Course › Unit 4 Unsupervised Learning (2 weeks) › Lecture 15. Generative Models ›
5. Maximum Likelihood Estimate

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

5. Maximum Likelihood Estimate

Maximum Likelihood Estimate

0:00 / 0:00 1.25x

Video

Download video file

Transcripts

Download SubRip (.srt) file
Download Text (.txt) file

Number of Parameters

1/1 point (graded)

For the following set of questions, let us consider generating documents that are English letter sequences (assume no spaces or punctuation), i.e. the vocabulary $W = \{a, b, c. . . , z\}$ is made up of all the letters in the English alphabet.

We would like to generate documents using this vocabulary using a multinomial model M. As described in the lecture, what is the minimal number of parameters that the model M should have? Enter your answer below.

25

Answer: 25

Solution:

Recall from the lecture that for multinomial generative models we have a parameter θ_w for each word $w \in W$. However, since the parameters should sum up to one, we can express one of the parameters as 1 minus the sum of all others. Since the vocabulary size for this example is 26, our model M can have 25 parameters to express the probability of each letter.

Submit

You have used 1 of 2 attempts

Answers are displayed within the problem

Note: For those of you who have completed **18.6501x (Fundamentals of Statistics)**, it may be useful at this point to review or recall the lectures on Maximum Likelihood Estimation, which is perhaps the most powerful and general estimation method.

Maximum Likelihood Estimate

1/1 point (graded)

Let $\theta^* = \theta^*_a, \theta^*_b, . . . , \theta^*_z$ be the parameters of the multinomial model M^* that maximize the likelihood of generating a document D.

Further, it is known that the letter 'e' is twice as likely to occur as the letter 'z' in document D.

Which of the following options is a correct expression relating θ_e^* and θ_z^*?

- $\theta_z^* = 2\theta_e^*$
- $\theta_e^* = 2\theta_z^*$
- $\theta_z^* = \theta_e^*$
- $\theta_z^* + \theta_e^* = 2$

Solution:

Recall from the lecture that for any $w \in W$, the maximum

$$\theta_w^* = \frac{\text{count}(w)}{\sum_{w' \in W} count(w')}$$

Since $count(e) = 2count(z)$, we can conclude that $\theta_e^* = 2\theta_z^*$

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Maximum Likelihood Estimate for Poisson Distribution

2/2 points (graded)

Maximum Likelihood Estimate (MLE) is a very general method that can be applied to both continuous and discrete distributions. In this problem, we assume we have a training data $x_1, x_2 \dots, x_n$ that are drawn from a Poisson distribution, with probability mass function (pmf)

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}.$$

We want to use MLE to fit the parameter λ with the training data. To do so, we first compute the log likelihood of our training data, or in other words, log of the probability of obtaining the sample $x_1, x_2, \ldots, x_n$ given the model and where x_i are independent. The log likelihood is...

- ○ $\log \lambda \sum_i x_i + n\lambda + \sum_i \log(x_i!)$
- ◉ $\log \lambda \sum_i x_i - n\lambda - \sum_i \log(x_i!)$
- ○ $\log \lambda \prod_i x_i - n\lambda - \prod_i \log(x_i!)$
- ○ $\log \lambda \prod_i x_i + n\lambda + \prod_i \log(x_i!)$

✔

In the next step, we maximize this log likelihood function by taking the derivative. What is the resulting estimate for λ?

- ◉ $\frac{1}{n} \sum_i x_i$
- ○ $\frac{1}{n} \prod_i x_i$
- ○ $\sum_i x_i$
- ○ $\prod_i x_i$

✔

Is it in accordance with the definition of λ in Poisson distribution? (There is no answer box for this question.)

Solution:

The loglikelihood of the data is:

$$\begin{aligned} \log \prod_i P(X = x_i) &= \log \prod_i \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} \\ &= \sum_i \log(\lambda^{x_i}) + \log(e^{-\lambda}) - \log(x_i!) \\ &= \log \lambda \sum_i x_i - n\lambda - \sum_i \log(x_i!) \end{aligned}$$

Take the derivative with respect to λ, we have

$$\frac{1}{\lambda} \sum_i x_i - n = 0$$

$$\lambda = \frac{1}{n} \sum_i x_i.$$

This is in accordance with the fact that λ is the expectation of a Poisson variable with parameter λ.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 4 Unsupervised Learning (2 weeks) :Lecture 15. Generative Models / 5. Maximum Likelihood Estimate

Add a Post

Show all posts by recent activity

Fund of stat course

I haven't completed the stats course yet, is that going to be a hinderance moving forward? 5