Jim Liu

Homework 3

ISyE 6420
October 4, 2020

Problem 1

Answer to the problem goes here.

(a)

$$
\begin{aligned}
P(\theta|y) &\propto P(y|\theta)P(\theta) \\
&= \prod_{i=1}^{n} \sqrt{2/\pi}\theta^{3/2} y_i^2 e^{-\theta y_i^2/2} \lambda e^{-\lambda\theta} \\
&\propto \theta^{\frac{3n}{2}} e^{-\lambda\theta - \frac{\theta}{2}\sum_{i=1}^{n} y_i^2} \\
&= \theta^{\frac{3n}{2}+1} - 1 e^{-\frac{\frac{\theta}{2}}{2\lambda + \sum_{i=1}^{n} y_i^2}} \\
&\sim Gamma(\alpha = \frac{3n}{2} + 1, \beta = \frac{2}{2\lambda + \sum_{i=1}^{n} y_i^2})
\end{aligned}
$$

(b)

$$
\begin{aligned}
E(\theta|y) &= (1.5n+1)(\frac{2}{2\lambda + \sum_{i=1}^{n} y_i^2}) \\
&= \frac{11}{2} \times \frac{2}{2 \times \frac{1}{2} + (1.4^2 + 3.1^2 + 2.5^2)} \\
&= \frac{11}{18.82} \\
&\approx 0.58 \\
\theta_{\text{MLE}} &= \frac{3}{\bar{y}^2} = \frac{3}{5.94} \approx 0.51 \\
\theta_{\text{prior}} &= E(\pi(\theta)) = \frac{1}{2} = 0.5
\end{aligned}
$$

(c) Problem 1 part 3 answer here.

```python
# Plot pdf to know range
from scipy.stats import gamma
fig, ax = plt.subplots(1, 1)
alpha = 11/2
beta = 0.10626993
x = np.linspace(gamma.ppf(0.01, a=alpha, scale=beta),
                gamma.ppf(0.99, a=alpha, scale=beta), 100)
ax.plot(x, gamma.pdf(x, a=alpha, scale=beta))
plt.axvline(x=gamma.ppf(0.025, a=alpha, scale=beta),color='r')
plt.axvline(x=gamma.ppf(0.975, a=alpha, scale=beta),color='r')
plt.title("Gamma 95% Equitailed Credible Sets")
plt.show()
```

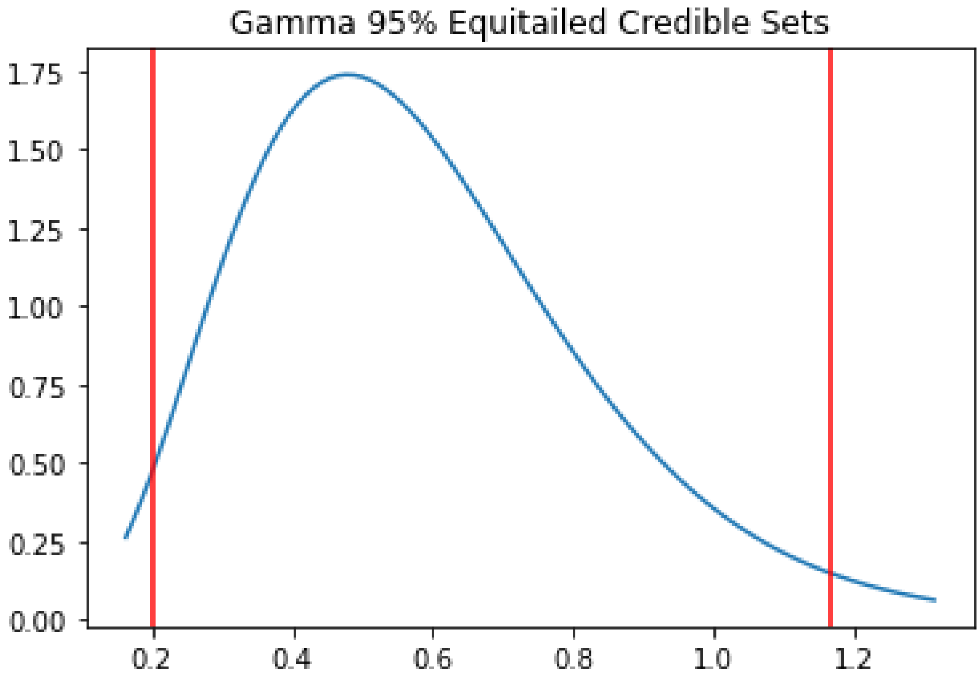

Figure 1: Gamma 95% Equitailed Credible set

θ range in 95% Equitailed Credible Set:[0.2027496498313763, 1.1647210502826375]

(d) Given y_4 a prediction,

$$
\begin{aligned}
m(x) &= \int f(y|\theta)\pi(\theta)\mathrm{d}\theta \\
f(y_4|y_1, y_2, y_3) &= \int f(y_4|\theta)\pi(\theta|y_1, y_2, y_3)\mathrm{d}\theta \\
\hat{y_4} &= \int f(y_4)f(y_4|y_1, y_2, y_3)\mathrm{d}y_4 \\
&= \int_0^\infty y_4\Big[\int_0^\infty \prod_{i=1}^{3}\sqrt{\frac{2}{\pi}}\theta^{1.5}y_i^2 e^{-\frac{\theta y_i^2}{2}}\lambda e^{-\lambda\theta}\mathrm{d}\theta\Big]\mathrm{d}y_4 \\
&= \int_0^\infty 0.013 f(y_4)\mathrm{d}y_4 \text{ by following integration codes} \\
&= 0.013E(y_4) = 0.013 \times 2\sqrt{\frac{2}{\pi\theta}} \\
&= 0.026\sqrt{\frac{2}{\pi\theta}}
\end{aligned}
$$

```
from scipy.integrate import quad
def integrand(theta):
    return (2/np.pi ** 1.5) * theta ** 4.5
    * 1.4 ** 2 * 3.1 ** 2 * 2.5 ** 2
    * np.exp(-(theta/2)*(1.4**2+3.1**2+2.5**2))

quad(integrand, 0, np.inf)
(0.013203651633076267, 5.038286154500041e-11)
```

Problem 2

(a)

$$
\begin{aligned}
\pi(\theta) &= \sum_{i=1}^{2} k_i \pi_i(\theta) = \epsilon\pi_1(\theta) + (1-\epsilon)\pi_2(\theta) \\
m(x) &= \int_\Theta f(x|\theta)\pi_i(\theta)\mathrm{d}\theta \\
&= \int_\Theta \epsilon f(x|\theta)\pi_1(\theta)\mathrm{d}\theta + \int_\Theta (1-\epsilon) f(x|\theta)\pi_2(\theta) \\
&= \epsilon m_1(x) + (1-\epsilon) m_2(x) \\
\pi(\theta|x) &= \frac{f(x|\theta)\pi(\theta)}{m(x)} \\
&= \frac{m_1(x)\epsilon\pi_1(\theta)f(x|\theta)/\epsilon m_1(x) + m_2(x)(1-\epsilon)\pi_2(\theta)f(x|\theta)/(1-\epsilon)m_2(x)}{m(x)} \\
&= \frac{m_1(x)\pi_1(\theta|x) + m_2(x)\pi_2(\theta|x)}{m(x)}
\end{aligned}
$$

To make it follow definitions of probability

$$
\begin{aligned}
&\implies \frac{m_1(x) + m_2(x)}{m(x)} = 1 \\
&\implies m_2(x) = 1 - \frac{m_1(x)}{m(x)} = 1 - \epsilon' \\
\epsilon' &= \frac{m_1(x)}{m(x)}
\end{aligned}
$$

(b) Find the posterior and Bayes estimator for θ if $X = 98$.

Posterior Probability: by Unit 4.2 for Normal likelihood + Normal Prior, we get

$$\begin{aligned}
\pi_1(\theta|x) &= \frac{2/3N(\theta,80)\times N(110,60)}{\int_\Theta N(\theta,80)2/3N(110,60)\mathrm{d}\theta} \\
&\propto N(\frac{80^2}{80^2+60^2}x+\frac{60^2}{80^2+60^2}110,\frac{80^2 60^2}{80^2+60^2}) \\
\pi_2(\theta|x) &= \frac{1/3N(\theta,80)\times N(100,200)}{\int_\Theta N(\theta,80)1/3N(100,200)\mathrm{d}\theta} \\
&\propto N(\frac{80^2}{80^2+200^2}x+\frac{200^2}{80^2+200^2}100,\frac{200^2 80^2}{200^2+80^2})
\end{aligned}$$

$$\begin{aligned}
\pi(\theta|x) &= \alpha\pi_1(\theta|x)+\beta\pi_2(\theta|x) \\
\alpha &\propto \int f(x|\theta)\pi_1(\theta) = \int \frac{1}{2\pi 80\times 60}exp(-0.5\frac{(x-\theta)^2}{80^2}-0.5\frac{(\theta-110)^2}{60^2})\mathrm{d}\theta \\
& exp(-0.5\frac{(x-\theta)^2}{80^2}-0.5\frac{(\theta-110)^2}{60^2}) \\
&= exp(-0.5(\frac{x^2-2\theta x+\theta^2}{80^2}+\frac{\theta^2-220\theta+110^2}{60^2}))
\end{aligned}$$

$$\begin{aligned}
&= exp(-0.5(\frac{1}{80^2 60^2}[10000\theta^2-(705600+1408000)\theta+3600\times 9604+6400\times 12100])) \\
&= exp(-0.5(\frac{1}{80^2 60^2}(10000(\theta-105.68)^2)-111682624+3600\times 9604+6400\times 12100)
\end{aligned}$$

$$\begin{aligned}
&= exp(-0.5(\frac{1}{80^2 60^2}[10000\theta^2-(705600+1408000)\theta+3600\times 9604+6400\times 12100])) \\
&= exp(-0.5(\frac{1}{80^2 60^2}(10000(\theta-105.68)^2)-111682624+3600\times 9604+6400\times 12100)
\end{aligned}$$

$$\begin{aligned}
\alpha &\propto \frac{1}{100\sqrt{2\pi}}e^{0.0144}\int \frac{1}{\sqrt{2\pi 80^2 60^2/100^2}}exp(-0.5\frac{(\theta-105.68)^2}{80^2 60^2/100^2}) \\
&\propto 100 * \frac{1}{100\sqrt{2\pi}}e^{0.0072} \\
&\propto 0.4
\end{aligned}$$

$$\beta = \int f(x|\theta)\pi_2(\theta) = \int \frac{1}{2\pi 80 * 200} exp(-0.5(\frac{x-\theta}{80})^2 - 0.5(\frac{\theta - 100}{200}^2))$$
$$= \int \frac{1}{2\pi 80 * 200} exp(-0.5(\frac{200^2(x^2 - 2\theta x + \theta^2) + 80^2(\theta^2 - 200\theta + 100^2)}{80^2 200^2}))$$
$$\int \frac{1}{2\pi 80 * 200} exp(-0.5\frac{1}{80^2 200^2} 46400\theta^2 - 9120000\theta + 200^2 x^2 + 80^2 100^2)$$
$$\int \frac{1}{2\pi 80 * 200} exp(-0.5\frac{1}{80^2 200^2}(46400(\theta - 98.2758621)^2 + 22069))$$
$$= \frac{1}{\sqrt{2\pi 46400}} e^{0.0000431} \int \frac{1}{\sqrt{2\pi 80^2 200^2/46400}} exp(-0.5(\theta - 98.2758621)^2/(80^2 200^2/46400)$$
$$\propto 100 * \frac{1}{\sqrt{2\pi 46400}} e^{0.0000431} \approx 0.18$$

$$\alpha' = 2/3 * 0.4/(2/3 * 0.4 + 1/3 * 0.18) = 0.81$$
$$\beta' = 1 - \alpha' = 0.19$$

So posterior probability is

$$\pi(\theta|x) = 0.81\pi_1(\theta|x) + 0.19\pi_2(\theta|x)$$

And Bayes estimator is

$$E(\pi(\theta|x)) = 0.81 * (\frac{80^2}{80^2 + 60^2} 98 + \frac{60^2}{80^2 + 60^2} 110) + 0.19 * (\frac{80^2}{80^2 + 200^2} 98 + \frac{200^2}{80^2 + 200^2} 100)$$
$$\approx 102$$

Problem 3

(a) What are prior and posterior means?

$$
\begin{aligned}
prior : \pi(p) &\sim Be(12, 4) \\
x|p &\sim Bin(n, p) \\
E(\pi(p)) &= 12/(12+4) = 0.75 \\
pi(p|x) &\propto f(x|p)\pi(p) \\
&= \prod_{i=1}^{m} \binom{n}{x} p^{x_i}(1-p)^{n-x_i} \frac{p^1 1(1-p)^3}{B(12,4)} \\
&= constant \times \frac{p^{\sum_{i=1}^{m} x_i + 12 - 1}(1-p)^{nm - \sum_{i=1}^{m} x_i + 3}}{B(\sum_i = 1^m x_i + 12, nm - \sum_i = 1^m x_i + 4)} \\
&\sim Beta(\sum_{i=1}^{m} x_i + 12, nm - \sum_{i=1}^{m} x_i + 4) = Beta(799, 281) m = 1, n = 1064, x_1 = 787 \\
E(\pi(p|x)) &= \frac{\sum_{i=1}^{m} x_i + 12}{\sum_{i=1}^{m} x_i + 12 + nm - \sum_{i=1}^{m} x_i + 4} \\
&= \frac{787 + 12}{787 + 12 + 1064 - 787 + 4} \approx 0.74
\end{aligned}
$$

(b) Find posterior probability of hypothesis $H_0 : p \leq 3/4$?

```
from scipy.stats import beta
# beta(787+12, 1064-787+4)
beta.cdf(3/4, 787+12, 1064-787+4, loc=0, scale=1)
0.7758595145276612
```

So $P(H_0 : p \leq 3/4) = 0.7758595145276612$

(c) Find a 95% equitailed credible set for the true proportion of tall height plants obtained from the given cross.

```
# Plot pdf to know range
fig3, ax3 = plt.subplots(1, 1)
xx = np.linspace(0, 1, 100)
a=787+12
b=1064-787+4
ax3.plot(xx, beta.pdf(xx, a, b))
plt.axvline(x=beta.ppf(0.025, a, b),color='r')
plt.axvline(x=beta.ppf(0.975, a, b),color='r')
plt.title("Beta 95% Equitailed Credible Sets")
plt.show()
```

```
print("p range in 95% Equitailed Credible Set
:[{}, {}]".format(beta.ppf(0.025, a, b), beta.ppf(0.975, a, b)))

p range in 95% Equitailed Credible Set:
[0.7132478379195061, 0.7655405496526497]
```

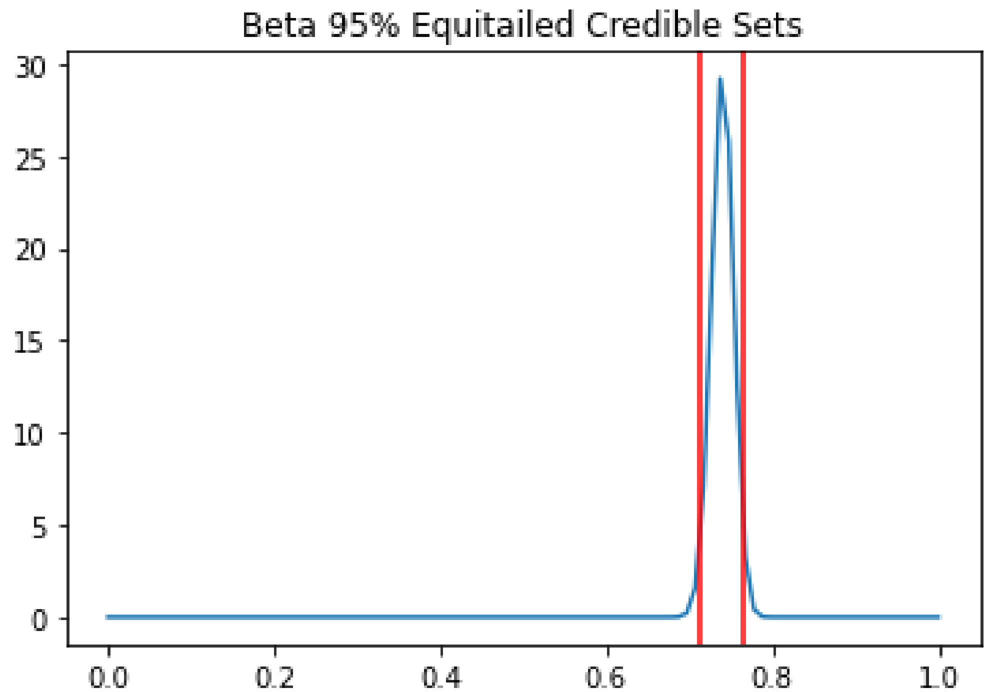

Figure 2: Beta 95% Equitailed Credible set