Assignment 1

Student ID: u7205329

Exercise 1.

1. Prove that $E[x+y] = E[x] + E[y]$

$$E[x+y] = \int_x \int_y (x+y)\, f(x,y)\, dx\, dy$$

$$= \int_x \int_y x\, f(x,y)\, dx\, dy + \int_x \int_y y\, f(x,y)\, dx\, dy$$

since X and Y are independent, $f(x,y) = f(x)\, g(y)$

$$= \int_x \int_y x\, f(x)\, f(y)\, dx\, dy + \int_x \int_y y\, f(x)\, f(y)\, dx\, dy$$

$$= \int_x x\, f(x)\, dx \int_y f(y)\, dy + \int_x f(x)\, dx \int_y y\, f(y)\, dy$$

$$= E[x] \,.\, 1 + 1 \,.\, E[y]$$

$$= E[x] + E[y]. \quad \square$$

2. Prove that $E[xy] = E[x]\, E[y]$

$$E[xy] = \int_x \int_y xy\, f(x,y)\, dx\, dy = \int_x \int_y xy\, f(x)\, f(y)\, dx\, dy$$

$$= \int_x x\, f(x)\, dx \int_y y\, f(y)\, dy$$

$$= E[x]\, E[y] \quad \square$$

3. Prove that $\mathrm{cov}[x,y] = 0$

$$\mathrm{cov}[x,y] = E[(x - E[x])(y - E[y])]$$

$$= E\left[xy - xE[y] - yE[x] + E[x]E[y] \right]$$

$$= E[xy] - E\left[xE[y]\right] - E\left[yE[x]\right] + E\left[E[x]E[y]\right]$$

$$= \underbrace{E[xy]}_{\text{from above}} - E[y]E[x] - E[x]E[y] + E[x]E[y]$$

$$= E[x]E[y] - E[x]E[y]$$

$$= 0. \qquad \square$$

Exercise 2.

1. Prove that $0 < B(a,b) \leq 1$ for all $a, b \geq 1$.

We have $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\,dx$ with $x \in [0,1]$ and $a, b \geq 1$

let $f(x) = (1-x)^{b-1}$ and $g(x) = x^{a-1}$

From the mean value theorem,

$$m \int_0^1 g(x)\,dx \leq \int_0^1 f(x)\,g(x)\,dx \leq M \int_0^1 g(x)\,dx \qquad (\star)$$

where $m = \min f(x)$ and $M = \max f(x)$ for $x \in [0,1]$

$\min f(x) = f(1) = 0$ and $\max f(x) = f(0) = 1$

continue from $(\star)$, we have

$$0 \int_0^1 g(x)\,dx \leq \int_0^1 f(x)\,g(x)\,dx \leq 1 . \int_0^1 g(x)\,dx$$

$$0 \leq \int_0^1 x^{a-1}(1-x)^{b-1}\,dx \leq 1 . \int_0^1 x^{a-1}\,dx$$

$$0 \leq \int_0^1 x^{a-1}(1-x)^{b-1}\,dx \leq \frac{1}{a}$$

since the minimum value of $a = 1 \Rightarrow \frac{1}{a} \leq 1$

Finally, $0 \leq \int_0^1 x^{a-1}(1-x)^{b-1}\,dx \leq 1$. $\square$

2. Prove that Beta$(x \mid a,b)$ is well-defined and a valid pdf.

Firstly, to prove Beta$(x \mid a,b)$ is well-defined, we show

$$\text{Beta}(x \mid a,b) \geq 0 \text{ for all } x,\ x \in [0,1]$$

As $\text{Beta}(x \mid a,b) = \dfrac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}$

we have shown that $B(a,b)$ is continuous on $[0,1] \Rightarrow \dfrac{1}{B(a,b)} \geq 1$.

However, for Beta$(x \mid a,b)$ to be defined, $B(a,b)$ has to be $\neq 0$.

We also have $x^{a-1} \geq 0 \quad \forall x \in [0,1]$ and $a \geq 1$.

Similarly $(1-x)^{b-1} \geq 0 \quad \forall x \in [0,1]$ and $b \geq 1$.

As Beta$(x \mid a,b)$ is a product of these 3 terms

$\Rightarrow \text{Beta}(x \mid a,b) \geq 0 \quad \forall x \in [0,1]$ and $a,b \geq 1$.

Secondly, to prove that Beta$(x \mid a,b)$ is valid

we need to show $\int_0^1 \text{Beta}(x \mid a,b)\,dx = 1$

$$\begin{aligned}\int_0^1 \text{Beta}(x \mid a,b)\,dx &= \int_0^1 \frac{1}{B(a,b)} \cdot x^{a-1}(1-x)^{b-1}\,dx \\ &= \frac{1}{B(a,b)} \int_0^1 x^{a-1}(1-x)^{b-1}\,dx \\ &= \frac{1}{B(a,b)} \cdot B(a,b) \\ &= 1. \quad \square\end{aligned}$$

3. Prove the identity $B(a+1, b) = \frac{a}{a+b} B(a,b)$

$$B(a+1, b) = \frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+b+1)} = \frac{a}{a+b} \cdot \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

$$= \frac{a}{a+b} \cdot B(a,b) . \quad \square$$

Exercise 3.

1. Express the uniform prior (θ) in terms of the beta prior.

$$p(\theta) = \frac{1}{B(a,b)} \theta^{a-1}(1-\theta)^{b-1}$$

2. Show that if the prior is chosen to be prior, and if we observe a sequence s of coin flips containing h ones and t zeros, the posterior distribution is $p(\theta | s) = \mathrm{Beta}(\theta | a+h, b+t)$ for $a, b \geqslant 1$.

We have for each coin flip x_i, the probability distribution for the outcome is $p(x_i | \theta) = \theta^{x_i}(1-\theta)^{1-x_i}$

Hence, we have the likelihood function

$$f(x_1, x_2, \dots, x_s | \theta)$$

$$= f(x_1 | \theta) \, f(x_2 | \theta) \dots f(x_s | \theta)$$

$$= \theta^{x_1}(1-\theta)^{1-x_1} \, \theta^{x_2}(1-\theta)^{1-x_2} \dots \theta^{x_s}(1-\theta)^{1-x_s}$$

$$= \theta^{\sum_{i=1}^{s} x_i}(1-\theta)^{s - \sum_{i=1}^{s} x_i}$$

$$= \theta^{h}(1-\theta)^{s-h}$$

$$= \theta^{h}(1-\theta)^{t} .$$

It follows that the joint likelihood $X_1, \dots, X_s, \theta$ is

$$\begin{aligned} f(x_1, \dots, x_s, \theta) &= f(x_1, \dots, x_s \mid \theta)\, f(\theta) \\ &= \theta^h (1-\theta)^t \frac{1}{B(a,b)} \theta^{a-1} (1-\theta)^{b-1} \\ &= \frac{1}{B(a,b)} \theta^{(a+h)-1} (1-\theta)^{(b+t)-1} \end{aligned}$$

and the marginal density of $X_1, \dots, X_s$ is

$$\begin{aligned} f(x_1, x_2, \dots, x_s) &= \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_s, \theta)\, d\theta \\ &= \int_0^1 \frac{1}{B(a,b)} \theta^{(a+h)-1} (1-\theta)^{(b+t)-1}\, d\theta \\ &= \frac{1}{B(a,b)} \int_0^1 \theta^{(a+h)-1} (1-\theta)^{(b+t)-1}\, d\theta \\ &= \frac{1}{B(a,b)} \cdot B(a+h, b+t) \end{aligned}$$

Finally, the posterior distribution

$$\begin{aligned} &f(\theta \mid x_1, \dots, x_s) \\ &= \frac{f(x_1, \dots, x_s, \theta)}{f(x_1, x_2, \dots, x_s)} = \frac{\frac{1}{B(a,b)} \theta^{(a+h)-1} (1-\theta)^{(b+t)-1}}{\frac{1}{B(a,b)} \cdot B(a+h, b+t)} \\ &= \frac{1}{B(a+h, b+t)} \theta^{(a+h)-1} (1-\theta)^{(b+t)-1} \\ &= \mathrm{Beta}(\theta \mid a+h, b+t). \end{aligned}$$

3. Prove that the mean of Beta $(\theta \mid a,b)$ is $\dfrac{a}{a+b}$

$$E[\theta] = \int_0^1 \theta \cdot \frac{1}{B(a,b)} \quad \theta^{a-1}(1-\theta)^{b-1}\, d\theta$$

$$= \frac{1}{B(a,b)} \int_0^1 \theta^{(a+1)-1}(1-\theta)^{b-1}\, d\theta$$

$$= \frac{1}{B(a,b)} \cdot B(a+1, b)$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+b+1)}$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{a\Gamma(a)\Gamma(b)}{(a+b)\Gamma(a+b)} = \frac{a}{a+b}. \qquad \square$$

4. Prove that the variance is $\dfrac{ab}{(a+b+1)(a+b)^2}$

$$E[\theta^2] = \int_0^1 \theta^2 \frac{1}{B(a,b)} \quad \theta^{a-1}(1-\theta)^{b-1}\, d\theta$$

$$= \frac{1}{B(a,b)} \int_0^1 \theta^{(a+2)-1}(1-\theta)^{b-1}\, d\theta$$

$$= \frac{1}{B(a,b)} \cdot B(a+2, b)$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a+2)\Gamma(b)}{\Gamma(a+b+2)}$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{(a+1)a\Gamma(a)\Gamma(b)}{(a+b+1)(a+b)\Gamma(a+b)} = \frac{a(a+1)}{(a+b)(a+b+1)}$$

Hence, the variance of Beta $(\theta \mid a,b)$ is

$$Var[\theta] = E[\theta^2] - E[\theta]^2$$

$$= \frac{a(a+1)}{(a+b)(a+b+1)} - \left(\frac{a}{a+b}\right)^2$$

$$= \frac{a(a+1)(a+b)}{(a+b+1)(a+b)^2} - \frac{a^2}{(a+b)^2}$$

$$= \frac{(a^2+a)(a+b)}{(a+b+1)(a+b)^2} - \frac{a^2(a+b+1)}{(a+b)^2(a+b+1)}$$

$$= \frac{a^3 + a^2 b + a^2 + ab - a^3 - a^2 b - a^2}{(a+b)^2(a+b+1)}$$

$$= \frac{ab}{(a+b)^2(a+b+1)}. \qquad \square$$

5. Consider a sequence $X_{1:2N} = 010101\ldots$ containing N zeros and N ones. What is the posterior distribution after observing this sequence, given a uniform prior? What is the variance?

Given a uniform prior: $p(\theta) = \frac{1}{1-0} = 1$.

Firstly, we have the likelihood function for $X_1, X_2, \ldots, X_{2N}$

$$f(x_1, \ldots, x_{2N} \mid \theta)$$

$$= f(x_1 \mid \theta)\ f(x_2 \mid \theta) \ \ldots \ f(x_{2N} \mid \theta)$$

$$= \theta^{x_1}(1-\theta)^{1-x_1}\ \theta^{x_2}(1-\theta)^{1-x_2} \ \ldots \ \theta^{x_{2N}}(1-\theta)^{1-x_{2N}}$$

$$= \theta^{\sum_{i=1}^{2N} x_i}(1-\theta)^{2N - \sum_{i=1}^{2N} x_i}$$

$$= \theta^N (1-\theta)^{2N-N} = \theta^N (1-\theta)^N .$$

The joint likelihood is

$$f(x_1, x_2, \ldots, x_{2N}, \theta) = f(x_1, x_2, \ldots, x_{2N} \mid \theta) \cdot f(\theta)$$

$$= \theta^N (1-\theta)^N \cdot 1 = \theta^N (1-\theta)^N .$$

Marginal density

$$f(x_1, \dots, x_{2N}) = \int_0^1 f(x_1, x_2, \dots, x_{2N}, \theta)\, d\theta$$

$$= \int_0^1 \theta^N (1-\theta)^N \, d\theta$$

$$= \int_0^1 \theta^{(N+1)-1} (1-\theta)^{(N+1)-1} \, d\theta$$

$$= B(N+1, N+1)$$

Posterior

$$f(\theta \mid x_1, x_2, \dots, x_{2N}) = \frac{f(x_1, x_2, \dots, x_{2N}, \theta)}{f(x_1, x_2, \dots, x_{2N})}$$

$$= \frac{\theta^N (1-\theta)^N}{B(N+1, N+1)}$$

$$= \frac{1}{B(N+1, N+1)} \cdot \theta^{(N+1)-1} (1-\theta)^{(N+1)-1}$$

$$= \mathrm{Beta}(\theta \mid N+1, N+1)$$

The variance can quickly be computed using the variance formula prove above. $\mathrm{Var}[\theta] = \dfrac{ab}{(a+b+1)(a+b)^2}$

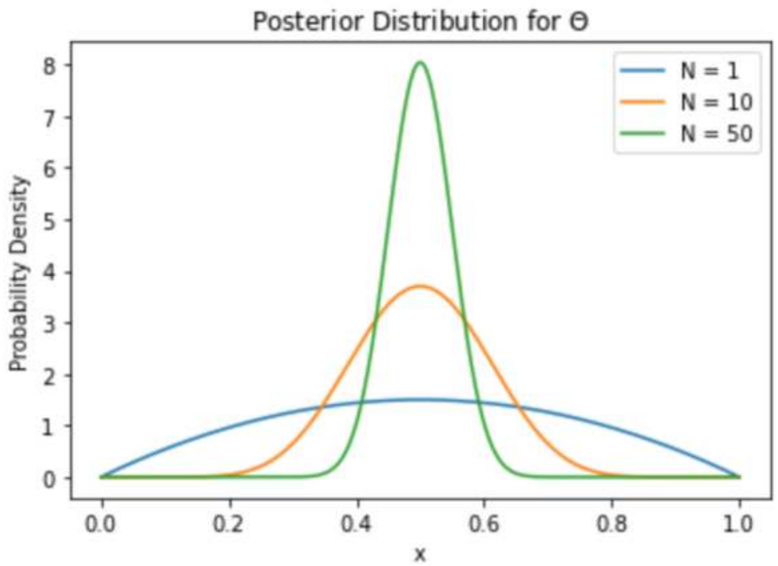
$$\Rightarrow \mathrm{Var}[\hat{\theta}_B] = \frac{(N+1)(N+1)}{[(N+1)+(N+1)+1][(N+1)+(N+1)]^2}$$

$$= \frac{(N+1)^2}{(2N+3)(2N+2)^2}$$

$$= \frac{(N+1)^2}{(2N+3)\, 4(N+1)^2}$$

$$= \frac{1}{4(2N+3)}.$$

6. Plot the posterior for N = 1, 10, 50



7. What's the smallest value for N, such that the probability that θ lies in the interval [0.49, 0.51] is at least 0.99?

We want to find the smallest N such that

$$P(0.49 \leq \theta \leq 0.51) \geq 0.99$$

Using Scipy, we found the smallest N to be 8292.

```
from scipy.stats import beta

n = 1
while True:
    p = beta.cdf(0.51, n+1, n+1) - beta.cdf(0.49, n+1, n+1)
    if p >= 0.99:
        print(n)
        break
    n += 1
```

8292

Exercise 4.

1. Show that $p(\hat{x}=1 \mid \theta, \phi) = (1-2\phi)\theta + \phi$

We have the probability of camera reports 1 is the sum of camera reports 1 when the coin flip was 1 and when the coin flip is actually 0.

$$
\begin{aligned}
& p(\hat{x}=1 \mid \theta, \phi) \\
&= p(\hat{x}=1, x=1 \mid \theta, \phi) + p(\hat{x}=1, x=0 \mid \theta, \phi) \\
&= p(x=1 \mid \theta) \cdot p(\hat{x}=1 \mid x=1, \theta, \phi) \\
&\qquad + p(x=0 \mid \theta) \cdot p(\hat{x}=1 \mid x=0, \theta, \phi) \\
&= \theta(1-\phi) + (1-\theta)\phi \\
&= \theta - \theta\phi + \phi - \theta\phi \\
&= \theta - 2\theta\phi + \phi \\
&= (1-2\phi)\theta + \phi .
\end{aligned}
$$

2. Let the prior $p(\theta \mid \phi) = \text{Beta}(\theta \mid a, b)$. Show that for all $a, b \geq 1$, and all $\phi \in [0, 0.5]$, there exists $\alpha \in [0,1]$ such that

$$p(\theta \mid \phi, \hat{x}=1) = (1-\alpha)\,\text{Beta}(\theta \mid a+1, b) + \alpha\,\text{Beta}(\theta \mid a, b)$$

Firstly, we have the joint likelihood

$$
\begin{aligned}
p(\hat{x}=1, \theta, \phi) &= p(\hat{x}=1 \mid \theta, \phi) \cdot p(\theta \mid \phi) \\
&= [(1-2\phi)\theta + \phi]\, p(\theta \mid \phi) \\
&= (1-2\phi)\theta\, p(\theta \mid \phi) + \phi\, p(\theta \mid \phi) \\
&= (1-2\phi)\theta\,\text{Beta}(\theta \mid a, b) + \phi\,\text{Beta}(\theta \mid a, b)
\end{aligned}
$$

before moving on, we have

$$\theta\,\text{Beta}(\theta \mid a, b) = \theta \cdot \frac{1}{B(a,b)}\, \theta^{a-1}(1-\theta)^{b-1}$$

$$= \frac{1}{B(a,b)} \theta^{(a+1)-1}(1-\theta)^{b-1} \frac{B(a+1,b)}{B(a+1,b)}$$

$$= \frac{1}{B(a,b)} \theta^{(a+1)-1}(1-\theta)^{b-1} \frac{a}{a+b} \frac{B(a,b)}{B(a+1,b)}$$

$$= \frac{a}{a+b} \frac{1}{B(a+1,b)} \theta^{(a+1)-1}(1-\theta)^{b-1}$$

$$= \frac{a}{a+b} \mathrm{Beta}(\theta \mid a+1, b)$$

and so

$$p(\hat{x}=1, \theta, \phi) = (1-2\phi) \frac{a}{a+b} \mathrm{Beta}(\theta \mid a+1, b) + \phi \, \mathrm{Beta}(\theta \mid a, b)$$

Hence, we have the posterior $p(\theta \mid \phi, \hat{x}=1)$

$$= \frac{p(\hat{x}=1, \theta \mid \phi)}{p(\hat{x}=1 \mid \phi)}$$

$$= \frac{(1-2\phi)\frac{a}{a+b} \mathrm{Beta}(\theta \mid a+1, b) + \phi \, \mathrm{Beta}(\theta \mid a, b)}{p(\hat{x}=1 \mid \phi)}$$

$$\therefore \frac{(1-2\phi)\frac{a}{a+b}}{p(\hat{x}=1 \mid \phi)} = 1 - \alpha \quad (1) \quad \text{and} \quad \frac{\phi}{p(\hat{x}=1 \mid \phi)} = \alpha \quad (2)$$

3. Find a closed form solution for α in terms of a, b, and ϕ.

From (2), we have $p(\hat{x}=1 \mid \phi) = \frac{\phi}{\alpha}$. Plugging this to (1)

$$\frac{(1-2\phi)\frac{a}{a+b}}{\frac{\phi}{\alpha}} = 1 - \alpha \iff \frac{(1-2\phi)a}{(a+b)} \cdot \frac{\alpha}{\phi} = 1 - \alpha$$

Let $c = \frac{(1-2\phi)a}{(a+b)\phi} \Rightarrow c\alpha = 1 - \alpha$

$$\iff \alpha = \frac{1}{c+1}$$

$$\Leftrightarrow \quad \alpha = \frac{1}{\frac{(1-2\phi)a}{(a+b)\phi} + 1}$$

$$\Leftrightarrow \quad \alpha = \frac{1}{\frac{(1-2\phi)a}{(a+b)\phi} + \frac{(a+b)\phi}{(a+b)\phi}}$$

$$\Leftrightarrow \quad \alpha = \frac{(a+b)\phi}{(1-2\phi)a + (a+b)\phi}$$

which confirms that $\alpha \in [0,1]$.

4. Show that the expression for α depends only on ϕ and b/a.

Continue from above, $\alpha = \frac{(a+b)\phi}{(1-2\phi)a + (a+b)\phi}$

$$= \frac{a\phi + b\phi}{a - 2a\phi + a\phi + b\phi}$$

$$= \frac{a\phi + b\phi}{a - a\phi + b\phi}$$

$$= \frac{\phi + (b/a)\phi}{1 - \phi + (b/a)\phi}. \quad \square$$