

## Quiz Submissions - Q4-1



Parsa Yadollahi (username: [parsa.yadollahi@mail.mcgill.ca](mailto:parsa.yadollahi@mail.mcgill.ca))

### Attempt 1

Written: Feb 3, 2022 12:38 PM - Feb 3, 2022 12:48 PM

### Submission View

Released: Feb 3, 2022 4:00 PM

### Question 1

0.667 / 1 point

Select all the true statments: [scoring is right-wrong]

- ☐ a Bernoulli distribution used to model a binary variable has a single parameter which is the probability of the variable being 1
- ☐ in maximum likelihood estimation, we find the model parameters that maximizes multiplication of the likelihood of each data point given the model
- ☐ in Bayesian inference we are estimating the distribution of model parameters instead of finding the best parameter
- ☐ prediction in Bayesian inference can be hard since it needs averaging over all possible values of our parameter
- ☐ prediction in MAP is a single point estimate similar to MLE and doesn't need averaging over all possible values, hence is more efficient
- ☐ with a uniform prior, MAP and MLE estimates give the same predictions

View Feedback

### Question 2

0 / 1 point

Suppose you are given a die which has sides as 1 to 6 where the probability of each of the sides facing up is given by the categorical distribution is represented in the dictionary {1: 1/12, 2: 1/4, 3:1/6, 4:1/6, 5: 1/4, 6:1/12 }. You random roll the die for four times (note each of the rolls is iid), what is the (approximate) likelihood of observing this sequence of faces: (1, 3, 5, 5):

☐ 0.00064

☐ 0.0104

☐ 0.0123

➡ ☒ 0.00087

▶ View Feedback

### Question 3

1 / 1 point

Suppose you are interested in predicting the weather in Montreal tomorrow. During the last 30 days,  $N(\text{sunny}) = 5$  days,  $N(\text{cloudy}) = 10$  days, and  $N(\text{windy}) = 15$  days. What are the chances of it being windy or rainy tomorrow when using MLE with Laplace add-one smoothing? (this is similar to assuming pseudo-counts of  $[1,1,1,1]$  in posterior predictive)

☐ 0.25 , 0.1

✓ ☒ 0.471 , 0.029

☐ 0.5 , 0.030

▶ View Feedback

### Question 4

1 / 1 point

Due to the assumption of Naive Bayes Classifier, the term

$$p(x_2|y, x_1)$$

becomes

☐

$$p(x_2)$$

☐

$$p(x_2|y) \times p(x_1)$$

✓ ☒

$$p(x_2|y)$$

☐

$$p(x_2) \times p(x_1)$$

 [View Feedback](#)

---

**Attempt Score:** 3 / 4 - 66.68 %

**Overall Grade (highest attempt):** 3 / 4 - 66.68 %

Done