

Homework 5

STA 360: Assignment 5, Fall 2020

Due Friday September 18, 5 PM Standard Eastern Time

Hoff Exercise 3.12

3.12 a)

$$Y \sim \text{binomial}(n, \theta) \rightarrow \binom{n}{y} \theta^y (1-\theta)^{n-y}$$

$$\log(p(y|\theta)) \rightarrow \log \binom{n}{y} + y \log(\theta) + (n-y) \log(1-\theta)$$

$$\frac{\partial \log(p(y|\theta))}{\partial \theta} = y * \frac{1}{\theta} + (n-y) * \frac{1}{1-\theta} * (-1)$$

$$= \frac{y}{\theta} - \frac{n-y}{1-\theta}$$

$$\frac{\partial^2 \log(p(y|\theta))}{\partial \theta^2} = -\frac{y}{\theta^2} - \frac{n-y}{(1-\theta)^2}$$

$$I(\theta) = -E \left[\frac{-y}{\theta^2} - \frac{n-y}{(1-\theta)^2} \mid \theta \right]$$

$E[y|\theta] = n\theta$ so $I(\theta)$ becomes

$$= -\frac{1}{\theta^2} E[y|\theta] - \frac{1}{(1-\theta)^2} E[n-y|\theta]$$

$$\frac{-n\theta}{\theta^2} - \frac{n(1-\theta)}{(1-\theta)^2} \rightarrow \text{simplifies to } \frac{-n}{\theta} - \frac{n}{(1-\theta)} \rightarrow \frac{-n}{\theta(1-\theta)}$$

$$I(\theta) = -\frac{-n}{\theta(1-\theta)} = \frac{n}{\theta(1-\theta)} \quad J_3(\theta) = \sqrt{\frac{n}{\theta(1-\theta)}}$$

$$\sqrt{\frac{n}{\theta(1-\theta)}} = n^{1/2} \theta^{-1/2} (1-\theta)^{-1/2}$$

$$\text{constant} \downarrow \theta^{-1/2+1-1} (1-\theta)^{-1/2+1-1} = \text{Beta}(-1/2+1, -1/2+1)$$

$$= \text{Beta}(1/2, 1/2)$$

a.

$$b) P(y|\psi) = \binom{n}{y} e^{\psi y} (1+e^\psi)^{-n}$$

$$\log(P(y|\psi)) = \log\binom{n}{y} + \psi y - n \log(1+e^\psi)$$

$$\frac{\partial \log(P(y|\psi))}{\partial \psi} = y - \frac{ne^\psi}{(1+e^\psi)}$$

$$\frac{\partial^2 \log(P(y|\psi))}{\partial \psi^2} = \frac{\partial}{\partial \psi} \left(-\frac{ne^\psi}{(1+e^\psi)} \right) \xrightarrow{\text{product rule}} -n \left(\frac{e^\psi(1+e^\psi) - e^\psi e^\psi}{(1+e^\psi)^2} \right)$$

simplify = $\frac{-ne^\psi}{(1+e^\psi)^2}$

$$I(\psi) = -E \left[\frac{-ne^\psi}{(1+e^\psi)^2} \middle| \psi \right] \xrightarrow{\text{condition on } \psi} \frac{ne^\psi}{(1+e^\psi)^2}$$

$$P_J(\psi) = \sqrt{\frac{ne^\psi}{(1+e^\psi)^2}} = \frac{n^{1/2} e^{1/2\psi}}{1+e^\psi} \propto \psi = \boxed{\frac{e^{1/2\psi}}{1+e^\psi}}$$

b.

c.
 $\theta \sim \text{Beta}(1/2, 1/2), \psi = \log\left(\frac{\theta}{1-\theta}\right)$

$$e^\psi = \frac{\theta}{1-\theta} \text{ solve for } \theta \rightarrow \theta = \frac{e^\psi}{1+e^\psi} = h(\psi)$$

So

$$p_\theta\left(\frac{e^\psi}{1+e^\psi}\right) = \frac{\Gamma(1)}{\Gamma(1/2)\Gamma(1/2)} \left(\frac{e^\psi}{1+e^\psi}\right)^{1/2} \left(1 - \frac{e^\psi}{1+e^\psi}\right)^{1/2}$$

$$\frac{dh(\psi)}{d\psi} = \frac{e^\psi(1+e^\psi) - e^{2\psi}}{(1+e^\psi)^2} \rightarrow \frac{e^\psi}{(1+e^\psi)^2}$$

Thus

$$p_\psi(\psi) = \frac{\Gamma(1)}{\Gamma(1/2)\Gamma(1/2)} \left(\frac{e^\psi}{1+e^\psi}\right)^{1/2} \left(\frac{1}{1+e^\psi}\right)^{1/2} * \frac{e^\psi}{(1+e^\psi)^2}$$

Simplify

$$\frac{(1+e^\psi)^{1/2} (1+e^\psi)^{1/2}}{e^{1/2} \cdot 1} \cdot \frac{e^\psi}{(1+e^\psi)^2} \rightarrow \frac{e^{1/2\psi}}{1+e^\psi}$$

$$\frac{\Gamma(1)}{\Gamma(1/2)\Gamma(1/2)} \frac{e^{1/2\psi}}{1+e^\psi} \propto \frac{e^{1/2\psi}}{1+e^\psi}$$

Constants

So

$$p_\psi(\psi) = \frac{e^{1/2\psi}}{1+e^\psi} \text{ Same as part b}$$

c.

Lab Component

a. Task 4

#Code taken from Professor Steort's Lab 5 RMD file

grid of points

```
x <- seq(0, 1, 10^-2)
```

```
fx <- function(x) sin(pi * x)^2
```

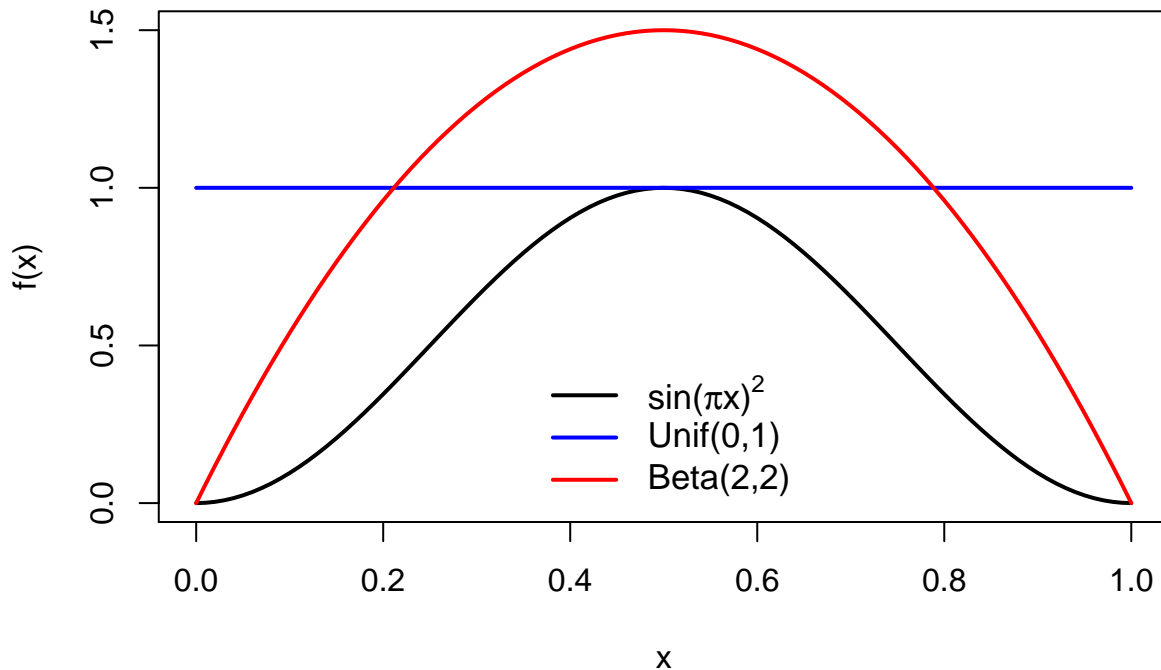
```
plot(fx, xlim = c(0,1), ylim = c(0,1.5), ylab = "f(x)", lwd = 2)
```

```
curve(dunif, add = TRUE, col = "blue", lwd = 2)
```

```
curve(dbeta(x,2,2), add = TRUE, col = "red", lwd = 2)
```

```
legend("bottom", legend = c(expression(paste("sin(", pi, "x)^2")), "Unif(0,1)",
```

```
"Beta(2,2)", col = c("black", "blue", "red"), lty = c(1,1,1), bty = "n", cex = 1.1, lwd = 2)
```

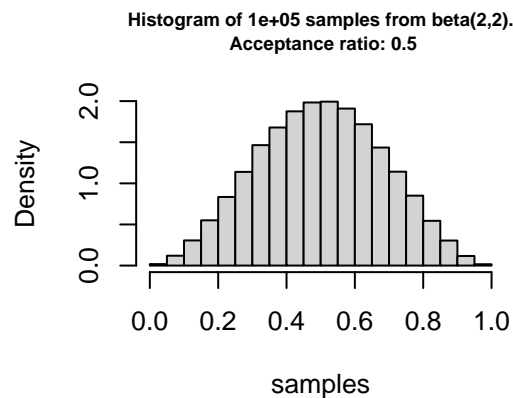
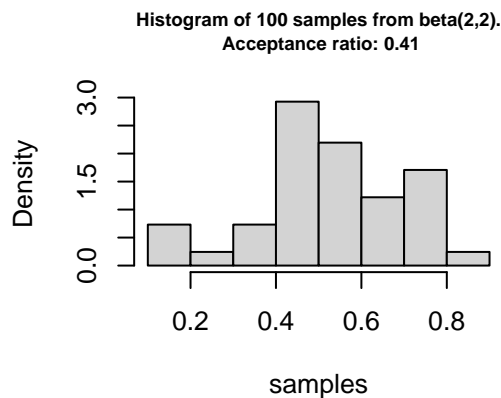
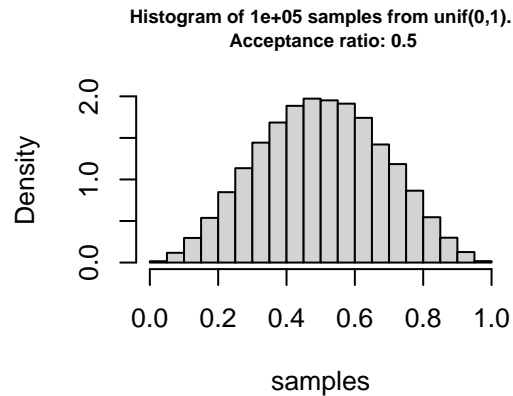
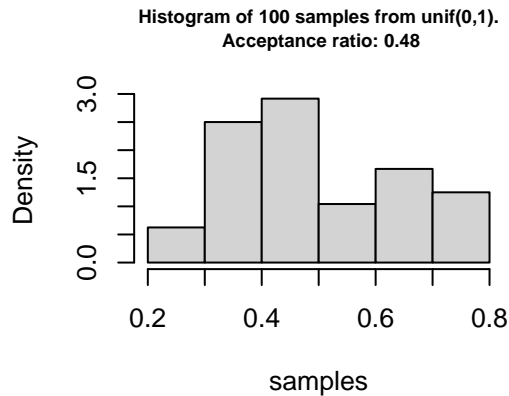


```
sim_fun <- function(f, envelope = "unif", par1 = 0, par2 = 1, n = 10^2, plot = TRUE){
  r_envelope <- match.fun(paste0("r", envelope))
  d_envelope <- match.fun(paste0("d", envelope))
  proposal <- r_envelope(n, par1, par2)
  if(envelope == "norm"){ #if normal, which is the better envelope function
    #multiply dnorm by the constant I chose
    density_ratio <- f(proposal) / (0.605* d_envelope(proposal, par1, par2))
  }else{
    #if not normal (uniform or beta), no constant is needed
    density_ratio <- f(proposal) / d_envelope(proposal, par1, par2)
  }
  samples <- proposal[runif(n) < density_ratio]
  acceptance_ratio <- length(samples) / n
  if (plot) {
    hist(samples, probability = TRUE,
          main = paste0("Histogram of ",
                        n, " samples from ",
                        envelope, "(", par1, ", ", par2,
                        ").\n Acceptance ratio: ",
                        round(acceptance_ratio,2)),
          cex.main = 0.75)
  }
  list(x = samples, acceptance_ratio = acceptance_ratio)
}
```

```
par(mfrow = c(2,2), mar = rep(4, 4))
```

```
unif_1 <- sim_fun(fx, envelope = "unif", par1 = 0, par2 = 1, n = 10^2)
unif_2 <- sim_fun(fx, envelope = "unif", par1 = 0, par2 = 1, n = 10^5)
```

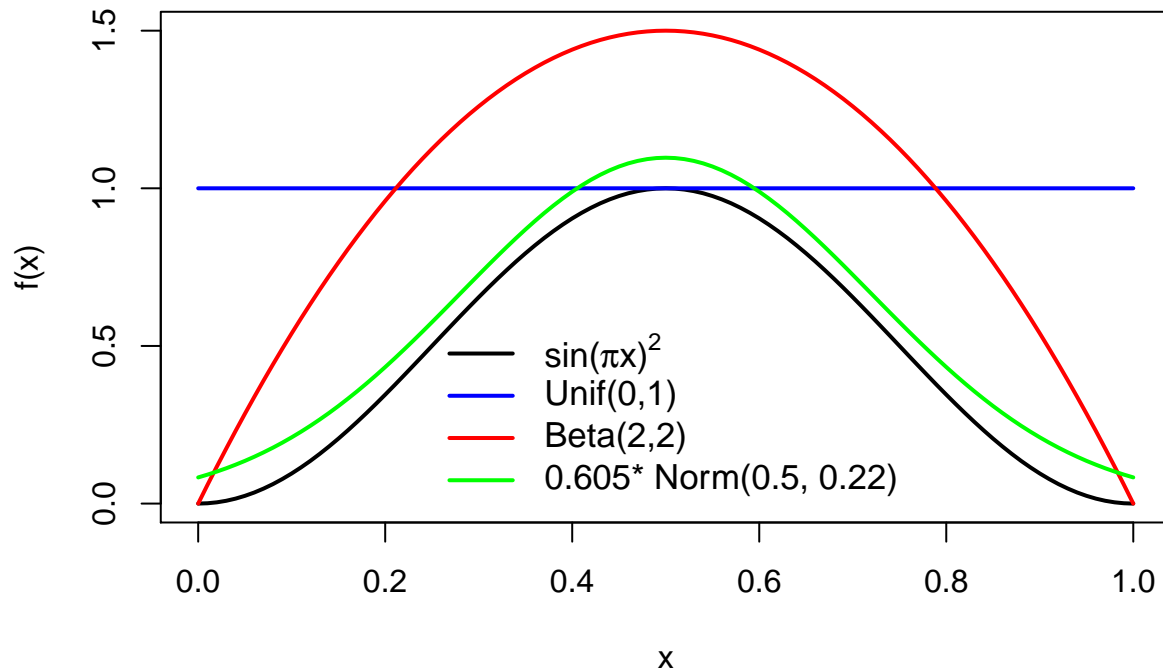
```
# Add in the Beta(2,2) Histograms with different amount of samples
beta_1 <- sim_fun(fx, envelop = "beta", par1 = 2, par2 = 2, n = 10^2)
beta_2 <- sim_fun(fx, envelop = "beta", par1 = 2, par2 = 2, n = 10^5)
```



From the histograms above, we can see that the more samples taken, the distribution of the accepted values looks more normal with most commonly accepted value being around 0.5. There is more variance with the 100 sample histograms with much more variance in the most commonly accepted value being either a bit higher than 0.5 or a bit lower. The acceptance ratio for the 100,000 samples for both the uniform and beta is 0.5 while the acceptance ratio of the 100 samples vary more with the uniform having a slightly higher or lower ratio each time you run the code above. However, even with more variance, the acceptance ratio of the 100 samples tends to hover in the range of 0.4-0.5.

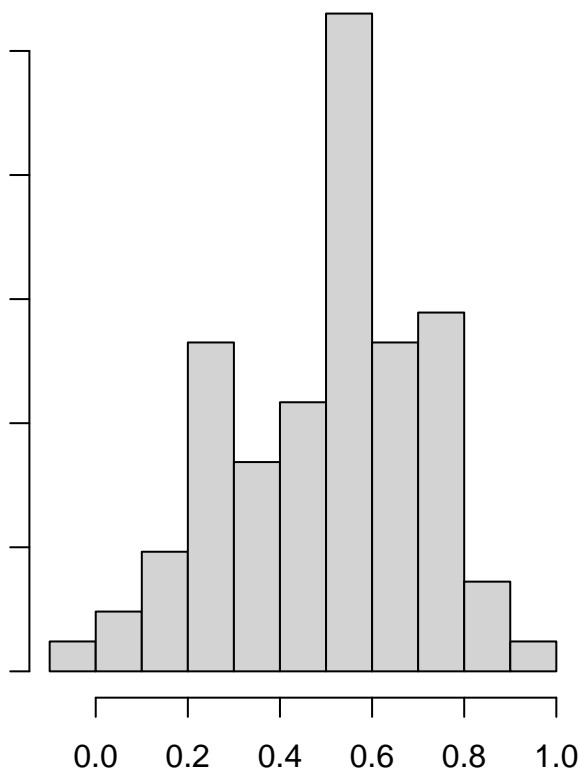
b. Task 5

```
#Plot new graph with the normal distribution, beta, and uniform
plot(fx, xlim = c(0,1), ylim = c(0,1.5), ylab = "f(x)", lwd = 2)
curve(dunif, add = TRUE, col = "blue", lwd = 2)
curve(dbeta(x,2,2), add = TRUE, col = "red", lwd = 2)
curve(0.605*dnorm(x, 0.5, 0.22), add = TRUE, col = "green", lwd = 2) #add in normal
legend("bottom", legend = c(expression(paste("sin(", pi, "x)"^2)), "Unif(0,1)",
"Beta(2,2)", "0.605* Norm(0.5, 0.22)"), col = c("black", "blue", "red", "green"),
lty = c(1,1,1), bty = "n", cex = 1.1, lwd = 2)
```

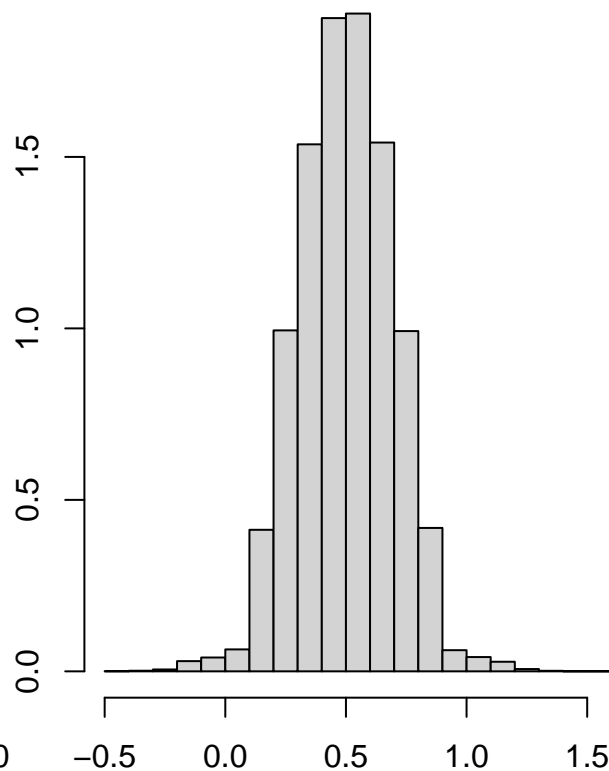


```
par(mfrow = c(1,2), mar = c(2,1,2,1))
#Adding in the normal distribution histograms
norm_1 <- sim_fun(fx, envelop = "norm", par1 = .5, par2 = .22, n = 10^2)
norm_2 <- sim_fun(fx, envelop = "norm", par1 = .5, par2 = .22, n = 10^5)
```

Histogram of 100 samples from norm(0.5,0.22).
Acceptance ratio: 0.83



Histogram of 1e+05 samples from norm(0.5,0.22).
Acceptance ratio: 0.84



As discussed in task 4, the two functions we chose are about the same when it comes to enveloping $f(x)$. An

acceptance ratio of 0.5 is not great and can be costly if you were to run many more samples as only half of the values will be accepted. So, a better enveloping function that I found is a normal function with mean 0.5, standard deviation of 0.22 and multiplied by a constant value of 0.605. With that density, we can see that the acceptance ratio of the 100,000 samples is around 0.84 and with the 100 samples showing, again, more variance, with values (that I've seen from running the code multiple times) from 0.73 to 0.89 which is still higher than both the beta and uniform densities. This is much better because you won't have to reject as many values since you can accept around 84% of the values with the $0.605 * \text{Normal}(0.5, 0.22)$ density.