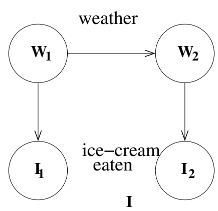
CS188 Fall 2015 Section 7: Sampling and Markov Models

Sampling

The Bayes Net in the diagram below describes a person's ice-cream eating habits. The nodes W_1 and W_2 stand for the weather on days 1 and 2, which can either be rainy R or sunny S, The nodes I_1 and I_2 stand for whether or not the person ate ice-cream that day. They can take the values T or F. The conditional probability distributions relevant to the graphical model are also given to you, note that there is a single conditional probability distribution $P(I|W)$, which I_1 and I_2 follow.

S	R
0.4	0.6



W_1 \ W_2	S	R
S	0.7	0.3
R	0.5	0.5

W \ I	T	F
S	0.9	0.1
R	0.2	0.8

Suppose we use prior sampling to produce the following samples from the weather/ice-cream model:

(W_1, I_1, W_2, I_2)
R, F, R, F excluded by rejection sampling
S, F, S, T excluded by rejection sampling
S, T, S, T excluded by rejection sampling
S, T, S, T excluded by rejection sampling
S, T, R, F

(W_1, I_1, W_2, I_2)
R, F, R, F excluded by rejection sampling
S, T, S, T excluded by rejection sampling
R, F, R, T excluded by rejection sampling
S, T, R, F
R, F, S, T excluded by rejection sampling

1. What is $\hat{P}(W_2 = R)$?
   Number of samples in which $W_2 = R$: 5. Total number of samples: 10. Answer $5/10 = 0.5$.

2. Cross off samples rejected by rejection sampling if we are computing $\hat{P}(W_2|I_1 = T, I_2 = F)$

   Now use likelihood weighting, and assume we've generated the following samples, given the evidence $I_1 = T$ and $I_2 = F$.

(W_1, I_1, W_2, I_2)	(W_1, I_1, W_2, I_2)
S, T, R, F	S, T, S, F
R, T, R, F	S, T, S, F
S, T, R, F	R, T, S, F

3. What is the weight of the first sample (S,T,R,F) above?

   The weight given to a sample in likelihood weighting is $\prod_{\text{evidence variables } e} \Pr(e|\text{Parents}(e))$. In this case, the evidence is $I_1 = T, I_2 = F$. The weight of the first sample is therefore $w = \Pr(I_1 = T|W_1 = S) \cdot \Pr(I_2 = F|W_2 = R) = 0.9 \cdot 0.8 = 0.72$.

4. Use likelihood weighting to estimate $\hat{P}(W_2|I_1 = T, I_2 = F)$

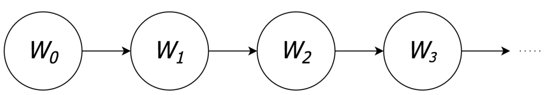
   The sample weights are given by

(W_1, I_1, W_2, I_2)	w	(W_1, I_1, W_2, I_2)	w
S, T, R, F	0.72	S, T, S, F	0.09
R, T, R, F	0.16	S, T, S, F	0.09
S, T, R, F	0.72	R, T, S, F	0.02

$$\hat{\Pr}(W_2 = R|I_1 = T, I_2 = F) = \frac{0.72 + 0.16 + 0.72}{0.72 + 0.16 + 0.72 + 0.09 + 0.09 + 0.02} = 0.889$$
$$\hat{\Pr}(W_2 = S|I_1 = T, I_2 = F) = 1 - 0.889 = 0.111.$$

Markov Model

We want to represent Berkeley's weather by Markov Model where we assume that the weather at day t is independent to the weather at day $0, ..., t-2$ given the weather at day $t-1$. We have the following initial distribution and transition model.



W_0	$P(W_0)$
sun	0.9
rain	0.1

W_i	W_{i-1}	$P(W_i|W_{i-1})$
sun	sun	0.8
rain	sun	0.2
sun	rain	0.3
rain	rain	0.7

Compute the following.

1. $P(W_1 = sun)$

   $\sum_{w_0} P(w_0, W_1 = sun) = \sum_{w_0} P(W_1 = sun|w_0)P(w_0) = 0.8 * 0.9 + 0.3 * 0.1 = .75$

2. $P(W_2 = sun)$

   $\sum_{w_1} P(w_1, W_2 = sun) = \sum_{w_1} P(W_2 = sun|w_1)P(w_1) = 0.8 * 0.75 + 0.3 * 0.25 = .675$

3. $P(W_\infty = sum)$ (stationary distribution)

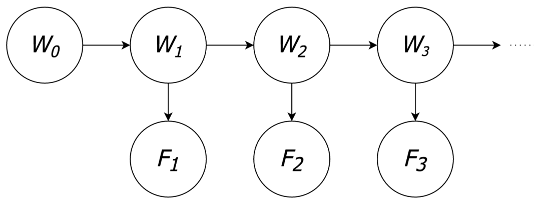
   By transition model,

$$P(W_\infty = sun) = P(W_{\infty+1} = sun) = \sum_{w_\infty} P(W_{\infty+1} = sun|w_\infty)P(w_\infty) = 0.8P(W_\infty = sun) + 0.3P(W_\infty = rain)$$

   Also, we know that $P(W_\infty = sun) + P(W_\infty = rain) = 1$. By solving a system of equations, we get $P(W_\infty = sun) = 0.6$

Hidden Markov Model

You are stuck on the second floor of Soda Hall working on the project for days. You have decided to not leave the building until you finish the project but wants to know what the weather is like outside. The only source of information you have is the thermometer in the room. We have a following initial distribution, transition model, and sensor model. Our initial distribution and transition model are same as the last problem.



W_0	$P(W_0)$
sun	0.9
rain	0.1

W_i	W_{i-1}	$P(W_i\|W_{i-1})$
sun	sun	0.8
rain	sun	0.2
sun	rain	0.3
rain	rain	0.7

F_i	W_i	$P(F_i\|W_i)$
high	sun	0.7
low	sun	0.3
high	rain	0.2
low	rain	0.8

1. We observed that $F_1 = high$ and computed $P(W_1)$ from previous section, what is the updated probability $P(W_1 = sum|F_1 = high)$?

$$P(W_1 = sun|F_1 = high) = \frac{P(W_1 = sun, F_1 = high)}{P(F_1 = high)}$$
$$= \frac{P(F_1 = high|W_1 = sun)P(W_1 = sun)}{\sum_{w_1} P(F_1 = high|w_1)P(w_1)} = \frac{0.7 * 0.75}{0.7 * 0.75 + 0.2 * 0.25} = 0.913$$

2. We want to predict tomorrow's weather based on our observation. What is $P(W_2 = sun|F_1 = high)$?

$$P(W_2 = sun|F_1 = high) = \sum_{w_1} P(W_2 = sun, w_1|F_1 = high)$$
$$= \sum P(W_2 = sun|w_1)P(w_1|F_1 = high)$$
$$= 0.8 * .913 + 0.3 * (1 - .913) = .757$$

3. Given $P(w_i|f_1, ..., f_i)$ for all w_i, find $P(W_{i+1}|f_1, ..., f_i)$.

$$P(W_{i+1}|f_1, ...f_i) = \sum_{w_i} P(W_{i+1}, w_i|f_1, .., f_i) = \sum_{w_i} P(W_{i+1}|w_i)P(w_i|f_1, ..., f_i)$$

4. **Forward Algorithm:** Given $P(w_i|f_1, ..., f_i)$ for all w_i and new observation f_{i+1}, find $P(W_{i+1}|f_1, .., f_{i+1})$.

By conditional probability,

$$
\begin{aligned}
P(W_{i+1}|f_1,...f_i,f_{i+1}) &= \frac{P(W_{i+1},f_{i+1}|f_1,...,f_i)}{P(f_{i+1}|f_1,...,f_i)} \\
&= \frac{P(f_{i+1}|W_{i+1},f_1,...,f_i)P(W_{i+1}|f_1,...,f_i)}{P(f_{i+1}|f_1,...,f_i)} \\
&= \frac{P(f_{i+1}|W_{i+1})\sum_{w_i}P(W_{i+1}|w_i)P(w_i|f_1,...,f_i)}{P(f_{i+1}|f_1,...,f_i)} \\
&= \alpha P(f_{i+1}|W_{i+1})\sum_{w_i}P(W_{i+1}|w_i)P(w_i|f_1,...,f_i), \alpha = \frac{1}{P(f_{i+1}|f_1,...,f_i)}
\end{aligned}
$$

On the first line we used conditional probability for W_{i+1} and f_{i+1}. For the second line, we used product rule for W_{i+1}, f_{i+1}. On the third line, we replaced $P(f_{i+1}|W_{i+1},f_1,...,f_i)$ with $P(f_{i+1}|W_{i+1})$ by Markov Chain assumption and $P(W_{i+1}|f_1,...,f_i)$ with what we computed on problem 3. Lastly, $\alpha = \frac{1}{P(f_{i+1}|f_1,...,f_i)}$ does not have to be computed directly, but can be normalized after computing all $P(W_{i+1}|f_1,...f_i,f_{i+1})$.