# Week 2 Quiz

10 questions

Correct  
1 / 1 points

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:

|  |  |
| --- | --- |
| *w* | *P*(*w*|*θ*) |
| the | 0.3 |
| global | 0.2 |
| warming | 0.2 |
| effects | X |

What is X, i.e., *P*(“effects”|*θ*) ?

1. **0.3**
2. 0
3. 0.1
4. 0.2

Correct  
1 / 1 points

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

1. *P*(“text mining”|*θ*)=0
2. *P*(“global warming”|*θ*)=0.04
3. ***P*(“global warming”|*θ*)>*P*(“warming global”|*θ*)**
4. *P*(“the global warming effects”|*θ*)<*P*(“global warming effects”|*θ*)

1  
point

3. Assume that words are being generated by a mixture of two unigram language models, *θ*1and *θ*2, where *P*(*θ*1)=0.5 and *P*(*θ*2)=0.5. The distributions of the two models are given in the table below:

|  |  |  |
| --- | --- | --- |
| *w* | *P*(*w*|*θ*1) | *P*(*w*|*θ*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*(“computer”)=

1. **0.45**
2. **0.4**
3. **0.05**
4. 0.25

Correct  
1 / 1 points

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*(*θ*1|“computer") and *P*(*θ*2|“computer"), respectively, then the values of *P*(*θ*1|“computer”) and *P*(*θ*2|“computer") are:

*Hint: Apply Bayes rule.*

1. 0.8 and 0.2
2. 0.1 and 0.9
3. 0.9 and 0.1
4. **0.2 and 0.8**

Correct  
1 / 1 points

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.

If *P*(*θ*1)=1 then which of the following statements is true?

1. *P*(*w*|*θ*1)=0, for any word w
2. *P*(*w*|*θ*2)=0, for any word w
3. ***P*(*w*)=*P*(*w*|*θ*1), for any word w**

Correct  
1 / 1 points

6. True or false? Let *Xtext*, *Xmining*, and *Xthe* 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*(*Xtext*|*Xmining*)>*H*(*Xtext*|*Xthe*).

1. **False**
2. True

Correct  
1 / 1 points

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

1. **True**
2. False

Correct  
1 / 1 points

8. Let w be a word and *Xw* be a binary random variable that indicates whether w appears in a text document in the corpus. Assume that the probability *P*(*Xw*=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. "the"
2. “photon"
3. **Both words have the same entropy.**

Correct  
1 / 1 points

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

1. H(X) ≤ 1
2. If P(X=0)=1, then H(X) = 0
3. **If P(X=1)=1, then H(X) = 1**
4. If P(X=0)=1, then H(X) = 1

Correct  
1 / 1 points

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

1. **True**
2. False