EECS545 Lecture 20 Quiz Solutions

1. Given a sequence of states `[Sunny, Sunny, Rainy, Sunny, Rainy, Rainy, Sunny, Sunny, Sunny]` and a sequence of observations (precipitation in inches) `[0, 0, 0.2, 0, 0.1, 0.3, 0, 0, 0]`, estimate the observation probability distribution of precipitation on Rainy days $P(\text{precipitation}|\text{Rainy}) = \mathcal{N}(\mu, \sigma^2)$.

> **Solution:**
>
> $$\begin{aligned}\mu &= \text{Sample mean of precipitation when it is Rainy}\\ &= \frac{1}{3}(0.2 + 0.1 + 0.3)\\ &= 0.2\\ \sigma &= \text{Sample standard deviation of precipitation when it is Rainy}\\ &= \sqrt{\frac{1}{3-1}((0.2-\mu)^2 + (0.1-\mu)^2 + (0.3-\mu)^2)}\\ &= \sqrt{\frac{1}{2}((-0.1)^2 + 0.1^2)}\\ &= 0.1\end{aligned}$$

2. Continued. Estimate the probability of transitioning `Sunny` $\rightarrow$ `Rainy`.

> **Solution:** We use Bayes Rule.
>
> $$P(\texttt{Sunny} \rightarrow \texttt{Rainy}) = \frac{\#\ \texttt{Sunny} \rightarrow \texttt{Rainy}}{\#\ \texttt{Sunny} \rightarrow \texttt{Sunny} + \#\ \texttt{Sunny} \rightarrow \texttt{Rainy}} = \frac{2}{5}$$

3. Which of the following statements are true about HMMs and RNNs? (Select all that apply)
   (a) HMM models a hidden state in data but RNN does not
   (b) HMM models an emission probability from hidden state but RNN does not
   (c) For HMM, you are guaranteed to converge to a local optimum, and it requires less tuning than RNNs.
   (d) HMM can model the data with richer set of hidden states and dynamics.
   (e) HMMs require fewer data to train

Solution: (c) and (e).

4. True or False. When learning an HMM for a fixed set of observations, assume we do not know the true number of hidden states (which is often the case), we can always monotonically increase the training data likelihood by permitting more hidden states.

Solution: True. Intuitively, if we increase the number of hidden states to the extreme, a hidden state for every timestep, then we can "fit" that the transition and emission probabilities perfectly for every state, making the likelihood 1! Increasing the number of hidden states in general similarly increases the likelihood but at a smaller scale.