ST540 HW3

8. We first calculate $P(Y_1 = 1)$ and $P(Y_2 = 1)$:

$$P(Y_1 = 1) = P(X_1 = 1 \wedge X_2 = 1) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6} \tag{1}$$

$$P(Y_2 = 1) = P(X_1 = 1 \vee X_2 = 1) = 1 - P(X_1 = 0 \wedge X_2 = 0) = 1 - \frac{1}{2} \cdot \frac{2}{3} = \frac{2}{3} \tag{2}$$

   (a) $P(X_1 = 1 \wedge X_2 = 1 | Y_1 = 1) = P(X_1 = 1 \wedge X_2 = 1 | X_1 = 1 \wedge X_2 = 1) = 1$

   (b)

$$\begin{aligned} P(X_1 = 1 \wedge X_2 = 1 | Y_2 = 1) &= \frac{P(Y_2 = 1 | X_1 = 1 \wedge X_2 = 1) P(X_1 = 1 \wedge X_2 = 1)}{P(Y_2 = 1)} \\ &= \frac{1 \cdot (1/6)}{2/3} = \frac{1}{4} \end{aligned} \tag{3}$$

   (c) $P(X_1 = 1 | Y_1 = 1) = P(X_1 = 1 | X_1 = 1 \wedge X_2 = 1) = 1$

   (d)

$$P(X_1 = 1 | Y_2 = 1) = \frac{P(Y_2 = 1 | X_1 = 1) P(X_1 = 1)}{P(Y_2 = 1)} = \frac{1 \cdot (1/2)}{2/3} = \frac{3}{4} \tag{4}$$

9. (a) Informative priors should be used if they can be supported by pre-existing evidence, such as common sense or previous research. In this case, we have a sample of 65 papers with known authorship. If this sample is a a good representation of the population(all the papers written by Hamilton or Madison), 51/(14+51) and 14/(14+51) would be reasonable priors. However, the sample can be biased. For example, the population may actually consists of 100 papers by Hamilton and 500 papers by Madison. If we have no evidence that the sample is unbiased, we should use non-informative priors.

   (b) Priors: $P(Author = H) = P(Author = M) = 0.5$
   The likelihood can be modeled using Binomial or Poisson distribution:
   $Y | Author = H \sim Bin(1000, 3.24/1000)$, $Y | Author = M \sim Bin(1000, 0.23/1000)$,
   or
   $Y | Author = H \sim Poisson(3.24)$, $Y | Author = M \sim Poisson(0.23)$,
   where Y is the frequency of 'upon' in a disputed text of length 1,000 words.

$$\begin{aligned} P(Author = H | Y = 3) &= \frac{P(Y = 3 | Author = H) P(Author = H)}{P(Y = 3)} \\ &= \frac{P(Y = 3 | Author = H) P(Author = H)}{P(Y = 3 | Author = H) P(Author = H) + P(Y = 3 | Author = M) P(Author = M)} \\ &= \frac{0.5 P(Y = 3 | Author = H)}{0.5 P(Y = 3 | Author = H) + 0.5 P(Y = 3 | Author = M)} \end{aligned} \tag{5}$$

   Using Binomial likelihood, $P(Y = 3 | Author = H) = \binom{1000}{3} (\frac{3.24}{1000})^3 (1 - \frac{3.24}{1000})^{1000-3} = 0.2223$. Similarly, $P(Y = 3 | Author = M) = 0.0016$. Thus, $P(Author = H | Y = 3) = 0.9928$.
   Using Poisson likelihood, $P(Y = 3 | Author = H) = 3.24^3 e^{-3.24}/3! = 0.2220$. Similarly, $P(Y = 3 | Author = M) = 0.0016$. Thus, $P(Author = H | Y = 3) = 0.9928$.
   It turns out that the result is very similar whether we use binomial or Poisson distribution. It's because $Bin(n, p)$ approaches $Poisson(\lambda)$ if $n \to \infty$, $p \to 0$ and $np \to \lambda$.

   (c) For the binomial likelihood, we actually assume that each word in the text has equal probability of being 'upon', and the words are independent of each other. These assumption are not likely to be true. For example, given that there is already an 'upon' in a sentence, it's not likely to have another 'upon' in the same sentence.

(d) The posterior probability in (b) will decrease, because $Y = 1$ is closer to Madison's expected rate.

(e) The posterior probability in (b) will increase. While the sample proportion of 'upon' stays the same ($3/1000 = 30/10000$), a larger sample size leads to smaller variance.