Week 2 Quiz

TOTAL POINTS 10

---

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 $\theta$is given in the table below:

1 point

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

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

- (•) 0.3
- ( ) 0
- ( ) 0.2
- ( ) 0.1

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

1 point

- ( ) $P(\text{“text mining”}|\theta) = 0$
- (•) $P(\text{“global warming”}|\theta) > P(\text{“warming global”}|\theta)$
- ( ) $P(\text{“the global warming effects”}|\theta) < P(\text{“global warming effects”}|\theta)$
- ( ) $P(\text{“global warming”}|\theta) = 0.04$

3.

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: $P(\text{“computer”}) =$

- ( ) 0.45
- ( ) 0.4
- (•) 0.25
- ( ) 0.05

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 point

Hint: Apply Bayes rule.

- ( ) 0.1 and 0.9
- ( ) 0.8 and 0.2
- (•) 0.2 and 0.8
- ( ) 0.9 and 0.1

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 point

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

- (selected) $P(w) = P(w|\theta_1)$, for any word w
- $P(w|\theta_1) = 0$, for any word w
- $P(w|\theta_2) = 0$, for any word w

6. 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 $H(X_{text}|X_{mining}) > H(X_{text}|X_{the})$.

1 point

- True
- (selected) False

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

1 point

- False
- (selected) True

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.

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?

1 point

- (selected) Both words have the same entropy.
- "the"
- "photon"

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

1 point

- [x] If P(X=1)=1, then H(X) = 1
- [ ] If P(X=0)=1, then H(X) = 0
- [ ] If P(X=0)=1, then H(X) = 1
- [ ] H(X) ≤ 1

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

1 point

- (•) True
- ( ) False

- [ ] I, **BAL KRISHNA NYAUPANE**, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.