Tutorial COMP1801: Introduction to Bayesian Thinking

1. RT-PCR essays in the UK have analytical sensitivity and specificity of 95%. Sensitivity is the percentage of people with COVID who (correctly) tested positive; while the specificity represents the percentage of people without COVID that correctly tested negative. By the most recent estimation from the Office of National Statistics (6th of November, 2020), 1 in 90 people within the population in England has COVID.

   (a) Assume that you are tested positive, what is the probability that you really have COVID?

   (b) To confirm your diagnosis you decide to take another test. If the test comes positive again, what will be the updated the probability that you have COVID? Use the probability calculated in part a) as the new prior probability of COVID.

   Solution.

   (a) The information given is:

   - $P(+|COVID) = 0.95$, therefore $P(-|COVID) = 0.05$
   - $P(-|NO\,COVID) = 0.95$, therefore $P(+|NO\,COVID) = 0.05$
   - $P(COVID) = 1/90$, therefore $P(NOCOVID) = 1 - \frac{1}{90}$

   We use Baye's Theorem to compute $P(COVID|+)$:

$$\begin{aligned} P(COVID|+) &= \frac{P(+|COVID)P(COVID)}{P(+|COVID)P(COVID) + P(+|NO\,COVID)P(NO\,COVID)} \\ &= \frac{0.95(1/90)}{0.95(1/90) + 0.05\left(1 - \frac{1}{90}\right)} \\ &= 0.1759 \end{aligned}$$

   The probability that you have COVID given that the test was positive is only 17.59%.

   (b) We use the Baye's Theorem but now the prior probability that you have COVID is 17.59%

$$\begin{aligned} P(COVID|+) &= \frac{P(+|COVID)P(COVID)}{P(+|COVID)P(COVID) + P(+|NO\,COVID)P(NO\,COVID)} \\ &= \frac{0.95(0.1759)}{0.95(0.1759) + 0.05\,(1 - 0.1759)} \\ &= 0.8021 \end{aligned}$$

   The probability that you have COVID after your second testj positive is 80.21%.

2. Assume a die that has probability of showing the number 6 equal to 25%. All other outcomes have equal probability to happen.

   (a) Let Y be a random variable that represents the outcome of rolling this die once. Write in a table the probability mass function of Y.

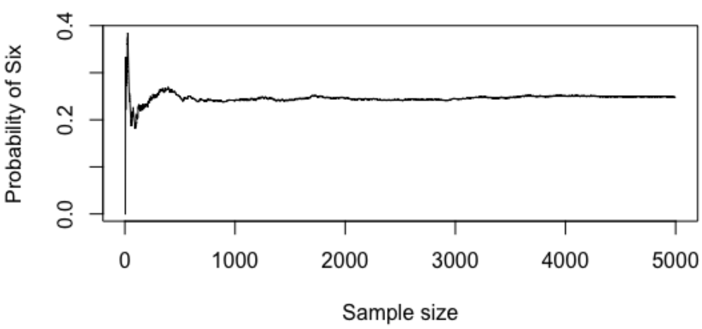
   (b) Simulate in R rolling the die $n = 5000$ times. Plot the proportion of 'sixes" as a function of the number of rolls. Does the proportion of "sixes" converge to the true probablity (25%)?

   Solution.

   (a) We are told that $P(Y = 6) = 0.25$ and that all the other outcomes have same probability to happen. Then, as $\sum_1^6 P(Y = i) = 1$, we know that the probability of not getting a number six is $1 - 0.25 = 0.75$. This, divided by the rest number of outcomes give us $0.75/5 = 0.15$.

y_i	1	2	3	4	5	6
$P(Y = y_i)$	0.15	0.15	0.15	0.15	0.15	0.25

   (b)
```
N = 5000 # sample size
sample = rbinom(N,1, 0.25) # simulation of coin tosses (1=Heads,
0=Tails)
index = 1:N
cum = cumsum(sample)
cum_mean = cum/index # proportion of Heads
plot(index[1:N], cum_mean[1:N], type="l", xlab="Iteration",  ylab="Probabili
 of Six")
```



3. Stanley Milgram investigated the propensity of people to obey orders from authority figures, even when those orders may harm other people [1]. The participants in the study were given the task of testing another participant (who was in truth a trained actor) on their ability to memorize facts. If the actor didn't remember a fact, the participant was ordered to administer a shock on the actor and to increase

[1] Milgram S. (1963) "Behavioral Study of Obedience", *The Journal of Abnormal and Social Psychology*, 67: 371–78.

the shock level with every subsequent failure. Unbeknownst to the participant, the shocks were fake and the actor was only **pretending** to register pain from the shock. Shockingly, among the 40 participants in Milgram's study, 26 (65%) administered what they thought to be the maximum shock to the actor). Our aim is to translate this study to the Bayesian framework [2].

(a) What is the parameter of interest?

(b) Assume that psychologists -experts on the field- propose a Beta$(1, 10)$ prior distribution for the parameter of interest. Plot the prior distribution. What does the prior model reveal about the psychologist's prior understanding of the parameter of interest?

(c) Using all the information given, write the appropriate Beta-Binomial model.

(d) Review the notes from the lecture to answer what is the posterior distribution of the parameter of interest? Specify the distribution and its updated parameter values.

(e) Plot the likelihood, prior and posterior distributions in R.

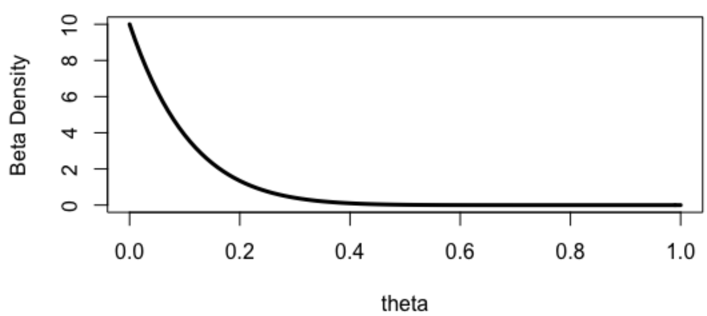
(f) Compare the mean of the prior and the posterior distribution to understand how the psychologists have updated their beliefs after the observation of the empirical data.

Solution.

(a) The parameter of interest is the proportion of people who will obey authorities even when those orders may harm other people, p.

(b) $p \sim B(1, 10)$

```
p = seq(0,1,0.01)
plot(p, dbeta(p, 1, 10), type="l", xlab="p", ylab="Beta Density",
lty=1, lwd=3)
```



[2] Adapted from *Bayes Rules! An Introduction to Bayesian Modeling with R* (Alicia A. Johnson, Miles Ott, Mine Dogucu).

The prior distribution of p shows that the experts are fairly certain that only a small proportion of people will do what authority tells them.

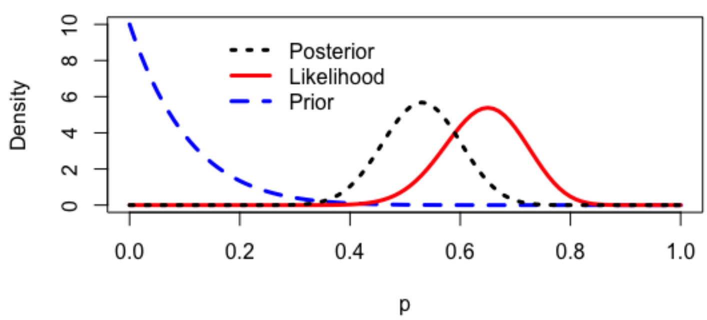
(c)

$$\begin{aligned} X|p &\sim \text{Binomial}(n = 40, p) \\ p &\sim \text{Beta}(a = 1, b = 10) \end{aligned}$$

(d) The Beta distribution Beta(a, b) is a conjugate prior for the Binomial likelihood. The posterior distribution will be Beta with parameters $p|x \sim \text{Beta}(a+x, b+n-x)$. In our case, $a =$, $b = 10$, $n = 40$ and $x = 26$. The posterior distribution of the parameter is:

$$p|x \sim \text{Beta}(27, 24)$$

(e)
```
alpha = 1
beta = 10
x = 26
n = 40
prior = dbeta(p, alpha, beta)
posterior = dbeta(p, (x + alpha), (n-x+beta))
like = dbeta(p, x+1, n-x+1)  # scaled likelihood
# likelihood = dbinom(x,n,p)  # likelihood function
plot(p, prior, type="l", xlab="p", ylab="Density", lty=2, lwd=3,
col='blue')
lines(p, like, lty=1, lwd=3, col='red')
lines(p, posterior, lty=3, lwd=3)
legend(0.15, 10, c("Posterior", "Likelihood", "Prior"), lty=c(3,1,2),
lwd=c(3,3,3), col=c("black", "red", "blue"), bty="n")
```



(f) The mean of a random variable X with distribution Beta(a, b) is $E(X) = \frac{a}{a+b}$. Therefore, the prior mean for p is $1/11 \approx 0.09$, and the posterior mean is $27/(27 + 24) \approx 0.53$. Initially psychologists (and I) thought that only a small proportion of people will obey orders from authorities when this could cause severe harm to other people. But after observing the data, now we believe that this proportion is about a half.

4. Consider the following model:

$$\begin{aligned} X|\lambda &\sim \text{Poisson}(\lambda) \\ \lambda &\sim \text{Gamma}(a, b) \end{aligned}$$

Recall the Poisson distribution

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!} \text{ for } x \in 0, 1, \ldots$$

with $E(X) = V(X) = \lambda$, and the Gamma distribution

$$f(x) = \frac{1}{\Gamma(a) b^a} x^{a-1} e^{-x/b} \text{ for } x \geq 0$$

with $E(X) = ab$ and $V(X) = ab^2$.

(a) Find analytically the posterior distribution of λ given the data.

(b) Suppose that $X \sim P(\lambda)$ is the number of patients arriving to an emergency ward in a given week. Experts from the hospital believe that the average value of λ is 30 and the variance is 180. With this information construct a Gamma prior distribution for λ (i.e. find the values of the parameters a and b).

(c) We collect data and find that in a given week 42 patients arrive to the emergency ward. With this information, plot the likelihood, prior and posterior distributions as function of λ in R.

(d) Find the posterior mean of λ and approximate the posterior median by simulating a sample from the posterior distribution.

(e) The posterior predictive distribution for the model presented above is

$$\tilde{x}|x \sim \text{Negative Binomial}\left(x + a, \frac{b}{2b+1}\right).$$

Use the command `dnbinom` to plot the posterior distribution for $\tilde{x} \in [0, 100]$ and find the posterior predictive probability that the number of patients arriving is between 40 and 45 (e.g. by using the R command `pnbinom`).

Solution.

(a) By Baye's Theorem we know that

$$\begin{aligned} p(\lambda|x) &\propto p(x|\lambda)p(\lambda) \\ &\propto \frac{\lambda^x}{x!} e^{-\lambda} \frac{1}{\Gamma(a) b^a} \lambda^{a-1} e^{-\lambda/b} \\ &\propto \lambda^x e^{-\lambda} \lambda^{a-1} e^{-\lambda/b} \\ &\propto \lambda^{a+x-1} e^{-\lambda(1+1/b)} \end{aligned}$$

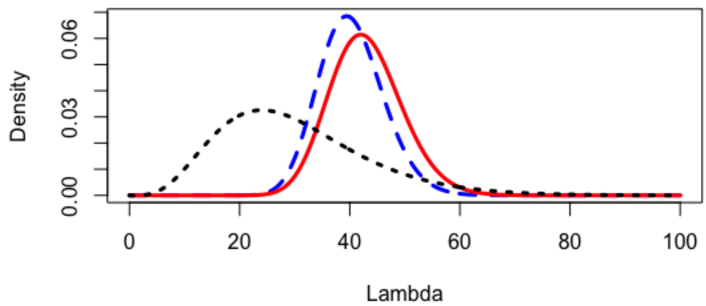
Thus, $\lambda|x \sim \text{Gamma}(a+x, \frac{1}{1+1/b})$, i.e. $\text{Gamma}(a+x, \frac{b}{1+b})$.

(b) We need to solve the following system of equations:

$$\begin{cases} E(\lambda) = ab = 30 \\ V(\lambda) = ab^2 = 180 \end{cases}$$

We find that $a = 5$ and $b = 6$. The prior distribution of λ is Gamma$(5, 6)$.

(c)
```
lambda=seq(0,100, length=500)
x = 42
a = 5
b = 6
prior = dgamma(lambda, 5, scale=6)
posterior = dgamma(lambda, (x + a), scale=b/(b+1))
like = dgamma(lambda, x+1, scale=1)  # scaled likelihood
# likelihood = dbinom(x,n,p)  # likelihood function
plot(lambda, posterior, type="l", xlab="Lambda", ylab="Density",
lty=2, lwd=3, col='blue')
lines(lambda, like, lty=1, lwd=3, col='red')
lines(lambda, prior, lty=3, lwd=3)
legend(0.15, 10, c("Prior", "Likelihood", "Posterior"), lty=c(3,1,2),
lwd=c(3,3,3), col=c("black", "red", "blue"), bty="n")
```



(d) By using the formula of the expectation for the Gamma distribution we find the posterior mean as:

$$E(\lambda|x) = (a+x)\frac{b}{1+b} = 40.29$$

To calculate the posterior median we do:

```
sample = rgamma(1000, (x + a), scale=b/(b+1))
median(sample)
```