

MATH 392 Problem Set 4

EJ Arce

23 February 2018

6.2

Analytical Solution

Let $x_1, \dots, x_n \sim \text{Poisson}(\lambda)$. Show the maximum likelihood estimate of λ is $\hat{\lambda} = \bar{x}$:

$$L(\lambda|x_1, \dots, x_n) = P(X_1 = x_1, \dots, X_n = x_n)$$

$$L(\lambda|x_1, \dots, x_n) = P(X = x_1) \dots P(X = x_n)$$

$$L(\lambda|x_1, \dots, x_n) = \prod_i^n \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

$$L(\lambda|x_1, \dots, x_n) = \frac{e^{-\lambda n} \lambda^{\sum_i^n x_i}}{\prod_i^n x_i!}$$

Take the natural log of both sides to make the derivation easier:

$$\ln(L(\lambda|x_1, \dots, x_n)) = \ln(e^{-n\lambda}) + \ln(\lambda^{\sum_i^n x_i}) - \ln(\prod_i^n x_i!)$$

$$\ln(L(\lambda|x_1, \dots, x_n)) = -n\lambda + \ln(\lambda) \sum_i^n x_i - \ln(\prod_i^n x_i!)$$

Derive with respect to λ and set equal to 0:

$$\frac{\partial \ln(L(\lambda|x_1, \dots, x_n))}{\partial \lambda} = -n + \frac{\sum_i^n x_i}{\lambda} = 0$$

$$\frac{\partial \ln(L(\lambda|x_1, \dots, x_n))}{\partial \lambda} = \lambda = \frac{\sum_i^n x_i}{n} = \bar{x}$$

Thus $\lambda_{MLE} = \bar{x}$

Empirical Solution

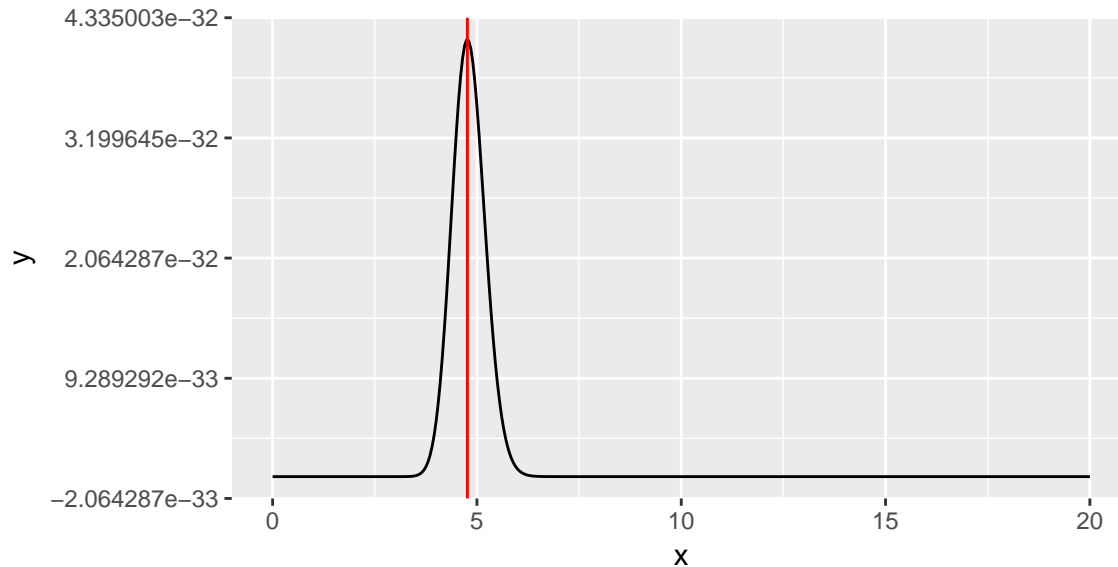
```

# Build function
e <- exp(1)
pois <- function(lambda, x){
  n <- length(x)
  sum.x <- sum(x)
  prod.fact <- prod(factorial(x))
  e^(-lambda*n)*(lambda^(sum.x))/(prod.fact)
}

# Create sequence of lambdas for MLE plot
lambdas <- seq(from = 0, to = 20, by = .05)
# Draw random sample of x's
x <- sample(1:10,30,replace=T)
# Apply pois function to find likelihoods
L_pois <- pois(lambdas,x)

# Plot
df <- data.frame(x = lambdas,
                 y = L_pois)
ggplot(df, aes(x=x,y=y)) +
  geom_line() +
  geom_vline(xintercept = mean(x), col = "red")

```



A red line above indicates the mean of the random sample. As expected, the highest likelihood for λ is at the sample mean.

6.8

Analytical Solution

$$f(x; \theta) = \frac{\sqrt{2/\pi} x^2 e^{-x^2/2\theta^2}}{\theta^3}$$

$$L(\theta|x_1, \dots, x_n) = \prod_i^n f(x_i; \theta)$$

$$L(\theta|x_1, \dots, x_n) = \prod_i^n \frac{\sqrt{2/\pi} x_i^2 e^{-x_i^2/2\theta^2}}{\theta^3}$$

$$L(\theta|x_1, \dots, x_n) = (\frac{\sqrt{2/\pi}}{\theta^3})^n \prod_i^n x_i^2 e^{-x_i^2/2\theta^2}$$

$$\ln(L(\theta|x_1, \dots, x_n)) = \ln((\frac{\sqrt{2/\pi}}{\theta^3})^n) + \sum_i^n \ln(x_i^2) + \sum_i^n \ln(e^{-x_i^2/2\theta^2})$$

$$\ln(L(\theta|x_1, \dots, x_n)) = n(\ln\sqrt{2/\pi} - \ln\theta^3) + \sum_i^n \ln(x_i^2) - \frac{1}{2\theta^2} \sum_i^n x_i^2$$

Derive with respect to θ and set to 0:

$$\frac{\partial(\ln L(\theta|x_1, \dots, x_n))}{\partial\theta} = -\frac{n}{\theta^3} 3\theta^2 + \frac{1}{\theta^3} \sum_i^n x_i^2 = 0$$

$$\frac{3n}{\theta} = \frac{\sum_i^n x_i^2}{\theta^3}$$

$$\theta = \sqrt{\frac{\sum_i^n x_i^2}{3n}}$$

Thus $\theta_{MLE} = \frac{\bar{x}\sqrt{n}}{\sqrt{3}}$.

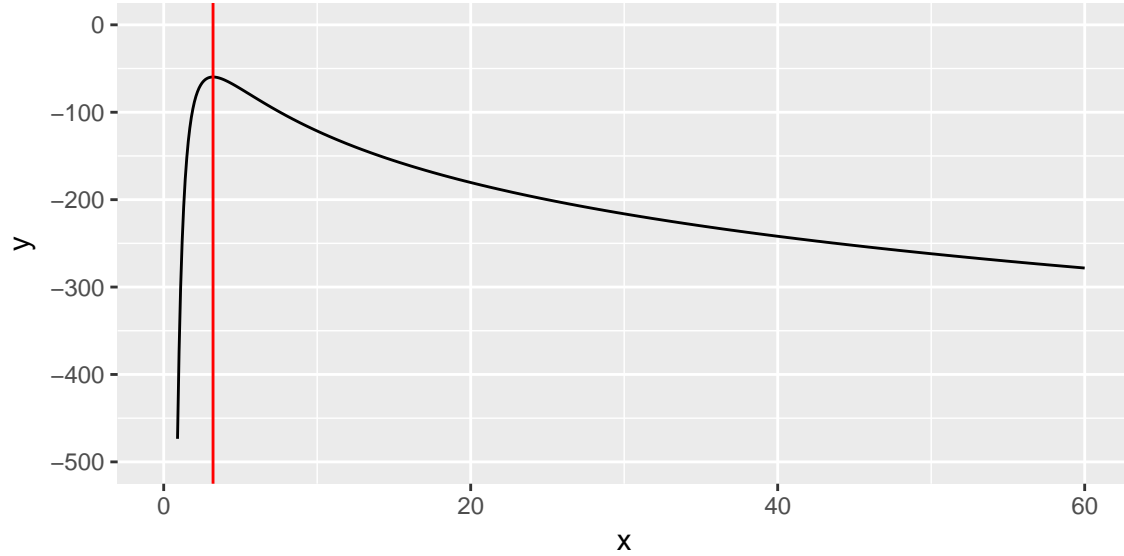
Empirical Solution

```
# Build log function
f <- function(theta, x){
  n <- length(x)
  n*(log(sqrt(2/pi)) - log(theta^3)) +
    sum(log(x^2)) -
    sum((x^2))/(2*(theta^2))
}

# Create sequence of thetas for MLE plot
thetas <- seq(from = .1, to = 60, by = .1)
# Reset seed and draw random sample of x's
set.seed(23)
x <- rnorm(30, 5, 2)
# Apply function to find likelihoods
L_theta <- f(thetas, x)
# Find x-intercept from analytical solution
n <- length(x)
analytical.sol <- sqrt(sum(x^2)/(3*n))
```

```
# Plot
df <- data.frame(x = thetas,
                 y = L_theta)
ggplot(df, aes(x=x,y=y)) +
  geom_line() +
  geom_vline(xintercept = analytical.sol, col = "red") +
  xlim(0,60) +
  ylim(-500,0)
```

```
## Warning: Removed 8 rows containing missing values (geom_path).
```



As expected, the maximum likelihood estimate of our simulation is at about the analytical solution.

6.11

Analytical Solution

$$L(\lambda) = \prod_i^n f(x_i) \prod_j^m f(y_j)$$

$$L(\lambda) = \prod_i^n \lambda e^{-\lambda x_i} \prod_j^m 2\lambda e^{-2\lambda y_j}$$

$$L(\lambda) = 2^m \lambda^{n+m} \lambda e^{-\lambda(\sum_i^n x_i + 2\sum_j^m y_j)}$$

$$\ln(L(\lambda)) = (n+m)\ln(\lambda) - \lambda(\sum_i^n x_i + 2\sum_j^m y_j)$$

$$\frac{\partial \ln(L(\lambda))}{\partial \lambda} = (n+m)/\lambda - (\sum_i^n x_i + 2\sum_j^m y_j) = 0$$

$$\lambda = \frac{n+m}{\sum_i^n x_i + 2 \sum_j^m y_j}$$

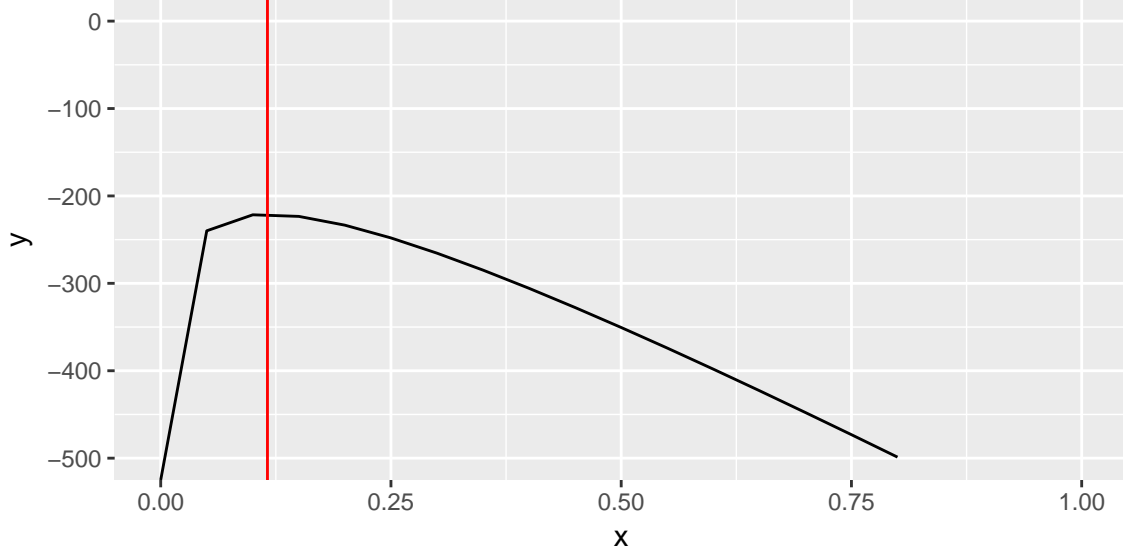
Thus, $\lambda_{MLE} = \frac{n+m}{\sum_i^n x_i + 2 \sum_j^m y_j}$.

Empirical Solution

```
# Build log function
pois <- function(lambda,x,y){
  n <- length(x)
  m <- length(y)
  (n+m)*log(lambda) - lambda*(sum(x) + 2*sum(y))
}

# Create sequence of lambdas for MLE plot
lambdas <- seq(from = 0, to = 20, by = .05)
# Draw random sample of x's
set.seed(11)
x <- sample(1:10,30,replace=T)
set.seed(35)
y <- sample(1:10,40,replace=T)
# Apply pois function to find likelihoods
L_pois <- pois(lambdas,x,y)
# Calculate analytical solution
obs <- (length(x)+length(y))/(sum(x) + 2*sum(y))
# Plot
df <- data.frame(x = lambdas,
                 y = L_pois)
ggplot(df, aes(x=x,y=y)) +
  geom_line() +
  geom_vline(xintercept = obs, col = "red") +
  xlim(0,1) +
  ylim(-500,0)
```

```
## Warning: Removed 384 rows containing missing values (geom_path).
```



The analytical solution (the red line) is at around the maximum of the log likelihood function.

6.13

a

$$L(\alpha|X_1, \dots, X_n; \beta) = \prod_i^n \alpha \beta X_i^{\beta-1} e^{-\alpha X_i^\beta} = (\alpha \beta)^n \prod_i^n X_i^{\beta-1} e^{-\alpha X_i^\beta}$$

$$\ln(L(\alpha)) = n \ln \alpha + n \ln \beta + (\beta - 1) \sum_i^n \ln X_i + \sum_i^n \ln(e^{-\alpha X_i^\beta})$$

$$\ln(L(\alpha|X; \beta)) = n \ln \alpha + n \ln \beta + (\beta - 1) \sum_i^n \ln X_i - \alpha \sum_i^n X_i^\beta$$

$$\frac{\partial \ln(L(\alpha|X; \beta))}{\partial \alpha} = \frac{n}{\alpha} - \sum_i^n X_i^\beta = 0$$

Thus $\alpha_{MLE} = \frac{1}{n} \sum_i^n X_i^\beta$.

b

$$L(\alpha; \beta|X_1, \dots, X_n) = (\alpha \beta)^n \prod_i^n X_i^{\beta-1} e^{-\alpha X_i^\beta}$$

Take the natural log of both sides:

$$\ln(L(\alpha; \beta|X)) = n \ln \alpha + n \ln \beta + (\beta - 1) \sum_i^n \ln X_i - \alpha \sum_i^n X_i^\beta$$

Differentiate with respect to α and β and set to 0. Solve simultaneously for α and β :

$$\frac{\partial \ln(L(\alpha; \beta|X))}{\partial \alpha} = \frac{n}{\alpha} - \sum_i^n X_i^\beta = 0$$

and

$$\frac{\partial \ln(L(\alpha; \beta|X))}{\partial \beta} = \frac{n}{\beta} + \ln\left(\sum_i^n X_i\right) - \alpha \sum_i^n X_i^\beta = 0$$

6.14

Let $X = 2, 3, 5, 9, 10 \sim \text{Unif}[\alpha, \beta]$. Find $\hat{\alpha}_{MOM}$ and $\hat{\beta}_{MOM}$:

1. Write first and second moments in terms of parameters α and β

$$\mu_1 = E[X] = \int_\alpha^\beta \frac{1}{\beta - \alpha} x dx = \frac{1}{2(\beta - \alpha)} x^2 \Big|_\alpha^\beta = \frac{\alpha + \beta}{2} = 5.8 = \bar{X}$$

Thus

$$\alpha + \beta = 11.6$$

Now consider the second moment:

$$\mu_2 = E[X^2] = \int_\alpha^\beta \frac{1}{\beta - \alpha} x^2 dx =$$

Solving the integral and setting μ_2 equal to its theoretical value gives

$$\mu_2 = \frac{\alpha^2 + \alpha\beta + \beta^2}{3} = \frac{\sum_i^5 X_i^2}{5} = 43.8$$

$$\mu_2 = 131.4 = \alpha^2 + \alpha\beta + \beta^2$$

Since $\alpha + \beta = 11.6$, solve for β and plug into the equation for μ_2 :

$$\beta = 11.6 - \alpha$$

$$131.4 = \alpha^2 + \alpha(11.6 - \alpha) + (11.6 - \alpha)^2$$

$$0 = 11.6\alpha + 134.6 - 23.2\alpha + \alpha^2 - 131.4$$

$$0 = \alpha^2 - 11.6\alpha + 4.2$$

Using the quadratic formula, solving for α gives

$$\alpha = .27 \text{ or } 11.32$$

Given the dataset, $\alpha = .27$ makes more sense, since $X_{max} = 10$. Now plug α in to solve for β :

$$.27 + \beta = 11.6$$

$$\beta = 11.33$$

Thus our method of moment estimates are $\hat{\alpha}_{MOM} = .27$, $\hat{\beta}_{MOM} = 11.33$.

6.20

a

$$L(\theta) = \prod_{i=1}^5 f(x_i|\theta) = \prod_{i=1}^5 \theta x_i^{\theta-1}$$

$$L(\theta) = \theta \left(\prod_{i=1}^5 x_i \right)^{\theta-1}$$

Differentiate, set equal to 0, and solve for θ :

$$\frac{\partial L(\theta)}{\partial \theta} = 5\theta^4 \left(\prod_{i=1}^5 x_i \right)^{\theta-1} + \theta^5 \left(\prod_{i=1}^5 x_i \right)^{\theta-1} \ln \left(\prod_{i=1}^5 x_i \right) = 0$$

$$\theta^4 \left(\prod_{i=1}^5 x_i \right)^{\theta-1} (5 + \theta \ln \left(\prod_{i=1}^5 x_i \right)) = 0$$

$$\theta = \frac{-5}{\ln \left(\prod_{i=1}^5 x_i \right)} \approx 1.57$$

Thus $\theta_{MLE} \approx 1.57$.

b

$$\mu_1 = E[X] = \int_0^1 \theta x^{\theta-1} x dx = \int_0^1 \theta x^{\theta} dx$$

$$\mu_1 = \frac{\theta}{\theta+1} x^{\theta+1} \Big|_0^1 = \frac{\theta}{\theta+1}$$

$$\hat{\mu}_1 = \frac{\theta}{\theta+1} = \bar{X}$$

Rearranging gives:

$$\theta = \frac{\bar{X}}{1 - \bar{X}} = \frac{.594}{1 - .594} = 1.463$$

Thus $\theta_{MOM} \approx 1.463$.

6.25

$$\forall i \in \{1, \dots, n\}, \sum_i a_i = 1$$

6.27

a

$$\text{bias} = E[\hat{\sigma}^2] - \sigma^2$$

$$E[\hat{\sigma}^2] = E\left[\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2\right] = \frac{1}{n} E\left[\sum_{i=1}^n X_i^2 - n\bar{X}^2\right]$$

$$E[\hat{\sigma}^2] = \frac{1}{n} \left(\sum_{i=1}^n E[X_i^2] - E[n\bar{X}^2] \right)$$

$$E[\hat{\sigma}^2] = \frac{n-1}{n} \sigma^2$$

$$\text{Thus bias} = \frac{n-1}{n} \sigma^2 - \sigma^2 = -\frac{\sigma^2}{n}$$

b

From theorem B.16, $\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$, where $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$. Also, from theorem B.12, $V\left[\frac{(n-1)S^2}{\sigma^2}\right] = 2(n-1)$. Thus

$$\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} \sim \chi_{n-1}^2$$

Replace with $\hat{\sigma}^2$ to get

$$\frac{\hat{\sigma}^2 n}{\sigma^2} \sim \chi_{n-1}^2$$

Thus

$$V\left[\frac{\hat{\sigma}^2 n}{\sigma^2}\right] = 2(n-1)$$

$$V[\hat{\sigma}^2] = \frac{2(n-1)\sigma^4}{n^2}$$

c

$$MSE[\hat{\sigma}^2] = bias[\hat{\sigma}^2]^2 + V[\hat{\sigma}^2]$$

$$MSE[\hat{\sigma}^2] = bias[\hat{\sigma}^2]^2 + V[\hat{\sigma}^2]$$

$$MSE[\hat{\sigma}^2] = \frac{\sigma^4}{n^2} + \frac{2(n-1)\sigma^4}{n^2}$$

$$MSE[\hat{\sigma}^2] = \frac{2n - \sigma^4}{n^2}$$

6.30

a

$$MSE_{\hat{\theta}_1} = bias_{\hat{\theta}_1}^2 + Var[\hat{\theta}_1] = 0 + 25 = 25$$

$$MSE_{\hat{\theta}_2} = bias_{\hat{\theta}_2}^2 + Var[\hat{\theta}_2] = 3^2 + 4 = 13$$

Thus $MSE_{\hat{\theta}_2} \leq MSE_{\hat{\theta}_1}$.

b

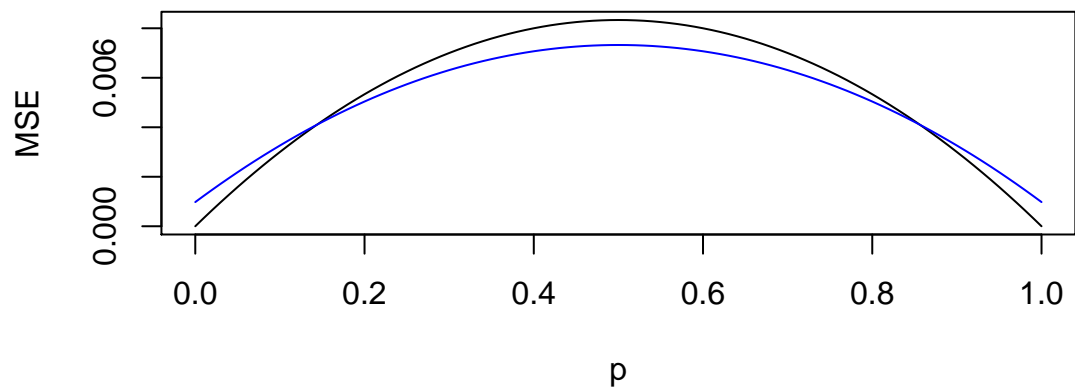
$$MSE_{\hat{\theta}_2} \leq MSE_{\hat{\theta}_1}$$

$$b^2 + 4 \leq 25$$

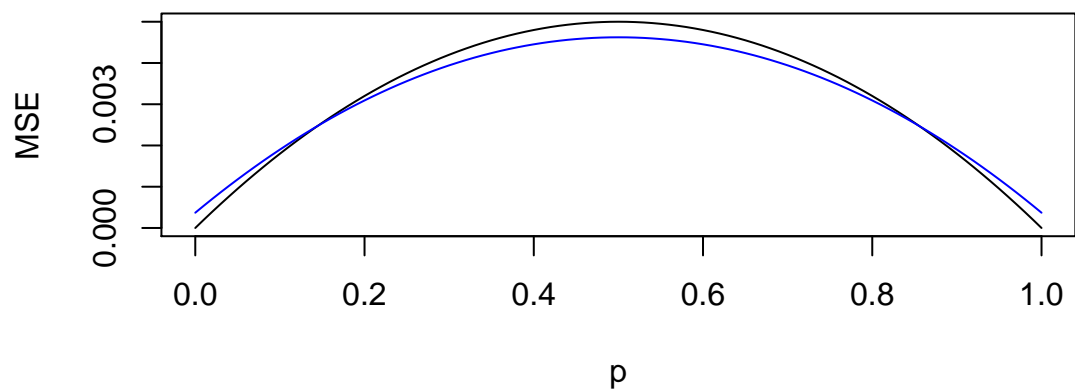
$$b \leq 4.583$$

6.31

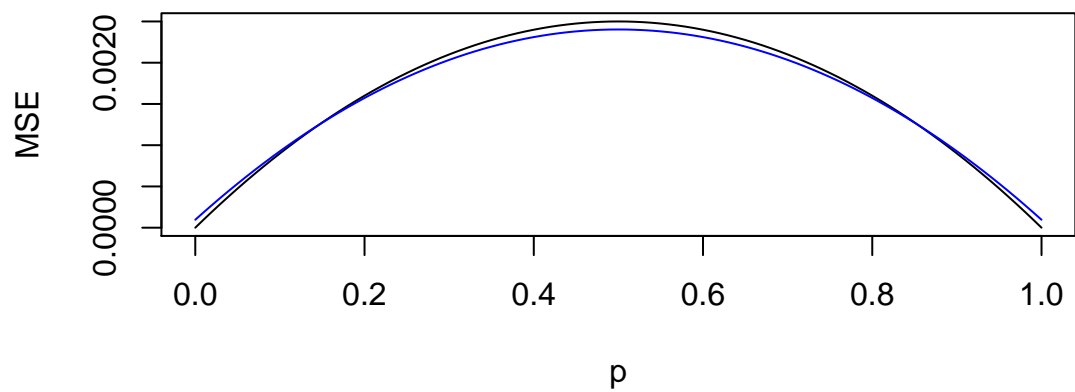
```
# Recreate plots using code from sec 6.3.3, changing the sample size
n <- 30
curve(x*(1-x)/n, from=0, to=1, xlab="p", ylab="MSE")
curve(n*(1-x)*x/(n+2)^2 + (1-2*x)^2/(n+2)^2, add=TRUE, col="blue", lty=1)
```



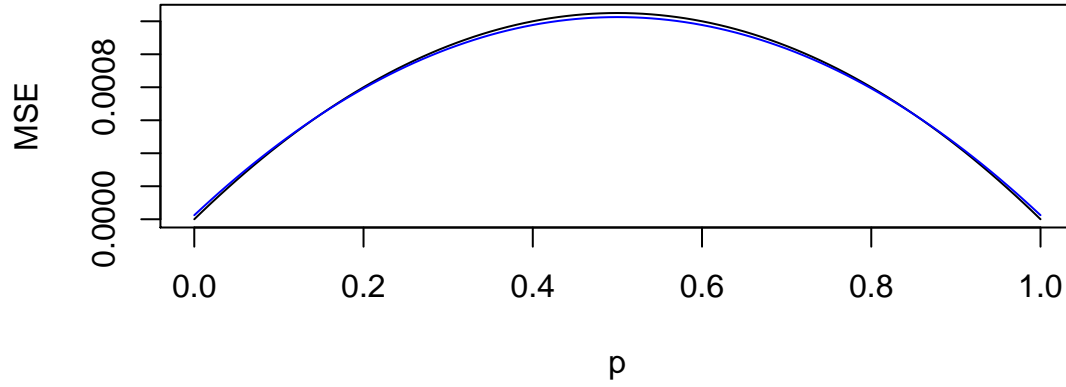
```
n <- 50
curve(x*(1-x)/n, from=0, to=1, xlab="p", ylab="MSE")
curve(n*(1-x)*x/(n+2)^2+(1-2*x)^2/(n+2)^2, add=TRUE, col="blue", lty=1)
```



```
n <- 100
curve(x*(1-x)/n, from=0, to=1, xlab="p", ylab="MSE")
curve(n*(1-x)*x/(n+2)^2+(1-2*x)^2/(n+2)^2, add=TRUE, col="blue", lty=1)
```



```
n <- 200
curve(x*(1-x)/n, from=0, to=1, xlab="p", ylab="MSE")
curve(n*(1-x)*x/(n+2)^2+(1-2*x)^2/(n+2)^2, add=TRUE, col="blue", lty=1)
```



As n increases, the MSE decreases. In each case, the second estimator (the blue line) has a lower MSE except for low and high estimates of p .

6.37

a

$$E[\bar{X}] = E\left[\frac{X_1 + X_2}{2}\right] = E[X_1]/2 + E[X_2]/2 = E[X]/2 + E[X]/2 = E[X] = 1/\lambda$$

b

$$Var[\bar{X}] = Var\left[\frac{X_1 + X_2}{2}\right] = \frac{1}{4}Var[X_1 + X_2] = \frac{1}{4}(Var[X] + Var[X]) = \frac{1}{4}\left(\frac{1}{\lambda^2} + \frac{1}{\lambda^2}\right) = \frac{2}{\lambda^2}$$

c

$$E[\sqrt{X_1 X_2}] = E[\sqrt{X_1} \sqrt{X_2}] = E[\sqrt{X_1}]E[\sqrt{X_2}]$$

By fact,

$$E[\sqrt{X_1 X_2}] = \sqrt{\pi}/2\sqrt{\lambda}(\sqrt{\pi}/2\sqrt{\lambda}) = \pi/4\lambda$$

d

$$bias = |E[\sqrt{X_1 X_2}] - 1/\lambda| = |\pi/4\lambda - 1/\lambda| = \left|\frac{\pi - 4}{4\lambda}\right|$$

Thus $bias = 0.2146/\lambda$.