CS 112: Modeling Uncertainty in Information Systems

Prof. Jenn Wortman Vaughan

April 23, 2012

Lecture 7

Reminders & Announcements

- Homework 2 was posted on the website last week and is due in class on Monday, April 30
- Solutions to homework 1 are available on Piazza (login required)

Reminders & Announcements

- The midterm is coming up on Wednesday, May 2 (one week from this Wednesday!)
- It will cover everything through the first half of Wednesday's class (including correlation and covariance)
- We'll start talking about the application of probability to data compression on Wednesday (and homework 3 will have a related programming component) but this will not be on the exam

Reminders & Announcements

- There will be no discussion section this Friday, but Jacob will hold several review sessions for the midterm on Wednesday and Thursday

  If you haven't emailed Jacob your availability yet, do it by the end of the day today!!

Today...

- Conditional PMFs
- Conditional expectation
- The Total Expectation Theorem (divide & conquer)

- Independence of random variables

Conditioning

- The conditional PMF of X given Y is denoted $p_{X|Y}$. For any values k of X and l of Y,

$$p_{X|Y}(k \mid l) = P(\{X=k\}|\{Y=l\}) = P(X=k \mid Y=l)$$

- We can compute this PMF using the definition of conditional probability

Conditioning

Conditional PMFs frequently give us a convenient method of calculating joint PMFs. Using the multiplication rule,

$$p_{X,Y}(k,l) = p_X(k)p_{Y|X}(l|k) = p_Y(l)p_{X|Y}(k|l)$$

Conditioning

Conditional PMFs frequently give us a convenient method of calculating joint PMFs. Using the multiplication rule,

$$p_{X,Y}(k,l) = p_X(k)p_{Y|X}(l|k) = p_Y(l)p_{X|Y}(k|l)$$

- Example: If my computer has a virus, which is true with probability 0.2, my anti-virus program detects it with probability 0.7. Otherwise, the program incorrectly identifies a virus with probability 0.1.

  Let X and Y be binary variables that are 1 if my computer has a virus and if the program says it has a virus respectively. What is $p_{X,Y}$?

Divide & Conquer: Total Probability

A divide & conquer technique to compute marginals:

$$p_Y(l) = \sum_k p_{X,Y}(k,l) = \sum_k p_X(k)\, p_{Y|X}(l \mid k)$$

Divide & Conquer: Total Probability

A divide & conquer technique to compute marginals:

$$p_Y(l) = \sum_k p_{X,Y}(k,l) = \sum_k p_X(k)\, p_{Y|X}(l \mid k)$$

- Example: If my computer has a virus, which is true with probability 0.2, my anti-virus program detects it with probability 0.7. Otherwise, the program incorrectly identifies a virus with probability 0.1.

  Let X and Y be binary variables that are 1 if my computer has a virus and if the program says it has a virus respectively. What is p_Y?

Conditioning

We can condition random variables on events too. For a random variable X and event A associated with the same experiment,

$$p_{X|A}(k) = P(X=k \mid A) = P(\{X=k\} \mid A)$$

Conditioning

We can condition random variables on events too. For a random variable X and event A associated with the same experiment,

$$p_{X|A}(k) = P(X=k \mid A) = P(\{X=k\} \mid A)$$

Example: X is "sum of two die rolls", A = "first roll is 1"

Conditioning

We can condition random variables on events too. For a random variable X and event A associated with the same experiment,

$$p_{X|A}(k) = P(X=k \mid A) = P(\{X=k\} \mid A)$$

Example: X is "sum of two die rolls", A = "first roll is 1"

Nothing fancy here. All of the same reasoning applies.

Conditioning

We can condition random variables on events too. For a random variable X and event A associated with the same experiment,

$$p_{X|A}(k) = P(X=k \mid A) = P(\{X=k\} \mid A)$$

Example: X is "sum of two die rolls", A = "first roll is 1"

Nothing fancy here. All of the same reasoning applies.

Remember: All of this is just notation. When in doubt, think about the underlying events.

Conditional Expectation

We can define a conditional version of expectation:

$$E[X \mid Y = l] = ???$$

Conditional Expectation

We can define a conditional version of expectation:

$$E[X \mid Y = l] = \sum_k k p_{X|Y}(k \mid l)$$

Conditional Expectation

We can define a conditional version of expectation:

$$E[X \mid Y = l] = \sum_k k p_{X|Y}(k \mid l)$$

For any function g:

$$E[g(X) \mid Y = l] = ???$$

Conditional Expectation

We can define a conditional version of expectation:

$$E[X \mid Y = l] = \sum_k k p_{X|Y}(k \mid l)$$

For any function g:

$$E[g(X) \mid Y = l] = \sum_k g(k) p_{X|Y}(k \mid l)$$

Conditional Expectation

We can define a conditional version of expectation:

$$E[X \mid Y = l] = \sum_k k p_{X|Y}(k \mid l)$$

For any function g:

$$E[g(X) \mid Y = l] = \sum_k g(k) p_{X|Y}(k \mid l)$$

Expectations can be conditioned on events A too.

Divide & Conquer: Total Expectation

The Total Expectation Theorem says that for any random variables X and Y,

$$E[X] = \sum_{l} p_Y(l) E[X \mid Y = l]$$

Divide & Conquer: Total Expectation

The Total Expectation Theorem says that for any random variables X and Y,

$$E[X] = \sum_{l} p_Y(l) E[X \mid Y = l]$$

More generally, for any disjoint events $A_1, \ldots, A_n$ that form a partition of Ω,

$$E[X] = \sum_{i=1}^{n} P(A_i) E[X \mid A_i]$$

Divide & Conquer: Total Expectation

The Total Expectation Theorem says that for any random variables X and Y,

$$E[X] = \sum_{l} p_Y(l) E[X \mid Y = l]$$

More generally, for any disjoint events $A_1, \ldots, A_n$ that form a partition of Ω,

$$E[X] = \sum_{i=1}^{n} P(A_i) E[X \mid A_i]$$

We could prove this using the multiplication rule and the definition of conditional expectation – Try this at home!

Divide & Conquer: Total Expectation

Let X be a geometric random variable with parameter p. What is E[X]?

Independence of Random Variables

Independence

- Recall that events A and B are independent if and only if

$$P(A \cap B) = P(A)\,P(B)$$

- If $P(B) > 0$, this condition is equivalent to

$$P(A \mid B) = P(A)$$

- If $P(A) > 0$, it is equivalent to

$$P(B \mid A) = P(B)$$

To prove or disprove independence, it is enough to prove or disprove any one of these conditions.

Independence

- X and Y are independent if and only if for all k and l

$$p_{X,Y}(k,l) = p_X(k)\, p_Y(l)$$

Independence

- X and Y are independent if and only if for all k and l

$$p_{X,Y}(k,l) = p_X(k)\, p_Y(l)$$

- This is equivalent to showing that for all k, l s.t. $p_Y(l)>0$

$$p_X(k) = p_{X|Y}(k \mid l)$$

Independence

- X and Y are independent if and only if for all k and l

$$p_{X,Y}(k,l) = p_X(k)\, p_Y(l)$$

- This is equivalent to showing that for all k, l s.t. $p_Y(l)>0$

$$p_X(k) = p_{X|Y}(k \mid l)$$

- This is equivalent to showing that for all k, l s.t. $p_X(k)>0$

$$p_Y(l) = p_{Y|X}(l \mid k)$$

Independence

- X and Y are independent if and only if for all k and l

$$p_{X,Y}(k,l) = p_X(k)\, p_Y(l)$$

- This is equivalent to showing that for all k, l s.t. $p_Y(l)>0$

$$p_X(k) = p_{X|Y}(k \mid l)$$

- This is equivalent to showing that for all k, l s.t. $p_X(k)>0$

$$p_Y(l) = p_{Y|X}(l \mid k)$$

To prove or disprove independence, it is enough to prove or disprove any one of these conditions.

Independence

		Y			
		l_1	l_2	l_3	l_4
X	k_1	0.05	0.15	0	0.2
	k_2	0.025	0.075	0	0.1
	k_3	0.05	0.15	0	0.2

Are X and Y independent? How can you tell?

Independence

		Y			
		l_1	l_2	l_3	l_4
X	k_1	0.05	0.15	0	0.2
	k_2	0.025	0.075	0	0.1
	k_3	0.05	0.15	0	0.2

Columns are multiples of each other, which implies that $p_{X|Y}(k \mid l)$ is the same for all l

Independence

It is easy to show that if X and Y are independent, then

$$E[XY] = E[X]\, E[Y]$$

Independence

It is easy to show that if X and Y are independent, then

$$E[XY] = E[X]\,E[Y]$$

Note that this is not true in general, though it *can hold* in some cases even if X and Y are not independent.

(See the example on Piazza…)

Independence

The definition of independence can be extended in the natural way to sets of more than two random variables

$$p_{X,Y,Z}(k, l, m) = p_X(k)\, p_Y(l)\, p_Z(m)$$

Independence

The definition of independence can be extended in the natural way to sets of more than two random variables

$$p_{X,Y,Z}(k, l, m) = p_X(k)\, p_Y(l)\, p_Z(m)$$

If X, Y, and Z are independent, then any three random variables of the form f(X), g(Y), and h(Z) would be independent too.

Independence

The definition of independence can be extended in the natural way to sets of more than two random variables

$$p_{X,Y,Z}(k, l, m) = p_X(k)\, p_Y(l)\, p_Z(m)$$

If X, Y, and Z are independent, then any three random variables of the form f(X), g(Y), and h(Z) would be independent too.

If $X_1, X_2, \ldots, X_n$ are independent, then

$$E[X_1\, X_2 \ldots X_n] = E[X_1]\, E[X_2] \ldots E[X_n]$$

One more useful fact…

If $X_1, X_2, \ldots, X_n$ are independent, then

$$\mathrm{var}(X_1 + X_2 + \ldots + X_n) = \mathrm{var}(X_1) + \mathrm{var}(X_2) + \ldots + \mathrm{var}(X_n)$$

One more useful fact…

If $X_1, X_2, \ldots, X_n$ are independent, then

$$\mathrm{var}(X_1 + X_2 + \ldots + X_n) = \mathrm{var}(X_1) + \mathrm{var}(X_2) + \ldots + \mathrm{var}(X_n)$$

This also is not true in general – need independence!

The Sample Mean

Suppose we would like to estimate the president's approval rating. We ask n random voters whether or not they approve of the president, and use the fraction of voters who say that they approve as our estimate.

How accurate is our estimate as a function of n?

The Sample Mean

Suppose we would like to estimate the president's approval rating. We ask n random voters whether or not they approve of the president, and use the fraction of voters who say that they approve as our estimate.

How accurate is our estimate as a function of n?

(We'll finish this example at the start of the next class…)