STAT 638: Solution for Homework #2

3.1 (a) Since $Y_1, \ldots, Y_{100}$ are, conditional on θ, i.i.d. binary random variables,

$$P(Y_1 = y_1, \ldots, Y_{100} = y_{100}|\theta) = \prod_{i=1}^{100} \theta^{y_i}(1-\theta)^{1-y_i} = \theta^{\sum_{i=1}^{100} y_i}(1-\theta)^{100-\sum_{i=1}^{100} y_i}$$

$P(\sum_{i=1}^{100} Y_i = y|\theta)$ is a binomial distribution and has form,
$P\left(\sum_{i=1}^{100} Y_i = y|\theta\right) = \binom{100}{y}\theta^y(1-\theta)^{100-y}$

(b) $P\left(\sum_{i=1}^{100} Y_i = 57|\theta\right) = \binom{100}{57}\theta^{57}(1-\theta)^{100-57}$
Plug $\theta = \{0, 0.1, \ldots, 1\}$ in, and we get the following table.

Table 1: Problem 3.1-b

θ	$P(\sum_{i=1}^{100} Y_i = 57\|\theta)$
0	0
0.1	4.107157e-31
0.2	3.738459e-16
0.3	1.306895e-08
0.4	2.285792e-04
0.5	3.006864e-02
0.6	6.672895e-02
0.7	1.853172e-03
0.8	1.003535e-07
0.9	9.395858e-18
1.0	0

(c) By Bayes' Rule $P\left(\theta\middle|\sum_{i=1}^{100} Y_i = 57\right) = \dfrac{P(\theta)P\left(\sum_{i=1}^{100} Y_i = 57\middle|\theta\right)}{\sum_\theta P(\theta)P\left(\sum_{i=1}^{100} Y_i = 57\middle|\theta\right)}$

Since $P(\theta) = \dfrac{1}{11}$ and $P\left(\sum_{i=1}^{100} Y_i = 57\middle|\theta\right)$ is given by part (b), we can get $P\left(\theta\middle|\sum_{i=1}^{n} Y_i = 57\right)$. See Table 2.

Table 2: Problem 3.1-c

θ	$P(\theta\|\sum_{i=1}^{100} Y_i = 57)$
0	0
0.1	4.153701e-30
0.2	3.780824e-15
0.3	1.321705e-07
0.4	2.311695e-03
0.5	3.040939e-01
0.6	6.748515e-01
0.7	1.874172e-02
0.8	1.014907e-06
0.9	9.502335e-17
1.0	0

(d) Since the prior distribution is Uniform[0,1], the posterior density is

$$P(\theta|Y) \propto P(\theta)P(Y|\theta) = \binom{100}{57}\theta^{57}(1-\theta)^{100-57}.$$

(e) The prior of part (c) is discrete uniform and that of part (d) is uniform(0,1). The posterior of part (c) is discrete but well approximated by the Beta(58,44), which is the posterior in (d).

3.3 (a)

$$P(\theta|Y) \propto P(\theta)P(Y|\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta} \times \frac{\theta^{\sum y_i}e^{-n\theta}}{\prod_{i=1}^n y_i!} \propto \theta^{\sum y_i+\alpha-1}e^{-\theta(n+\beta)}$$

Hence, the posterior distribution of θ_A is Gamma(237,20).

$$E(\theta_A|Y_A) = \frac{\alpha}{\beta} = 11.85$$
$$Var(\theta_A|Y_A) = \frac{\alpha}{\beta^2} = 0.5925$$

95% quantile based Confidence Interval : (10.39,13.41).

The posterior distribution of θ_B is Gamma(125,14).

$$E(\theta_B|Y_B) = \frac{\alpha}{\beta} = 8.9286$$
$$Var(\theta_B|Y_B) = \frac{\alpha}{\beta^2} = 0.6378$$

95% quantile based confidence interval : (7.43,10.56).

(b) $\theta_B \sim Gamma(12 \times n_0, n_0)$
Then the posterior density is $Gamma(12 \times n_0 + 113, n_0 + 13)$
In order for the posterior expectation of θ_B to be close to the expectation of θ_A, we need

$$\frac{12 \times n_0 + 113}{n_0 + 13} = 11.85 \quad \Rightarrow n_0 = 274.$$

That means in order for the posterior expectation of θ_B to be close to that of θ_A, the variance of the prior of θ_B should be small. In other words, strong beliefs about θ_B are necessary.

(c) Although Type B mice are unknown, type B mice are related to type A mice. Therefore, tumor counts for type B mice may not be independent of those for type A mice. It may not be sensible to have $P(\theta_A, \theta_B) = P(\theta_A)P(\theta_B)$.

3.4 (a) $P(\theta) = \frac{1}{B(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}$

$P(y|\theta) = \binom{n}{y}\theta^y(1-\theta)^{n-y}$

$P(\theta|y) \propto \frac{1}{B(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}\binom{n}{y}\theta^y(1-\theta)^{n-y} \propto \theta^{\alpha+y-1}(1-\theta)^{n-y+\beta-1}$

which is Beta($\alpha + y$, $n + \beta - y$). Hence, the posterior distribution is Beta(17,36).

$$E(\theta|Y) = \frac{\alpha}{\alpha+\beta} = 0.3208$$

$$\text{Mode}(\theta|Y) = \frac{\alpha-1}{\alpha+\beta-2} = 0.3137$$

$$\text{sd}(\theta|Y) = \sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta_1)}} = 0.06352$$

95% quantile-based confidence interval = (0.2033, 0.4510)

(b) Since $\theta \sim$ Beta(8,2), the posterior distribution is Beta(23,30),

$E(\theta|Y) = 0.4340$
$\text{mode}(\theta|Y) = 0.4314$
$\text{sd}(\theta|Y) = 0.06745$

95% quantile-based confidence interval = (0.3047, 0.5680)

(c) Based on historical or outside information, we believe there is a 75% chance that θ is near 0.2 and a 25% chance that it is near 0.8.

(d) i) $P(\theta)P(Y|\theta) = \frac{1}{4}\frac{\Gamma(10)}{\Gamma(2)\Gamma(8)}\left[3\theta(1-\theta)^7 + \theta^7(1-\theta)\right]\binom{43}{15}\theta^{15}(1-\theta)^{28}$
$\propto 3\theta^{16}(1-\theta)^{35} + \theta^{22}(1-\theta)^{29}$

ii) By part (a), the posterior distribution is a mixture of Beta(17,36) and Beta(23,30)

iii)

Table 3: Problem 3.4-d

	Part (a)	Part (b)	Part (c)
Mode	0.314	0.431	0.314

The posterior mode in part (c) is 0.314, which is close to posteior mode in part (a). The prior of part (c) put more weight on Beta(2,8), therefore, the posterior mode in part (c) will be close to that in part (a).

(e) $$P(\theta|y) = \frac{\frac{1}{4}\frac{1}{B(2,8)}\frac{1}{B(16,29)}\left(3\theta^{16}(1-\theta)^{35} + \theta^{22}(1-\theta)^{29}\right)}{P(y)}$$

where $$P(y) = \frac{1}{4}\frac{1}{B(2,8)}\frac{1}{B(16,29)}\int\left(3\theta^{16}(1-\theta)^{35} + \theta^{22}(1-\theta)^{29}\right)d\theta$$
$$= \frac{1}{4}\frac{1}{B(2,8)}\frac{1}{B(16,29)}\left(3B(17,36) + B(23,30)\right)$$

Hence,

$$P(\theta|y) = \frac{3B(17,36)\frac{1}{B(17,36)}\theta^{16}(1-\theta)^{35} + B(23,30)\frac{1}{B(23,30)}\theta^{22}(1-\theta)^{29}}{3B(17,36) + B(23,30)}$$

The weights are respectively,
$$\frac{3B(17,36)}{3B(17,36) + B(23,30)} = 0.9849 \text{ and } \frac{B(23,30)}{3B(17,36) + B(23,30)} = 0.0151.$$

3.7 a. Since $P(\theta) = 1$ and $P(y|\theta) = \binom{n}{y}\theta^y(1-\theta)^{n-y}$ the posterior distribution of θ given y is $P(\theta|y) \propto \theta^y(1-\theta)^{n-y}$, which is Beta($y+1$,$n-y+1$). Hence our posterior density is Beta$(3, 14)$.

$$E(\theta|y) = \frac{\alpha}{\alpha+\beta} = 0.1765$$

$$\text{Mode}(\theta|y) = \frac{\alpha-1}{\alpha+\beta-2} = 0.133$$

$$\text{sd}(\theta|y) = \sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}} = 0.0899$$

b. i) The conditional independence of Y_1 and Y_2 given θ is needed.

$$\begin{aligned}
P(Y_2 = y_2|Y_1 = y_1) &= \frac{\int_0^1 P(Y_1 = y_1, Y_2 = y_2, \theta)d\theta}{P(Y_1 = y_1)} \\
&= \frac{\int_0^1 P(Y_1 = y_1, Y_2 = y_2|\theta)P(\theta)d\theta}{P(Y_1 = y_1)} \\
&= \frac{\int_0^1 P(Y_1 = y_1|\theta)P(Y_2 = y_2|\theta)P(\theta)d\theta}{P(Y_1 = y_1)} \\
&= \frac{\int_0^1 P(Y_2 = y2|\theta)P(\theta|Y_1 = y_1)P(Y_1 = y_1)}{P(Y_1 = y_1)} \\
&= \int_0^1 P(Y_2 = y_2|\theta)P(\theta|Y_1 = y_1)d\theta
\end{aligned}$$

ii)

$$\begin{aligned}
&\int_0^1 P(Y_2 = y_2|\theta)P(\theta|Y_1 = 2)d\theta \\
&\quad = \int_0^1 \binom{278}{y_2}\theta^{y_2}(1-\theta)^{278-y_2}\frac{\Gamma(17)}{\Gamma(3)\Gamma(14)}\theta^2(1-\theta)^{13}d\theta
\end{aligned}$$

iii)

$$\begin{aligned}
&\int_0^1 \binom{278}{y_2}\theta^{y_2}(1-\theta)^{278-y_2}\frac{\Gamma(17)}{\Gamma(3)\Gamma(14)}\theta^2(1-\theta)^{13}d\theta \\
&\quad = \binom{278}{y_2}\frac{\Gamma(17)}{\Gamma(3)\Gamma(14)}\frac{\Gamma(y_2+3)\Gamma(292-y_2)}{\Gamma(295)}
\end{aligned}$$

c)

$$\begin{aligned}
&E(Y_2|Y_1 = 2) = \sum y_2 P(Y_2 = y_2|Y_1 = 2) = 49.06 \\
&Var(Y_2|Y_1 = 2) = \sum y_2^2 P(Y_2 = y_2|Y_1 = 2) - (E(Y_2|Y_1 = 2))^2 = 662.134 \\
&\text{sd}(Y_2|Y_1 = 2) = 25.73
\end{aligned}$$

d)

$$\begin{aligned}
&P(Y_2 = y_2|\hat{\theta}) = \binom{n}{y_2}\left(\frac{2}{15}\right)^{y_2}\left(\frac{13}{15}\right)^{n-y_2} \\
&E(Y_2|\hat{\theta}) = 272\left(\frac{2}{15}\right) = 37.067 \\
&Var(Y_2|\hat{\theta}) = 272\left(\frac{2}{15}\right)\left(\frac{13}{15}\right) = 32.1285 \\
&\text{sd}(Y_2|\hat{\theta}) = 5.668
\end{aligned}$$

By comparing the plots, $P(Y_2 = y_2|\hat{\theta})$ has smaller variation than $P(Y_2 = y_2|Y_1 = 2)$. I will choose $P(Y_2 = y_2|Y_1 = 2)$ for predicting Y_2 since it takes into account uncertainty about θ.

3.12 a)

$$\frac{\partial^2 \log P(y|\theta)}{\partial\theta^2} = -\frac{y}{\theta^2} - \frac{n-y}{(1-\theta)^2}, \quad E\left(\frac{Y}{\theta^2} + \frac{n-y}{(1-\theta)^2}\,\middle|\,\theta\right)^{\frac{1}{2}} = \sqrt{\frac{n}{\theta(1-\theta)}}$$

Hence, Jeffreys' prior is Beta$\left(\frac{1}{2}, \frac{1}{2}\right)$.

b)

$$\frac{\partial^2 \log P(y|\varphi)}{\partial \varphi^2} = -\frac{ne^{\varphi}}{(1+e^{\varphi})^2}, \quad E\left(\frac{ne^{\varphi}}{(1+e^{\varphi})^2} \,\middle|\, \varphi\right)^{\frac{1}{2}} = \frac{\sqrt{ne^{\varphi}}}{(1+e^{\varphi})}$$

Hence, Jeffreys' prior is proportional to $e^{\frac{\varphi}{2}}(1+e^{\varphi})^{-1}$.

c) We have $\varphi = \log\left(\frac{\theta}{1-\theta}\right)$, $\theta = h(\varphi) = \frac{e^{\varphi}}{1+e^{\varphi}}$ and the Jacobian is $J = \frac{e^{\varphi}}{(1+e^{\varphi})^2}$. Therefore, $P_J(\varphi) = P_\theta(h(\varphi))|J| \propto e^{\frac{\varphi}{2}}(1+e^{\varphi})^{-1}$, which is the same as in (b).