Homework 1

An Bui

2023-01-16

Bayesian Statistical Modeling Winter 2023

Homework, Week 1

When is homework due? Homework is due on Mondays (Tuesday if Monday is a holiday). This assignment is due on Monday, 1/16/2023. Submit homework to gradescope.

Before submitting your work You should use this file as a template for completing the assignment. You should change the author to your name. You should use the features of Rmarkdown to write your text answers below and include R code and equations where necessary. You should turn in both the .Rmd file as well as the pdf file that you generate by using the `Knit` button.

Running the shiny app

This assignment is a walkthough of using the virutal machine to run a shiny app. Once you have your VM up and running you can navigate the filesystem and find the file app.R under "Distributed Class Material". This app performs the Bayesian calculation we discussed in class, namely what is the probability of having covid given cold-like symptoms.

Question 1: Baye's rule

The calculation uses Bayes rule and expresses relationships between conditional probabilities.

1A: What is the conditional probability that we are trying to calculate? Write this in words and symbols, as best you can.

The conditional probability we are trying to calculate is the **probability that an individual has covid given that they have cold-like symptoms (e.g. a runny nose)**. In symbols, this would be: $\Pr(\mathrm{C}|\mathrm{R_n})$ where C represents covid and $\mathrm{R_n}$ represents cold-like symptoms, or a runny nose.

1B: I gave two versions of the calculation in class. The simpler version assumed that you could not have both a cold and covid, while the more complicated one assumed you could have both. Write out one of these as best you can.

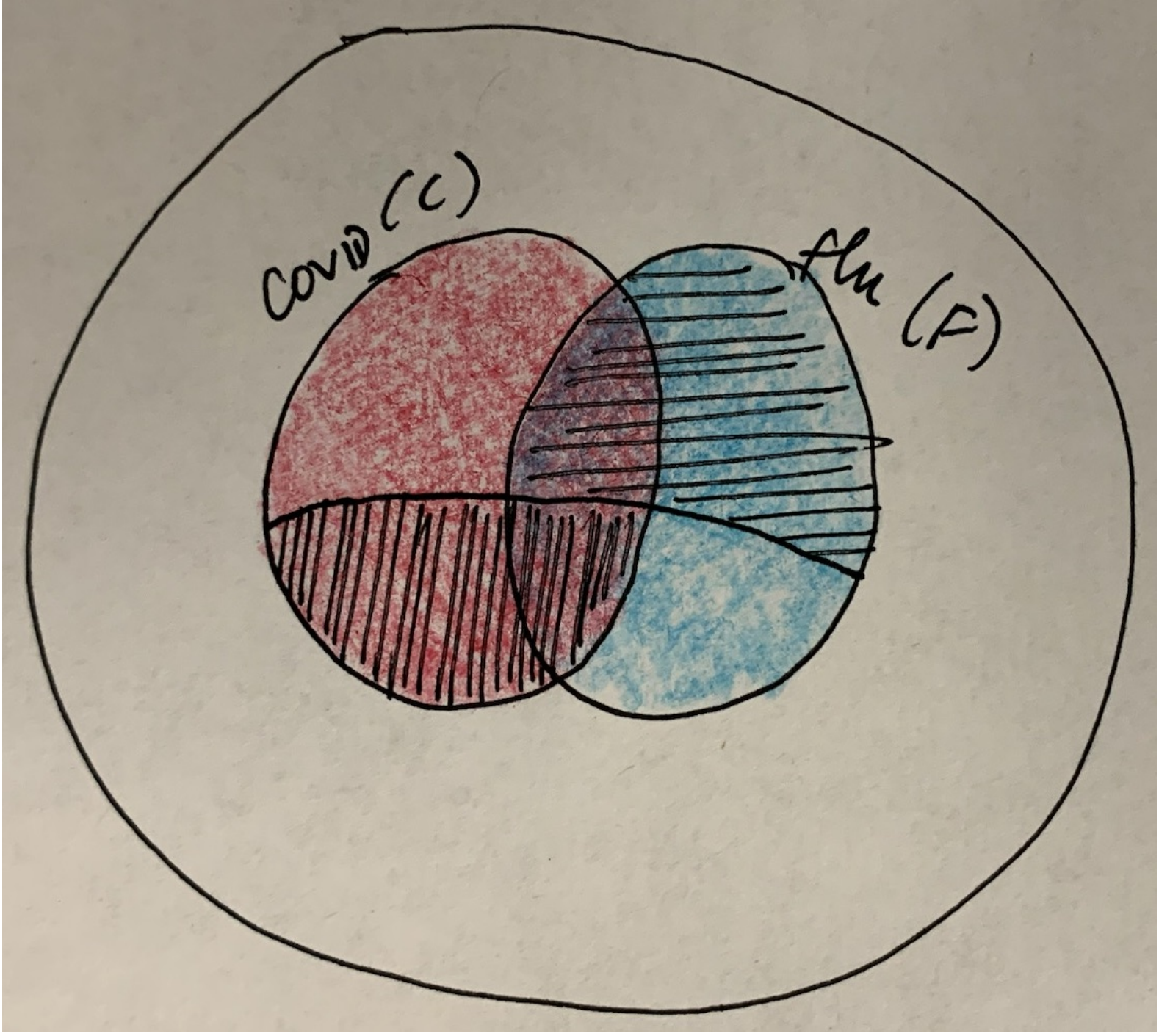
Simple version (cannot have both a cold and covid):

$$\Pr(\mathrm{C}|\mathrm{R_n}) = \frac{\Pr(\mathrm{R_n}|\mathrm{C}) * \Pr(\mathrm{C})}{\Pr(\mathrm{R_n}|\mathrm{C}) * \Pr(\mathrm{C}) + \Pr(\mathrm{R_n}|\mathrm{F}) * \Pr(\mathrm{F})}$$

where F represents a cold (or a flu).

Complicated version (can have cold *and* covid):

In this scenario, 1) you can have a runny nose ($\mathrm{R_n}$) from either covid (C) or a cold (F), but not both, and 2) if you have both, then you have to have a runny nose.



conceptual diagram

$$\Pr(C|R_n) = \frac{\Pr(R_n|C) * \Pr(C)}{\begin{array}{l}[\Pr(R_n|C) * \Pr(C) - \Pr(R_n|C) * \Pr(F) * \Pr(C)] + \\ [\Pr(R_n|C) * \Pr(C) * \Pr(F)] + [\Pr(R_n|F) * \Pr(F) * \Pr(C)] + \\ [\Pr(R_n|F) * \Pr(F) - \Pr(R_n|F) * \Pr(F) * \Pr(C)]\end{array}}$$

Question 2: Bayesian updated covid odds

2A: Low covid prevalence

Set the slider to a low value for covid prevalence. Report the value you chose, and describe the graph and explain why you beleive it has the shape and values.

For low covid prevalence, I chose **0.05**. The graph has a line that falls steeply as the probability of having a cold increases. At $\Pr(\text{cold}) = 1, \Pr(C|\text{symptoms}) = 0.05$. The steep slope at the beginning is due to the fact that covid prevalence is low in the population, so it is not very likely that a person has covid if they have symptoms; rather, it is likely they have a cold. The

line reaches 0.05 at $\Pr(\text{cold}) = 1$ because there is still a probability that an individual has covid *and* a cold, equal to the covid prevalence in the population; however, it is more likely that they have a cold rather than covid.

2B: High covid prevalence

Set the slider to a high value for covid prevalence. Report the value you chose, and describe the graph and explain why you beleive it has the shape and values. How is it different from the low value?

For high covid prevalence, I chose **0.85**. The graph has a line with a shallow slope (rather than a steep slope with low covid prevalence), and at $\Pr(\text{cold}) = 1, \Pr(\text{C}|\text{symptoms}) = 0.85$ (rather than 0.05, with low covid prevalence). The shallow slope comes from the fact that, because covid prevalence is high in the population, an individual is likely to have covid if they have symptoms. $\Pr(\text{C}|\text{symptoms}) = 0.85$ at $\Pr(\text{cold}) = 1$ because, similarly to the low covid prevalence example, there is still a probability that an individual has covid and a cold; therefore, the probability of having covid if presenting symptoms is equal to the prevalence in the population.