# Quiz Submissions - Quiz 3

**Mike Gao (username: zenghao.gao@mail.mcgill.ca)**

**Attempt 1**

Written: Sep 28, 2020 8:22 AM - Sep 28, 2020 4:34 PM

**Submission View**

Released: Sep 27, 2020 3:21 PM

## Question 1

Suppose x~N(0, 2), y = 2x, z = 3y. Select all correct options.

- ✓ ☐ Corr(x,y) < Corr(x, z)
- ✓ ☑ Corr(-x, y) < Corr(x, y)
- ✓ ☐ Corr(y, z) > Corr(x, y)

## Question 2

Suppose you have a multi-class classification dataset with 4 possible classes and 30 binary features. How many parameters do you need to learn to fit a Naive Bayes model to this dataset?

- ✓ ◉ 123
- ○ 61
- ○ 125
- ○ 122

▼ Hide Feedback

We need to learn the class distribution P(y) and the conditional distribution of the features given each class. Since there are 4 classes, P(y) requires 3 parameters to learn, as specifying the marginal probability of three of classes determines the probability of the fourth class. (Note, for instance, that in the binary classification case we only needed 1 feature to represent P(y) when there was two classes).

In addition, to estimate P(x|y=k) for each class, we need 30 parameters. Thus, in total, there are 3+4*30=123 parameters to learn.

## Question 3

X is a random normal variable, with mean μ and variance σ^2. The “standardised form” of X is Z = (X−μ)/ σ .

What are the mean and variance, respectively, of Z ?

- ✔ (●) 0,1
- ( ) 2, 1
- ( ) 1, 0
- ( ) 2, 0

## Question 4

If we train a Naive Bayes classifier using infinite training data that satisfies all of its modeling assumptions (e.g., conditional independence), then in general, what can we say about the training error (error in training data) and test error (error in held-out test data)?

- ✔ (●) It may not achieve either zero training error or zero test error
- ( ) It will always achieve zero training error and zero test error
- ( ) It will always achieve zero training error but may not achieve zero test error
- ( ) It may not achieve zero training error but will always achieve zero test error

## Question 5

Select all correct statements

- ✔ [✔] If p(x| y, z) = p(x|y), then x is independent of z given y
- ✔ [✔] If p(x, z|y) = p(x|y) p(z|y) then x is independent of z given y
- ✔ [ ] If x and y are dependent then they have non-zero correlation
- ✔ [✔] Naive Bayes is a generative classifier

## Question 6

An employee at a movie production company is prototyping a Naive Bayes model to predict whether a movie will be successful (a binary classification task). So far in the prototype there are three binary features:

- fresh, which is 1 if the movie is “certified fresh” on Rotten Tomatoes and 0 otherwise.
- summer, which is 1 if the movie was released in the summer and 0 otherwise.
- rock, which is 1 if the movie is starring Dwayne “The Rock” Johnson and 0 otherwise.

Suppose the model is trained with Laplace add-one smoothing on the following data:

- success=1, [fresh=0, summer=0, rock=1]
- success=1, [fresh=0, summer=0, rock=0]
- success=1, [fresh=1, summer=1, rock=1]
- success=0, [fresh=0, summer=1, rock=0]
- success=0, [fresh=1, summer=0, rock=0]

Would this model predict success or failure for a movie with the following attributes: [fresh=0, summer=0, rock=1]

- ✓ (●) Success
- ( ) Failure
- ( ) Impossible to tell (i.e. not enough information given)

▼ Hide Feedback

---

Using the notation from lecture, the maximum likelihood parameters for this model are:

$$\theta_1 = 3/5,\ \theta_{1,fresh} = 2/5,\ \theta_{1,summer} = 2/5,\ \theta_{1,rock} = 3/5,\ \theta_0 = 2/5,$$
$$\theta_{0,fresh} = 2/4,\ \theta_{0,summer} = 2/4,\ \theta_{0,rock} = 1/4$$

And from these we can get that

$$P(success = 1|[fresh = 0, summer = 0, rock = 1])\theta_1(1 - \theta_{1,fresh})(1 - \theta_{1,summer})\theta_{1,rock} = 0.129$$

and that

$$P(success = 0|[fresh = 0, summer = 0, rock = 1]) \propto \theta_0(1 - \theta_{0,fresh})(1 - \theta_{0,summer})\theta_{0,rock} = 0.112$$

---

## Question 7

What is the number of parameters needed to represent a Naive Bayes classifier with n Boolean variables and a Boolean label ?

- ✔ 2n+1
- 2n
- n
- n+1

**Attempt Score:** 7 / 7 - 100 %

**Overall Grade** (highest attempt): 7 / 7 - 100 %

Done