CHEM 116 February 16, 2016

Unit 2, Lecture 2

Numerical Methods and Statistics

1 Working with Marginals, Condtionals, and Joints

1.0.1 Seasons Example

The sample space is a product space of the seasons T (Winter (0), Spring (1), Summer (2), Fall (3)) and W if the weather is nice or not (N=nice, S=not nice). We know that

$$\begin{array}{ll} \mathrm{P}(W=N|T=0)=0 & \mathrm{P}(W=N|T=1)=0.4 \\ \mathrm{P}(W=N|T=2)=0.8 & \mathrm{P}(W=N|T=3)=0.7 \end{array}$$

$$P(T=t)=\frac{1}{4}$$

What is the probability that the weather is not nice? Use marginalization of conditional:

$$\mathrm{P}(W=S)=\sum_{\mathcal{T}}\mathrm{P}(W=S|T)P(T)=0.25(1-0)+0.25(1-0.4)+0.25(1-0.8)+0.25(1-0.7)$$

$$\mathrm{P}(W=S)=0.525$$

What is the probability that it is fall given that it is nice? Start with definition of conditional :

$$\mathrm{P}(T=3|W=N)=\frac{\mathrm{P}(T=3,W=N)}{\mathrm{P}(W=N)}$$

We know that $\mathrm{P}(T=3,W=N)=\mathrm{P}(W=N|T=3)\mathrm{P}(T=3)$ due to definition of conditional:

$$\mathrm{P}(T=3|W=N)=\frac{\mathrm{P}(W=N|T=3)\mathrm{P}(T=3)}{\mathrm{P}(W=N)}$$

Finally, we can use $\mathrm{P}(W=N)=1-\mathrm{P}(W=S)$, the **NOT** rule:

$$\mathrm{P}(T=3|W=N)=\frac{0.7\times 0.25}{1-0.525}=0.368$$

1.1 Bayes' Theorem

We derived a well-known equation in that example called Bayes' Theorem:

$$\mathrm{P}(A|B)=\frac{\mathrm{P}(B|A)\mathrm{P}(A)}{\mathrm{P}(B)} \tag{1}$$

This is useful to swap the order of conditionals

1.2 Independence

Finally, we are ready to define independence. If the random variables X and Y are independent, then

$$\mathrm{P}(x|y) = \mathrm{P}(x) \tag{2}$$

This implies via Bayes' Theorem

$$\mathrm{P}(y|x) = \mathrm{P}(y) \tag{3}$$

And also implies via CPF definition

$$\mathrm{P}(x, y) = \mathrm{P}(x)\mathrm{P}(y) \tag{4}$$

which was our **AND** rule from last lecture.

In our weather example, is the season and weather independent?

$$\mathrm{P}(W = N|T = 0) \neq P(W = N|T = 1)$$

so no.

1.3 Compound Conditionals

When writing conditionals, this is a common short-hand:

$$\mathrm{P}(x = 2 \,|\, Y = A, Z = 0)$$

for

$$\mathrm{P}([x = 2] \,|\, [Y = A, Z = 0])$$

The conditional is always evaluated last, for example:

$$\mathrm{P}(x = 2, Y = A \,|\, Z = 0)$$

is also possible and means the probability of the joint $(x = 2, Y = A)$ given that $Z = 0$

1.4 Conditional Independence

X_0 and X_1 are conditionally independent given Z if

$$\mathrm{P}(X_0|Z, X_1) = \mathrm{P}(X_0|Z) \tag{5}$$

which is equivalent to

$$\mathrm{P}(X_0, X_1|Z) = \mathrm{P}(X_0|Z)\mathrm{P}(X_1|Z) \tag{6}$$

To use conditional independence, you must condition on Z. For example, if I want to calculate $\mathrm{P}(X_0, X_1)$, I'll need to condition it. Marginalizing a conditional is one way to get that quantity, using a compound conditional:

$$\mathrm{P}(X_0, X_1) = \sum_{\mathcal{Z}} \mathrm{P}(X_0, X_1|Z)\mathrm{P}(Z)$$

Now it is in a form where the conditional independence applies:

$$\mathrm{P}(X_0, X_1) = \sum_{\mathcal{Z}} \mathrm{P}(X_0|Z)\mathrm{P}(X_1|Z)$$

This is common for sequential trials, where the trials are independent when conditioned on some underlying property, but dependent if we do not know the property. For example, let's say I have two dice, one that is fair and one that follows the biased die model we saw in class. Let's further assume that $P(D = 0) = 0.1$, where D indicates the chosen die.

$$\mathrm{P}(X = x|D = 0) = \frac{1}{6}$$
$$\mathrm{P}(X = x|D = 1) = \frac{x}{21}$$

If I know which die I'm rolling, then every roll is independent as expected:

$$\mathrm{P}(X_0 = 6, X_1 = 1|D) = \mathrm{P}(X_0 = 6|D)\mathrm{P}(X_1 = 1|D)$$

however, consider

$$\mathrm{P}(X_1 = 1|X_0 = 6) = \frac{\mathrm{P}(X_1 = 1, X_0 = 6)}{\mathrm{P}(X_0 = 6)}$$

As above, we'll try to condition it to exploit conditional independence.

$$\mathrm{P}(X_1 = 1, X_0 = 6) = \mathrm{P}(X_1 = 1|D = 0)\mathrm{P}(X_0 = 6|D = 0)\mathrm{P}(D = 0)+$$
$$\mathrm{P}(X_1 = 1|D = 1)\mathrm{P}(X_0 = 6|D = 1)\mathrm{P}(D = 1)$$
$$\mathrm{P}(X_1 = 1, X_0 = 6) = \frac{1}{6}\frac{1}{6}\frac{1}{10} + \frac{1}{21}\frac{6}{21}\frac{9}{10} = 0.0150$$
$$\mathrm{P}(X_0 = 6) = \sum_{X_1} \mathrm{P}(X_1, X_0 = 6) = \frac{1}{6}\frac{1}{10}\underbrace{\sum_x \frac{1}{6}}_{=1} + \frac{6}{21}\frac{9}{10}\underbrace{\sum_x \frac{x}{21}}_{=1}$$
$$\mathrm{P}(X_0 = 6) = 0.274$$
$$\mathrm{P}(X_1 = 1|X_0 = 6) = \frac{0.0150}{0.274} = 0.0549$$

Finally, to show they are not independent:

$$\mathrm{P}(X_1 = 1) = \sum_{X_1} \mathrm{P}(X_1 = 1, X_0) = \frac{1}{6}\frac{1}{10}\underbrace{\sum_x \frac{1}{6}}_{=1} + \frac{1}{21}\frac{9}{10}\underbrace{\sum_x \frac{x}{21}}_{=1} = 0.0595$$

The intuition here is that the marginal of $X_1 = 1$ considers both die according to the die marginals. However, knowing that $X_0 = 6$ shifts it so that the biased die is more likely. This can be seen with Bayes theorem to:

$$\mathrm{P}(D = 1|X_0 = 6) \neq \mathrm{P}(D = 1)$$

and thus $\mathrm{P}(X_1|X_0 = 6)$ changes from $\mathrm{P}(X_1)$.

1.5 Table of Equations

$P(x) = \sum_y P(x, y)$	Definition of Marginal
$P(x\|y) = \frac{P(x,y)}{P(y)}$	Definition of Conditional
$P(x, y) = P(x\|y)P(y)$	Definition of Conditional
$\sum_x P(x\|y) = 1$	Law of Total Probability/Normalization
$\sum_y P(x\|y)P(y) = P(x)$	Marginilzation of Conditional
$P(x\|y) = \frac{P(y\|x)P(x)}{P(y)}$	Bayes' Theorem
$P(x, y) = P(x)P(y)$	Definition of Independence
$P(x\|y) = P(x)$	Independence Property (x,y independent)
$P(x, y\|z) = P(x\|z)P(y\|z)$	Conditional Independence