Congratulations! You passed!

TO PASS 70% or higher

<div>Keep Learning</div>

GRADE

97.50%

Week 2 Quiz

LATEST SUBMISSION GRADE

97.5%

1. You are given a unigram language model θ distributed over a vocabulary set V composed of **only** 4 words: "the", "global", "warming", and "effects". The distribution of θ is given in the table below:

   1 / 1 point

w	$P(w|\theta)$
the	0.3
global	0.2
warming	0.2
effects	X

   What is X, i.e., $P(\text{"effects"}|\theta)$?

   ✓ **Correct**

2. Assume you are given the same unigram language model as in Question 1. Which of the following is **not** true?

   1 / 1 point

   ✓ **Correct**

3.

   1 / 1 point

Assume that words are being generated by a mixture of two unigram language models, θ_1 and θ_2, where $P(\theta_1) = 0.5$ and $P(\theta_2) = 0.5$. The distributions of the two models are given in the table below:

w	$P(w|\theta_1)$	$P(w|\theta_2)$
sports	0.35	0.05
basketball	0.2	0.05
fast	0.3	0.3
computer	0.1	0.4
smartphone	0.05	0.2

Then the probability of observing "computer" from this mixture model is: $P(\text{"computer"}) =$

✓   **Correct**

4.   Assume the same given as in Question 3. We now want to infer which of the two word distributions, θ_1 and θ_2, has been used to generate "computer", and would thus like to compute the probability that it has been generated using θ_1 and θ_2, i.e., $P(\theta_1|\text{"computer"})$ and $P(\theta_2|\text{"computer"})$, respectively, then the values of $P(\theta_1|\text{"computer"})$ and $P(\theta_2|\text{"computer"})$ are: **1 / 1 point**

Hint: Apply Bayes rule.

✓   **Correct**

5.   Suppose words are being generated using a mixture of two unigram language models θ_1 and θ_2. Let $P(w)$ denote the probability of generating a word w from this mixture model.   **1 / 1 point**

If $P(\theta_1) = 1$ then which of the following statements is true?

✓   **Correct**

6.   **1 / 1 point**

True or false? Let X_{text}, X_{mining}, and X_{the} be binary random variables associated with the words "text", "mining", and "the", respectively. Assume that the probabilities of the random variables are estimated based on a large corpus. Then we should expect $H(X_{text}|X_{mining}) > H(X_{text}|X_{the})$.

✓ **Correct**

7. True or false? I(X;Y)=0 if and only if X and Y are independent.   **1 / 1 point**

✓ **Correct**

8. Let w be a word and X_w be a binary random variable that indicates whether w appears in a text document in the corpus. Assume that the probability $P(X_w = 1)$ is estimated by Count(w)/N, where Count(w) is the number of documents w appears in and N is the total number of documents in the corpus.   **1 / 1 point**

   You are given that "the" is a very frequent word that appears in 99% of the documents and that "photon" is a very rare word that occurs in 1% of the documents. Which word has a higher entropy?

✓ **Correct**

9. Let X be a binary random variable. Which of the following is **not** true? Select all that apply.   **0.75 / 1 point**

❗ **Incorrect**

10. True or false? An unbiased coin has a higher entropy than any biased coin.   **1 / 1 point**

✓ **Correct**