

Homework 5

STA-360-602

Total points: 10 (reproducibility) + 15 (Q1) + 25 (Q2) = 50 points.

1. (15 points, 5 points each) Hoff, 3.12 (Jeffrey's prior).
 - (a) $Y \sim \text{binomial}(n, \theta)$. Obtain Jeffreys' prior distribution $p_J(\theta)$ for this model.

Q1a Ans:

$$\begin{aligned}
I(\theta) &= -E[\partial^2 \log p(Y|\theta)/\partial\theta^2 \mid \theta] \\
&= -E\left[\frac{\partial^2 \log\left(\binom{n}{Y}\theta^Y(1-\theta)^{n-Y}\right)}{\partial\theta^2} \mid \theta\right] \\
&= -E\left[\frac{\partial^2(\log\binom{n}{Y} + Y\log(\theta) + (n-Y)\log(1-\theta))}{\partial\theta^2} \mid \theta\right] \\
&= -E\left[\frac{\partial\left(\frac{Y}{\theta} - \frac{n-Y}{1-\theta}\right)}{\partial\theta} \mid \theta\right] \\
&= -E\left[\frac{\partial\left(\frac{Y-Y\theta-n\theta+Y\theta}{\theta(1-\theta)}\right)}{\partial\theta} \mid \theta\right] \\
&= -E\left[\frac{\partial\left(\frac{Y-n\theta}{\theta(1-\theta)}\right)}{\partial\theta} \mid \theta\right] \\
&= -E[(Y-n\theta)\theta^{-1}(-1)(-1)(1-\theta)^{-2} + (Y-n\theta)(-1)\theta^{-2}(1-\theta)^{-1} + (-n)\theta^{-1}(1-\theta)^{-1} \mid \theta] \\
&= -E\left[\frac{(Y-n\theta)\theta - (Y-n\theta)(1-\theta) - n\theta(1-\theta)}{\theta^2(1-\theta)^2} \mid \theta\right] \\
&= -E\left[\frac{Y\theta - n\theta^2 - Y + Y\theta + n\theta - n\theta^2 - n\theta + n\theta^2}{\theta^2(1-\theta)^2} \mid \theta\right] \\
&= -E\left[\frac{Y\theta - n\theta^2 - Y + Y\theta + n\theta - n\theta^2 - n\theta + n\theta^2}{\theta^2(1-\theta)^2} \mid \theta\right] \\
&= E\left[\frac{n\theta^2 + Y - 2Y\theta}{\theta^2(1-\theta)^2} \mid \theta\right] \\
&= \frac{n\theta^2 + (n\theta) - 2(n\theta)\theta}{\theta^2(1-\theta)^2} \\
&= \frac{n\theta - n\theta^2}{\theta^2(1-\theta)^2} \\
&= \frac{n}{\theta(1-\theta)} \\
&\propto \frac{1}{\theta(1-\theta)}
\end{aligned}$$

$$\text{Hence, } p_J(\theta) \propto \sqrt{I(\theta)} \propto \sqrt{\frac{1}{\theta(1-\theta)}}$$

- (b) Reparameterize with $\psi = \log[\theta/(1-\theta)]$ so that $p(y|\psi) = \binom{n}{y} e^{\psi y} (1+e^\psi)^{-n}$.
Obtain Jeffreys' prior distribution $p_J(\psi)$ for this model.

Q1b Ans:

$$\begin{aligned}
I(\psi) &= -E[\partial^2 \log p(y|\psi) / \partial \psi^2 \mid \psi] \\
&= -E \left[\frac{\partial^2 \log(\binom{n}{y} e^{\psi y} (1+e^\psi)^{-n})}{\partial \psi^2} \mid \psi \right] \\
&= -E \left[\frac{\partial^2 (\log \binom{n}{y} + \psi y - n \log(1+e^\psi))}{\partial \psi^2} \mid \psi \right] \\
&= -E \left[\frac{\partial (y - n(1+e^\psi)^{-1} e^\psi)}{\partial \psi} \mid \psi \right] \\
&= -E [-n(-1)(1+e^\psi)^{-2} e^\psi e^\psi - n(1+e^\psi)^{-1} e^\psi \mid \psi] \\
&= nE \left[\frac{e^\psi}{(1+e^\psi)} - \frac{e^{2\psi}}{(1+e^\psi)^2} \mid \psi \right] \\
&= nE \left[\frac{e^\psi}{(1+e^\psi)^2} \mid \psi \right] \\
&\propto \frac{e^\psi}{(1+e^\psi)^2}
\end{aligned}$$

$$\text{Hence, } p_J(\psi) \propto \sqrt{I(\psi)} \propto \sqrt{\frac{e^\psi}{(1+e^\psi)^2}}$$

- (c) Take the prior distribution from a) and apply the change of variables formula from Exercise 3.10 to obtain the induced prior density on ψ . This density should be the same as the one derived in part b) of this exercise. This consistency under reparameterization is the defining characteristic of Jeffrey's' prior.

Q1c Ans:

Since $\psi = \log[\theta/(1 - \theta)]$,

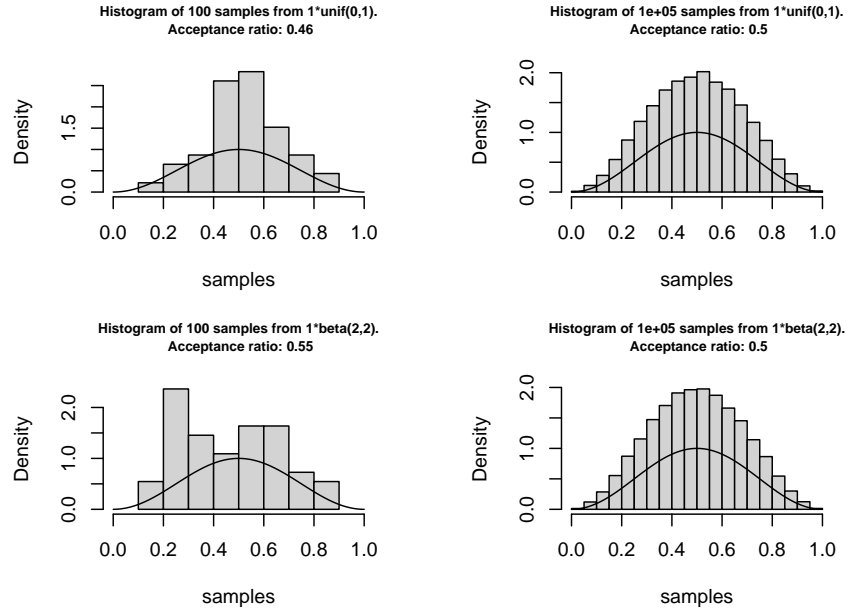
$$\theta = \frac{e^\psi}{1 + e^\psi} = h(\psi)$$

Based on 3.10,

$$\begin{aligned} p_{J,\psi}(\psi) &= p_{J,\theta}(\theta) = p_{J,\theta}(h(\psi)) \times \left| \frac{dh}{d\psi} \right| \\ &\propto \sqrt{\frac{1}{h(\psi)(1-h(\psi))}} \times |e^\psi(1+e^\psi)^{-1} + e^\psi(-1)(1+e^\psi)^{-2}e^\psi| \\ &= \sqrt{\frac{1}{\frac{e^\psi}{1+e^\psi}(1-\frac{e^\psi}{1+e^\psi})}} \times \left| \frac{e^\psi + e^{2\psi} - e^{2\psi}}{(1+e^\psi)^2} \right| \\ &= \sqrt{\frac{1}{\frac{e^\psi}{(1+e^\psi)^2}}} \times \left| \frac{e^\psi}{(1+e^\psi)^2} \right| \\ &= \sqrt{\frac{e^\psi}{(1+e^\psi)^2}} \end{aligned}$$

2. *Lab component* (25 points total) Please refer to lab 5 and complete tasks 4–5.

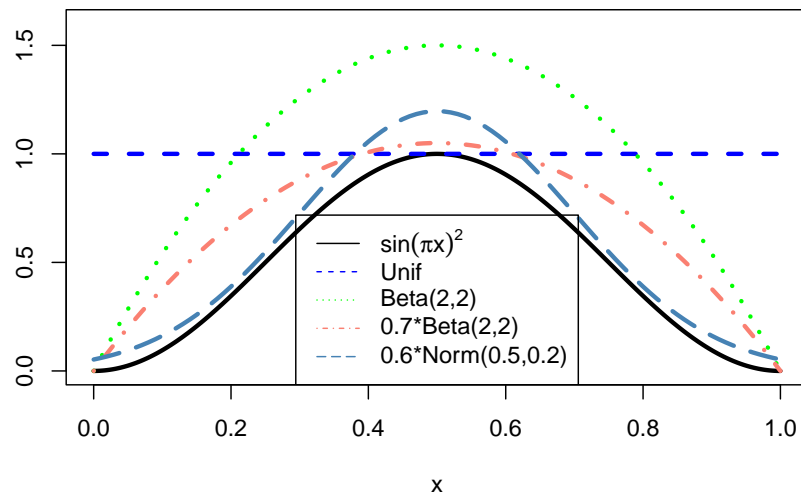
(a) (10) Task 4



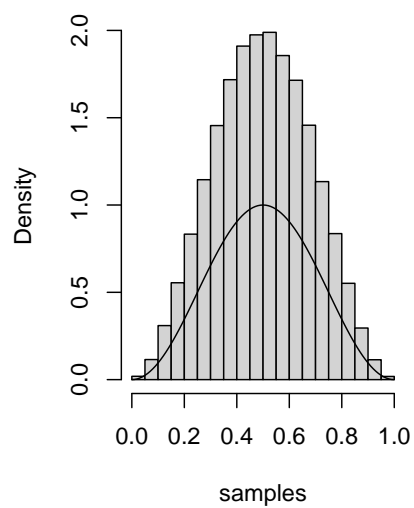
- With only 100 samples, it appeared that the Beta(2,2) enveloping proposal worked slightly better than Unif(0,1) with a slightly higher acceptance ratio of 0.55, in comparison to 0.46. However, Beta(2,2) generated an approximation that is skewed, unlike what was generated using Unif(0,1) and the actual $f(x)$.
- The above mentioned differences in acceptance ratios and skewness went away when the number of simulations was increased to 100,000 and both enveloping proposals generated essentially the same histogram which approximates $f(x)$ with the same acceptance ratio of 0.5.
- In all cases, the histograms are higher than $f(x)$ because $f(x)$ is not a normalized density while the histograms approximates the normalized density.

(b) (15) Task 5

- (i) They were about the same in terms of acceptance ratio when the number of simulations is sufficiently high (10^5), so neither is better in that aspect. In terms of speed, however, it may be more efficient to use the uniform density function rather than the beta density function since computing the probability density is easier with a uniform distribution.
- (ii)
 - To get higher acceptance ratio, I tried two new proposal distributions
 - The first one is $0.7 \cdot \text{Beta}(2,2)$ and the second one is $0.6 \cdot \text{Norm}(0.5, 0.2)$ (truncated to focus on $0 \leq x \leq 1$).
 - These choices were made such that the the density functions are closer to $f(x)$, i.e., the area under the enveloping proposals and above $f(x)$ is decreased.
 - Both new enveloping proposals are shown on the first figure, as well as the old ones.
 - The acceptance ratios of the new proposals are 0.72 and 0.84, indeed higher than the previous values of 0.5.



Histogram of 1e+05 samples from $0.7 \cdot \text{beta}(2,2)$.
Acceptance ratio: 0.72



Histogram of 1e+05 samples from $0.6 \cdot \text{norm}(0.5, 0.1)$.
Acceptance ratio: 0.84

