Group Homework 2 Stat 100 Spring 2022

5/17/2022

100 Possible Points

Attempt 1

Next Up: Review Feedback

Offline Score: N/A

Add Comment

Unlimited Attempts Allowed

5/8/2022 to 5/17/2022

Details

Group Homework 2 Stat 100 Spring 2022

This second homework is again focused on the use of conditional, joint, total probability and Bayes theorem. But now, we are going to investigate how those are used when conducting statistical data analysis and using probability models.

Do not hesitate to contact Dr. Sanchez, if some of the instructions below are not clear. Use Piazza to ask your questions, so that other students in class will answer them and get points.

Reading section with embedded questions the whole group must answer.

We learn in weeks 6 and 7 how to model distributions of two random variables and distributions derived from them. In doing so, we repeat the study of joint probability, total probability, conditional probability, but now with distributions for random variables. The question then is, how do the formulas that we learned in Chapter 3, translate into formulas with distributions? We see them in lectures. But I summarize them here and go a little beyond.

Multiplication rule (in general)

$f(x,y) = f(x)f(y|X=x)$ we saw in week 6 that if x,y, independent f(x,y)=f(x)f(y)

Conditional formula

Submit Assignment

(https://bruinlearn.ucla.edu/courses/133166/modules/items/4815548)

Bayes' Rule

$$f(x|y) = \frac{f(y|X=x)f(x)}{f(y)}$$

In Bayes formula, f(x) is called the prior probability density function of X (unconditional pdf). And f(x|y) is called the posterior probability density function of X (posterior because it depends on what y is, after we know what y is), a conditional density function.

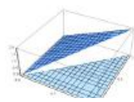
Similar formulas apply to discrete probability mass functions. We just have to replace the f for P (big P).

Reading question 1.

(a) We study the following bivariate density function in Lecture 23:

Gasoline is to be stocked in a bulk tank once each week and then sold to customers. Let X_1 denote the proportion of the tank that is stocked in a particular week, and let X_2 denote the proportion of the tank that is sold in th same week. Due to limited supplies, X_1 is not fixed in advance but varies from week to week. Suppose that a study of many weeks shows the joint relative frequency behavior of X_1 and X_2 to be such that the following join density function provides and adequate model:

$$f(x_1, x_2) = \begin{cases} 3x_1 & 0 \le x_2 \le x_1 \le 1 \\ 0 & \text{elsewhere} \end{cases}$$



Calculate all the density functions needed and plug them in the formula for Bayes theorem given above. Confirm that Bayes theorem formula holds. Show work for the ones not shown in the lecture, and use the formulas given in the lecture if they have been already calculated there.

(b) We study in lecture 24 the bivariate Gaussian distribution. We would like to find out what distribution the posterior density function of x given y is. Calculate all the density functions needed and plug them in the formula for Bayes theorem given above. Confirm that Bayes theorem formula holds. You may use formulas given already in the lecture. Make sure you do the necessary simplifications in order to identify the model for the posterior density.

Bayes rule when we have data

When used with data, to learn about populations about which we do not know the distribution parameters, the terms in Bayes formula receive special names and play a very specific role. We will explain with an example in sports.

Note: little p is a parameter usually, big P is used to denote probability of a random variable, or an event. You must keep that distinction. Do not use little p when you are saying, for example, p(A).

Let's do an example:

<

[(https://bruinlearn.ucla.edu/courses/133166/modules/items/4815548)](https://bruinlearn.ucla.edu/courses/133166/modules/items/4815548)

Submit Assignment

the proportion of the games won by A in the year. I summarize what we know from past years in columns 1 and 2 below. Model means the proportion of games won by A in a year. Prior means knowledge of proportion of past years where A won that proportion of games.

Model	Prior	Likelihood	Prior × Likelihood	Posterior	Model × Posterior
0	0	0	0	0	0
.1	.02	.081	.00162	162/8,616 = .02	162/86,160
.2	.03	.128	.00384	384/8,616 = .04	768/86,160
.3	.05	$\boxed{.147}$	.00735	735/8,616 = .09	2,205/86,160
.4	.10	.144	.01440	1,440/8,616 = .17	5,760/86,160
.5	.15	.125	.01875	1,875/8,616 = .22	9,375/86,160
.6	.20	$\boxed{.096}$	.01920	1,920/8,616 = .22	11,520/86,160
.7	.25	.063	.01575	1,575/8,616 = .18	11,025/86,160
.8	.15	.032	.00480	480/8,616 = .06	3,840/86,160
.9	.05	.009	.00045	45/8,616 = .01	405/86,160
1	0	0	0	0	0
Sum	1.00		.08616	1.00	45,060/86,160

Those two first columns are saying, for example,that in 2% of the past years player A won 10 percent of the games against B; in 25% of the years, player A won 70% of the games; and so on. Those are my prior beliefs, that is why I call them prior probability. **The first column plays the role of our variable x in the Bayes formula given earlier. The second column is P(x)**. The P(x) given in the table is the population model, what we know without further information. It is what we would use to predict the probability of A's proportion of wins next year if we did not know anything. But what if we know more information?

? Reading section question 2. Based solely on the prior probabilities for the models -the past history(looking at columns 1 and 2), what is the probability that player A is better than player B? Which of the players would you bet on? Explain why referring to the table.

Suppose that the new year starts, and I have observed now in the current year that of the **three** times that A,B played. That is additional information, data. I no longer have just the information from last year, the population model. A won the first game and lost the second and third. This is data, y. We could calculate what is called the *likelihood* of the model given that data. *The likelihood of a model is the probability of the data calculated assuming that model.* This function is occupying the place of f(y|x) in our Bayes formula above, but when it is likelihood of just one specific observed data point, it is not really a conditional probability, it will not have probabilities add up to 1 (adding up likelihoods over all possible data will add up to 1, though). Since we have observed only that the first was won by A, we have observed only one possible value of the likelihood. For example, in column 3, in the row where the model is p=0.6, the likelihood is 0.6(0.4)(0.4) = 0.096, using the product rule for independent events. If the model is p=0.3, the likelihood of

(https://bruinlearn.ucla.edu/courses/133166/modules/items/4815548)

Submit Assignment

With columns 1,2,3 figured out, we are now ready to calculate the numerator of the Bayes formula given earlier, what plays the role of f(y| X=x)f(x). The numerator under the different models is given in column 4. So we multiply columns 3 by column 2 to obtain column 4. For example, when p=0.3, the prior probability 0.05 times the likelihood is 0.147(0.05)= 0.00735. These values are given in column 4 of the table.

We still have to calculate the P(y), the probability of the data, in the denominator of Bayes formula. According to the *law of total probability*, this will be:

P(y) =P(y|model=0)P(model=0)+ P(y|model=0.1)P(model=0.1) + P(y|model=0.2)P(model=0.2)+...........+P(y|model 0.9)P(model=0.9)+P(x|model=1)P(model=1)= 0.081(0.02)+ 0.128(0.03)+.............+0.009(0.05)=0.08616.

The individual values of the numerator of Bayes formula for each prior model are given in column 4. At the bottom of column 4, you have the value of the P(y) just calculated.

We have now all the information needed to calculate the posterior probability of the models considered. Applying Bayes formula, column 5 calculates the posterior

Posterior probability of a model = $\frac{Likelihood \times Prior}{\texttt{Total probability of the data}}$

Now that we have updated the probabilities of various model proportions, we can calculate the posterior probabilities of various events. For example, the posterior probability, given the data observed (that player A won the first of 3 games), that player A is better than B is (1,920+1,575 +480+45)/8,616 = 0.4665738 or approximately 47%.

? Reading section question 3. Based on the posterior probabilities, given the observed data, what is the probability that A is better than B? Which player would you bet on?

What is the posterior probability that A and B are equally good?

After seeing the posterior distribution, did you change the bet you suggested in question 1? Explain why using the information in the table and your answers to reading questions 1 and 2.

? Reading section question 4.

- Plot the prior probability against model in one graph, using only lines to represent the probabilities (See the graph in the title slide of Lecture 16, week 5; you must use that type of graph, not bar graphs). Plot the Likelihood against model in a separate graph right below the first graph, and plot the posterior probability right below the other two, also against the

(https://bruinlearn.ucla.edu/courses/133166/modules/items/4815548)

Submit Assignment

Use instead a line graph like the one on the front of the pre-recorded lecture on the negative binomial.

- Compare the posterior with the prior distribution plots. Has the data had any effect on the prior distribution? Explain why you answer the way you answer referring to your graphs.

Reading section question 5. Generalization to when we have density models to represent the information.

Suppose that X can be modeled by a *Beta distribution* and suppose that Y follows a Binomial model P(Y=y |X=x, n=3). What would be the formula for the posterior distribution of X given Y? To answer, substitute the formulas for these models in the formula for Bayes theorem given above, in the numerator. The Beta is a density function, and the Binomial is a pmf. So notice that we can mix distributions in the Bayes formula too. By looking at the formula for the posterior, and comparing with the formulas used, can you identify the density model for the posterior?

Use the Wikipedia parameterization of the Beta density function (at the top of Wikipedia's page after you google Beta distribution. Observe in Wikipedia's plots what the density looks like.

Reading section question 6.

Suppose that instead of having the data "A wins the first game" we had had the data "A wins 1 game of the three games already played this current year." That is very different information than the one we used to construct the table provided. Modify the calculations in the table, and write the whole table again, with the new information for all columns that will need change. Compare the new prior and posterior probabilities and determine which player is better now, A or B, based on this information. How do these new calculations change your answers to questions 3 and 4?

<

(https://bruinlearn.ucla.edu/courses/133166/modules/items/4815548)

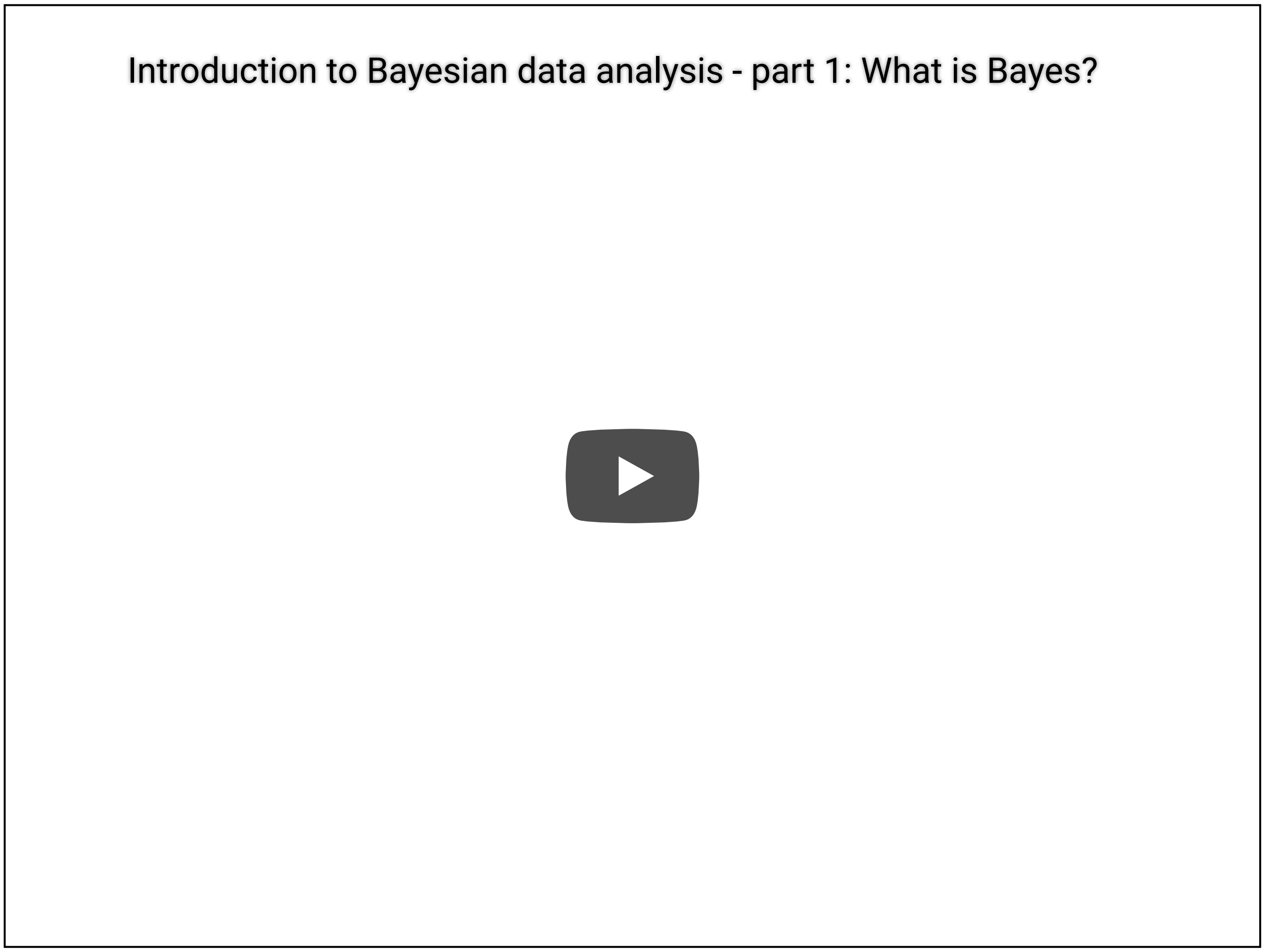
Submit Assignment

Watch the following videos. The first one is a little historical, and very short. The second shows an example different from the one we have described in the Reading section.

(http://Understand%20how%20Bayes's%20theorem%20can%20make%20educated%20mathematical%

https://www.britannica.com/video/193403/Bayes-theorem-guesses (https://www.britannica.com/video/193403/Bayes-theorem-guesses)

Watch also this video introducing Bayesian data analysis



(http://Understand%20how%20Bayes's%20theorem%20c

(https://www.britannica.com/video/193403/Bayes-theorem-guesses)

(https://bruinlearn.ucla.edu/courses/133166/modules/items/4815548)

Submit Assignment

In Module Week 4 lesson 11B, we fitted a Log normal density model to the radon data. We said there that we used a method called maximum likelihood, which is a method used by statisticians to estimate parameters of probability models from data. The function calculated is the likelihood function, which is the joint distribution of the data observed given the parameters, the f(y|x)-equivalent seen in the Bayes formula given at the beginning of the reading section in this homework. That is obtained by doing the calculation we did in week 6, the video supplementing lecture 21, but looking at the joint distribution as a function of the parameters.

(a) After you watch the video Introduction to Bayesian Data Analysis Part 1, posted right above, compare the two ways of estimating the parameters of a distribution, and indicate what else would we need in the estimation of the log normal distribution for the radon data in order to do the estimation of the parameters in a Bayesian one. Refer specifically to the examples seen in lecture 11B and the "Introduction to Bayesian Data Analysis part 1 video" and the example of the two players in the reading.

(b) The video before this "Introduction to Bayesian Data Analysis video" posted above mentions several applications of Bayesian data analysis that are crucial to today's progress in discovery in the world. Give those examples mentioned in the video. As you watch that video you probably are also thinking about the slide in the video posted in our "Getting Ready" module, the video titled "Why is probability so important in machine learning?" There, the authors mention conditional probability. When the author says in this last video that machine learning uses conditional probability, which conditional probabilities is it referring to in the Bayes formula ? Justify your answer by providing examples of applications of that conditional probability mentioned in the video.

(c) Give a couple of generative model examples given in the video posted above (Introduction to Bayesian Data Analysis). What is the difference between a generative model and the models we are studying in this class (probability mass functions and density models).

Additional Activities

Additional activities question 1.

<

(https://bruinlearn.ucla.edu/courses/133166/modules/items/4815548)

Submit Assignment

order to perform a Bayesian data analysis. In your description, be specific as to where you would look for that data and what that data might look like. Specific examples preferred.

(b) Gathertown has a game going on now, Escape the Island, that can be played with your team. I think you have to login with a guest account in order to access it. **https://app.gather.town/app** **(https://app.gather.town/app)**

- I am writing here the link to access it. **https://app.gather.town/invite?token=dt0hGXoMQB58dTkeAFvxFsuKXTGj5Uf1** **(https://app.gather.town/invite?token=dt0hGXoMQB58dTkeAFvxFsuKXTGj5Uf1)** Preferably 4 people should play, so coordinate with your team to play it.
- After playing it, do you think Bayesian statistics could be used to predict the outcome of this game? Why or why not? You may want to introduce how the game goes first, because the reader needs to know.

Assignment and what to submit

- Create a document where you will copy paste the questions that are posed in all the sections above (**Q 1-6 in Reading Section, Question Lecture Video Question 1 and Additional Activities question 1**) and answer those questions, in the same order in which they appear. The questions should be discussed with all members of the group, because solving one helps solve the others better than if you solve them separately.

Submission of group work (before 5/17/2022, 11:55 PM)

- Submit **one pdf file per group** (one of you submits it). All members of the group are responsible, so you must make sure that your designated person for submission has submitted it.
- The file name should be your three last names .pdf. For example, smith-lee-bastani.pdf, for students with those last names.
- Your three full names must be written at the beginning of your document.
- As indicated in the instructions the question must be copy-pasted before the answer, and the questions and answers must appear in the order in which they show in this homework.
- No more than 10 pages must be submitted.
- Notice that it is possible that upon reading it, you might encounter a typo or something needs clarification. Please, email me so that I can share the correction with the class.

<

(https://bruinlearn.ucla.edu/courses/133166/modules/items/4815548)

Submit Assignment

View Rubric

(https://bruinlearn.ucla.edu/courses/133166/modules/items/4815548)

Submit Assignment

Some Rubric (1)		
Criteria	**Ratings**	**Pts**
Each of the 4 resources mentioned is summarized accurately, and briefly, with the main points well understood and explained (10 points each) view longer description		/ 40 pts
Description of criterion view longer description		/ 40 pts
The file contains the names of the three students in the group at the very beginning of the file (5 pts) view longer description		/ 5 pts
The file is neatly typed and the parts are separated and sources labeled with (a), (b), ..etc.), no ambiguous language is used that hides lack of undestanding (10 pts) view longer description		/ 10 pts
The file is pdf and name of three		

Submit Assignment

(https://bruinlearn.ucla.edu/courses/133166/modules/items/4815548)

Some Rubric (1)

Criteria	Ratings	Pts
Indication of plagiarism of other groups' work (0 points and a trip to the Dean of students' office) view longer description		/ 0 pts
		Total Points: 0

Keep in mind, this submission will count for everyone in your 100A Homework 2 Spring 2022 group.

Choose a submission type

Upload

More

Canvas Files

or

(https://bruinlearn.ucla.edu/courses/133166/modules/items/4815548)

Submit Assignment

I agree to the tool's **End-User License Agreement (https://api.turnitin.com/api/lti/1p0/user/static_eula)**
This assignment submission is my own, original work

(https://bruinlearn.ucla.edu/courses/133166/modules/items/4815548)

Submit Assignment