistant:** Jonathan Brophy  
Email: jbrophy@cs.uoregon.edu

**Office hours:**  
Monday 3-4:30pm (Jonathan, location TBD)  
Wednesday 10-11:30am (Jonathan, location TBD)  
Friday 11:30-1:00pm (Daniel, Deschutes 262)

**Piazza:** [http://piazza.com/uoregon/spring2018/cis473573](http://piazza.com/uoregon/spring2018/cis473573)  
(Primary source of up-to-date info.)

**Web:** [http://www.cs.uoregon.edu/classes/18S/cis473/](http://www.cs.uoregon.edu/classes/18S/cis473/)  
(Static page that’s sometimes out of date.)

Prerequisites

This course assumes a basic understanding of:

- Python programming
- Algorithms
- Discrete mathematics

If you have weaknesses in any of these areas, plan to put in extra time.

The following is not required, but very, very helpful:

- Probability
Textbook

- Complements:

Evaluation

- 6-8 homeworks: 40%
- Midterm: 25%
- Final: 35%

You may discuss homeworks with other students, but please write up your answers independently. Feel free to ask for help in office hours.
Expectations

• Study
  – Attend lecture. Take notes.
  – Read the book: Look at it before and after lecture.
  – Post questions on Piazza.
  – Recommended: Come to office hours.
  – Recommended: Watch videos on Coursera.

• Homeworks
  – Due weekly. Start early. Ask questions.
  – Workload: 5-15 hours/week outside of class
  – Late policy: You must ask in advance. Further in advance is better. I will usually say yes. Don’t abuse this policy, or I will start saying no.

Graduate vs. Undergraduate

Undergraduates: Understand how the methods work and when to use them.

Graduates: Understand why the methods work and how to adapt them.

Some of the homework and test questions will only be for graduate students to answer.
This Course in One Slide

You will learn:
• Basics of discrete probability (1.5 weeks)
• How to represent complex probability distributions using probabilistic graphical models (3 weeks)
• How to answer probability queries using probabilistic graphical models (3 weeks)
• How to fit probabilistic models to data (1 week)
• Motivating applications (throughout)

Applications
• Machine learning
• Data mining
• Automated reasoning and planning
• Vision and graphics
• Robotics
• Natural language processing and speech
• Information retrieval
• Databases and data management
• Networks and systems
• Ubiquitous computing
• Human-computer interaction
• Simulation
• Computational biology
• Computational neuroscience
• Etc.
Section 0: Probability

How do we formally describe and reason about uncertainty? (Focusing on discrete distributions over finite numbers of outcomes.)

- Sample spaces and events
- Random variables
- Conditional probability
- Bayes’ rule
- Independence and conditional independence

Section 1: Representation

How do we represent problems as probability distributions and encode them efficiently when there are many variables?

- Bayesian networks
- Markov networks
- Briefly, as time allows:
  - Template models (hidden Markov models, dynamic Bayesian networks, and relational models)
  - Mixture models and topic models
  - Tractable models (arithmetic circuits and sum-product networks)
  - Gaussian graphical models
Section 2: Inference

How do we use these probability distributions to answer questions efficiently?

• Why inference is hard
• Exact inference
  – Enumeration (brute force!)
  – Variable elimination
• Approximate inference
  – Loopy belief propagation
  – Sampling methods
  – Variational inference (as time allows)

Section 3: Learning (if time)

How can we use data to automatically pick the model’s structure and parameters?

• The basics of statistical estimation
• Learning Bayesian networks (parameters and structure)
• Learning Markov networks (parameters and structure)
• Learning with missing data
Applications Throughout

- Bayesian networks
  - Genetics
  - Evolutionary biology: Phylogenetic trees
  - Diagnosis: Medical diagnosis, fault diagnosis
  - Natural language processing: Language modeling, document clustering
- Markov networks
  - Computer vision: Image denoising, stereo vision
  - Molecular biology: Protein structure prediction
  - Social network analysis

Relations to Other Classes

- CIS 471/571: Artificial Intelligence
- CIS 472/572: Machine Learning
- CIS 453/553: Data Mining
- Application classes (e.g., Bioinformatics)
- Probability and Statistics