EECS 391: Introduction to AI

Soumya Ray

Website: http:// engr . case . edu / ray _ soumya / eecs391_sp16/
Email: sray @ case . edu
Office: Olin 516
Office hours: M 12:30-2pm
Announcements

• HW3 due
  – Solution posted tomorrow

• Quiz 3 Thursday
  – FOL and inference, planning
  – Syntax, semantics, lifted resolution, MGU
  – STRIPS, assumptions, state space and plan space planning

• Read Chapter 14 sections 1, 2, 4.1, 4.2, 5.1 (Bayesian networks)
Today

• Basic Probability and probabilistic inference (Ch 13)
• Review of midterm grades
# Recap

## Joint Probability Density Function

<table>
<thead>
<tr>
<th>CloudyTomorrow</th>
<th>RainTomorrow</th>
<th>WetGrass</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>No</td>
<td>No</td>
<td>0.4</td>
</tr>
<tr>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>0.01</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>0.01</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>0.01</td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>0.15</td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>0.02</td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>0.01</td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Atomic Event

Event

Sample Space
Recap: Random Variable (R.V.)

• A variable that refers to an uncertain fact
  – Analogous to proposition symbol
  – Has a domain that can be discrete or continuous
    • For this class, focus on discrete case

• For each value (or set of values), we can specify a degree of belief that shows how much we believe the stated fact---this is the probability associated with the fact
Recap: Atomic Event

• If the state of the world is described by $n$ r.v.’s and we assign values to all of them, this defines an atomic event
  – (Analog of a row in a truth table)
Recap: Axioms of Probability

• For any event $E$, $0 \leq \Pr(E) \leq 1$
• The probability of the sample space is 1
• For mutually disjoint events, the probability of the union is given by:

$$\Pr(\bigcup E_i) = \sum_{i=1}^{\infty} \Pr(E_i)$$

In particular this must apply to atomic events.
Recap: Joint PDF

- Using joint probability, we can define joint density functions for collections of random variables

\[ p_{R,C}(R = x, C = y) = \begin{cases} 
0.5 & \text{if } x = \text{Yes}, y = \text{Yes} \\
0.2 & \text{if } x = \text{No}, y = \text{Yes} \\
0.2 & \text{if } x = \text{Yes}, y = \text{No} \\
0.1 & \text{if } x = \text{No}, y = \text{No} 
\end{cases} \]
Conditional Probability

- The conditional probability of $X$ given $Y$ is:

$$p_{X|Y}(X = x \mid Y = y) = \frac{p_{X,Y}(X = x, Y = y)}{p_Y(Y = y)}$$

$X = x$, $Y = y$ (""," means AND)
Product Rule

• From the definition of conditional probability:

\[ p_{X,Y}(X = x, Y = y) = \]
\[ p_Y(Y = y) p_{X|Y}(X = x \mid Y = y) \]
Marginalization

• For any two random variables $X$ and $Y$:

$$p_X(X = x) = \sum_y p_{X,Y}(X = x, Y = y)$$
Conditioning

\[ p(X = x) = \sum_y p(X = x, Y = y) \quad \text{Marginalization} \]

\[ = \sum_y p(X = x \mid Y = y)p(Y = y) \quad \text{Product Rule} \]
Bayes’ Rule
(Rev. Thomas Bayes 1763)

\[ p(C = c \mid E = e) = \frac{p(C = c, E = e)}{p(E = e)} \]

\[ = \frac{p(E = e \mid C = c) p(C = c)}{p(E = e)} \]

\[ = \frac{p(E = e \mid C = c) p(C = c)}{\sum_{c'} p(E = e \mid C = c') p(C = c')} \]
The importance of Bayes Rule

- Let $C$ be a random variable with values that are possible “causes”
- Let $E$ denote a random variable with values that are possible effects of each cause
- It is often easy to specify $p(E=e|C=c)$, much harder to specify $p(C=c|E=e)$
- Bayes Rule therefore allows us to reason backwards over uncertain events---fundamental to learning
Example

• Lung cancer can be caused by smoking or by a genetic defect. 5% of the population are smokers. 2 in 3 who smoke and 1 in 100 who don’t get the disease.

• Suppose X has lung cancer. What is the probability X smokes?
Example

\[
P(S) = 0.05, \quad P(LC \mid S) = 0.67, \quad P(LC \mid \bar{S}) = 0.01
\]

\[
P(S \mid LC) = \frac{P(LC \mid S)P(S)}{P(LC \mid S)P(S) + P(LC \mid \bar{S})P(\bar{S})}
\]

\[
= \frac{0.67 \times 0.05}{0.67 \times 0.05 + 0.01 \times 0.95} = 0.78
\]
Statistical Independence

• Two r.v.’s $X$ and $Y$ are statistically independent if

$$p_{X,Y}(X = x, Y = y) = p_X(X = x) p_Y(Y = y)$$

• If so, we can reason separately about $x$ and $y$ and then combine results---key factor in gaining efficiency (later)
Consequence

\[ p_{X|Y}(X = x \mid Y = y) = \frac{p_{X,Y}(X = x, Y = y)}{p_Y(Y = y)} \]

\[ = \frac{p_X(X = x)p_Y(Y = y)}{p_Y(Y = y)} \]

\[ = p_X(X = x) \]
Conditional Independence

• Two r.v.’s $X$ and $Y$ are conditionally independent given a third, $R$, if

$$p_{X,Y|R}(X = x, Y = y \mid R = r) = p_{X|R}(X = x \mid R = r)p_{Y|R}(Y = y \mid R = r)$$
Summarizing a PDF

• A PDF is a large table of numbers

• But generally, we don’t need to know the entire thing; often the “highlights” are enough
  – *Expectation* and *Variance*
  – (statistics)
Expectation

- The expectation of r.v. $X$ is defined as:

$$E(X) = \sum_{x} x \cdot p_{X}(x)$$

- The “average value” of $X$ under $p_{X}(x)$
Expectation example

• A coin has 0.99 probability of showing heads. You get $0 if the coin shows heads, and $10 else. How much do you expect to get if I toss the coin?

\[
E(X) = \sum x p_X(x) = (0 \times 0.99 + 10 \times 0.01) = \$0.1
\]

>> A=rand(1000,1); A=(A<=0.01)*10; sum(A)/1000
ans =
0.1100
Expectation of a function

- The expectation of a function $g(X)$ is defined as:

$$E(g(X)) = \sum_{x} g(x) p_X(x)$$
Variance

- The variance of r.v. $X$ is defined as:

$$V(X) = E([X - E(X)]^2)$$

$$= \sum_x (x - E(X))^2 p_X(x)$$

- The “average spread” of values of a r.v. around the average of the r.v.
Variance example

• A coin has 0.99 probability of showing heads. You get $0 if the coin shows heads, and $10 else. What is the variance of your takings?

\[
E(X) = \sum_{x} x p_X(x) = (0 \times 0.99 + 10 \times 0.01) = 0.1
\]

\[
V(X) = E([X - E(X)]^2)
\]

\[
= (0 - 0.1)^2 \times 0.99 + (10 - 0.1)^2 \times 0.01
\]

\[
= 0.99
\]
Variance example

- A coin has 0.99 probability of showing heads. You get $10 if the coin shows heads, and $0 else. What is the variance of your takings?

\[
E(X) = \sum_x xp_X(x) = (10 \times 0.99 + 0 \times 0.01) = $9.9
\]

\[
V(X) = E([X - E(X)]^2)
\]

\[
= (10 - 9.9)^2 \times 0.99 + (0 - 9.9)^2 \times 0.01
\]

\[
= 0.99
\]
Variance example 3

• A coin has 0.5 probability of showing heads. You get $0 if the coin shows heads, and $10 else. What is the variance of your takings?

\[ E(X) = \sum \limits_x x p_X(x) = (0 \times 0.5 + 10 \times 0.5) = $5 \]

\[ V(X) = E([X - E(X)]^2) \]

\[ = (0 - 5)^2 \times 0.5 + (10 - 5)^2 \times 0.5 \]

\[ = 25 \]