Session 2 - Bayesian Inference and Markov Chain Monte Carlo - Exercises

1. [**Advanced**] In the session notes we showed that the posterior distribution that results from applying Bayes rule with an uninformed prior and a Bernoulli likelihood function is a beta probability density function. Show that this is also the result when the prior is already a beta probability density function and it is updated with a further Bernoulli likelihood function.

2. A coin is flipped 8 times and comes up heads just two times. Use PyMC3 to model this setting and generate a trace representing the posterior probability of the coin coming up heads. Use that trace to calculate:

   (a) The probability of the next coin flip coming up heads.

   (b) The probability of the next two coin flips both coming up heads.

   (c) The probability of any three out of the next five coin flips coming up heads.

3. Consider the case where a coin has been flipped 6 times and has come up heads twice. This can be represented as a prior with a beta probability density function using the built-in beta function as below:

   ```
   p = pm.Beta('p', 3, 5)
   ```

   Build an appropriate model in PyMC3 using this prior and generate a trace representing the posterior probability of the coin coming up heads assuming that two more coin flips are made and both of them come up heads.

4. For what values of a and b does the beta probability density function have a value greater than 0 at $p = 0$ and $p = 1$. What does this mean in terms of the results of the coin flipping experiments discussed in this session.

Session 2 - Bayesian Inference and Markov Chain Monte Carlo - Solutions

1. We assume that we have a prior distribution represented by a beta probability density function, $\mathrm{Beta}(p|a,b)$, and a likelihood function given by a Bernoulli distribution, $P(n|p,N)$. We can write out Bayes rule to describe the posterior probability:

$$p(p|n,N) = \frac{P(n|p,N)\mathrm{Beta}(p|a,b)}{\int_0^1 P(n|p,N)\mathrm{Beta}(p|a,b)dp} \tag{1}$$

We can substitute in the expressions for the likelihood function and the prior:

$$p(p|n,N) = \frac{\frac{N!}{n!(N-n)!}p^n(1-p)^{N-n}\frac{1}{B(a,b)}p^{a-1}(1-p)^{b-1}}{\int_0^1 \frac{N!}{n!(N-n)!}p^n(1-p)^{N-n}\frac{1}{B(a,b)}p^{a-1}(1-p)^{b-1}dp} \tag{2}$$

This looks very messy but the binomial coefficients and the beta functions do not depend on p and thus will cancel to give:

$$p(p|n,N) = \frac{p^n(1-p)^{N-n}p^{a-1}(1-p)^{b-1}}{\int_0^1 p^n(1-p)^{N-n}p^{a-1}(1-p)^{b-1}dp} \tag{3}$$

We can then combine together terms:

$$p(p|n,N) = \frac{p^{a+n-1}(1-p)^{b+N-n-1}}{\int_0^1 p^{a+n-1}(1-p)^{b+N-n-1}dp} \tag{4}$$

Now we can change variable such that, $a' = a+n$ and $b' = b+N-n$, to give:

$$p(p|n,N) = \frac{p^{a'-1}(1-p)^{b'-1}}{\int_0^1 p^{a'-1}(1-p)^{b'-1}dp} \tag{5}$$

We are now back to the same situation that we had in the second session and we can simply recognise this as the definition of a beta probability density function such that:

$$p(p|n,N) = \mathrm{Beta}(p|a',b') \tag{6}$$

where the two parameters of the beta probability density function are simply determined by adding the number of successful Bernoulli trials to a and the number of unsuccessful trials to b.

2. See `problem_solutions.ipynb`.

   (a) 0.2992 : `np.mean(posterior)`

   (b) 0.1087 : `np.mean(posterior**2)`

   (c) 0.1394 : `np.mean(likelihood(3,posterior,5))`

3. See `problem_solutions.ipynb`.

4. When $a \geq 1$ we are specifiying that we have seen at least one head arise from the coin flips. This means that p cannot be equal to zero. Thus the probability density function goes to zero at $p = 0$. The same applies for when $b \geq 1$.