

#### Unit 4 Unsupervised Learning (2

Course > weeks)

> Lecture 15. Generative Models >

3. Simple Multinomial Generative model

### **Audit Access Expires May 11, 2020**

You lose all access to this course, including your progress, on May 11, 2020.

# 3. Simple Multinomial Generative model

## Simple Multinomial Generative model



Video

Download video file

Transcripts

<u>Download SubRip (.srt) file</u>

<u>Download Text (.txt) file</u>

Note: For those who have taken 18.6501x (Fundamentals of Statistics): The concept of generative model introduced in the above video that models the probabilistic nature of data generation is the same as what we learnt as a statistical model in 18.6501x. With parameter  $\theta$ , the analagous notation that we saw for  $p\left(w\mid\theta\right)$  in the statistics course is  $\left(E,\left\{P_{\theta}\right\}_{\theta\in\Theta}\right)$ , where E is the sample space of the data and  $\left\{P_{\theta}\right\}_{\theta\in\Theta}$  is the family of distributions parameterized by  $\theta$ .

## Simple Multinomial Generative model

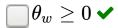
0/1 point (graded)

Consider a very simple multinomial model M to generate text in documents.

Let us assume that this model M has a fixed vocabulary W and that we generate a document by sampling one word at a time from this vocabulary. Furthermore, all the words that are generated by M are independent of each other.

We would like to capture the fact in our generative model M that some words in W are more likely to occur in any given document than the others. So, the first thing that M models is how likely it is to generate certain word  $w \in W$ . We denote this probability by  $P\left(w|\theta\right) = \theta_w$ , where  $\theta_w$  is a parameter in our model M.

Which of the following option(s) is/are true about the model parameters  $\theta_w$ ? Choose all that apply from the statements below:



$$lacksquare heta_w \geq 1$$

$$lefta \sum_{w \in W} heta_w = 1$$
 🗸

$$igsqcup \sum_{w \in W} heta_w > 1$$



#### **Solution:**

Note that  $\theta_w$  denotes the probability of model M choosing the word w. Since it's a probability, its value must lie between 0 and 1. Therefore,  $0 \le \theta_w \le 1$ .

Further, all the above probability values must also sum up to 1. That is,  $\sum_{w \in W} \theta_w = 1$ .

Submit

You have used 1 of 1 attempt

**1** Answers are displayed within the problem

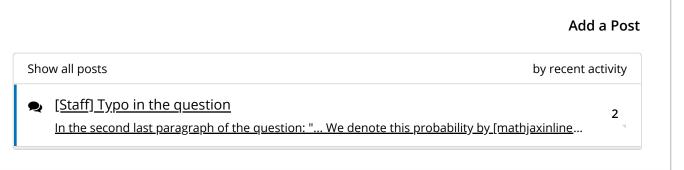
### Discussion

**Hide Discussion** 

**Topic:** Unit 4 Unsupervised Learning (2 weeks) :Lecture 15. Generative Models / 3. Simple Multinomial Generative model

3. Simple Multinomial Generative model | Lectur...

https://courses.edx.org/courses/course-v1:MITx+...



© All Rights Reserved