Problem 1

(a) Let $f(x)$ be the pdf of Uniform $(0,1)$

$f(x) = 1\, \mathbb{1}_{(0<x<1)}$

$\therefore h(x) = x^2$

$\therefore \hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} h(x_i)$

$Var(\hat{\mu}_n) = \frac{1}{n-1} \times \frac{1}{n} \sum_{i=1}^{n} [h(x_i) - \hat{\mu}_n]^2$

1. Generate $X \sim U(0,1)$
2. Calculate $h(x)$
3. Calculate $\hat{\mu}_n$ and $Var(\hat{\mu}_n)$

For $n = 1\times10^6$

* The estimated integral is calculated as 0.3333605, variance is 8.885079×10^{-8}.

(b) Let $f(x,y)$ be the joint pdf of Uniform $(-2,2)$ and Uniform $(0,1)$

$f(x,y) = \frac{1}{4} \mathbb{1}_{(-2<x<2,\ 0<y<1)}$

$\therefore h(x,y) = 4x^2\cos(xy)$

$\therefore \hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} h(x_i, y_i)$

$Var(\hat{\mu}_n) = \frac{1}{n-1} \times \frac{1}{n} \sum_{i=1}^{n} [h(x_i, y_i) - \hat{\mu}_n]^2$

1. Generate $X \sim U(-2,2)$, $Y \sim U(0,1)$
2. Calculate $h(x,y)$
3. Calculate $\hat{\mu}_n$ and $Var(\hat{\mu}_n)$

For $n = 1\times10^6$

* The estimated integral is calculated as 3.485711, variance is 1.274877×10^{-5}.

(c) Let $f(x)$ be the pdf of Exponential $(\frac{1}{4})$, $x > 0$

$f(x) = \frac{1}{4} e^{-\frac{x}{4}}$

$\therefore h(x) = 3x^4 e^{-\frac{1}{4}(x^3 - x)}$

$\therefore \hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} h(x_i)$

$Var(\hat{\mu}_n) = \frac{1}{n-1} \times \frac{1}{n} \sum_{i=1}^{n} [h(x_i) - \hat{\mu}_n]^2$

1. Generate $X \sim Exp(\frac{1}{4})$
2. Calculate $h(x)$
3. Calculate $\hat{\mu}_n$ and $Var(\hat{\mu}_n)$

For $n = 1\times10^6$

* The estimated integral is calculated as 2.276883, variance is 1.348142×10^{-5}.

Problem 2

Let $f(x)$ be the pdf of Uniform $(1,2)$

$$f(x) = 1 \, \mathbb{1}_{(1<x<2)}$$

$$I = \int_1^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \int_1^2 h(x) f(x) dx$$

$$\therefore h(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$w^*(x) = \frac{f(x)}{g(x)} = \frac{1}{N(1.5, \upsilon^2)} \mathbb{1}_{(1<x<2)} \quad ; = 0 \quad o.w.$$

$\therefore$ 1. Generate $X \sim g$

2. estimate I by $\hat{I}_g = \frac{1}{m} \sum_{i=1}^{m} h(x_i) w^*(x_i)$

$$= \frac{1}{m} \sum_{i=1}^{m} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \times \sqrt{2\pi\upsilon^2} \, e^{\frac{(x-1.5)^2}{2\upsilon^2}} \quad (1<X<2) \; ; = 0 \quad o.w$$

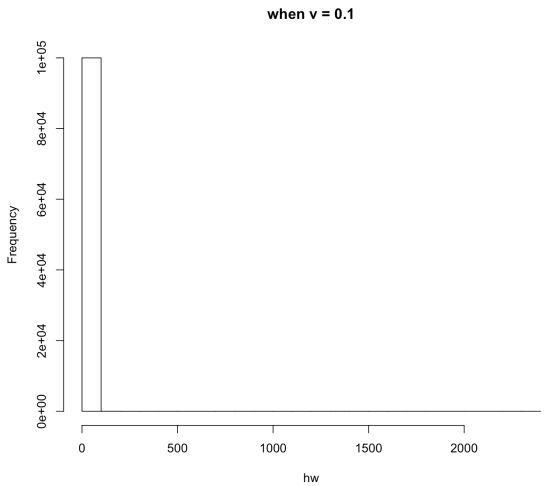
$$= \frac{1}{m} \sum_{i=1}^{m} \upsilon e^{\left(-\frac{x^2}{2} + \frac{(X-1.5)^2}{2\upsilon^2}\right)} \quad (1<X<2) \; ; = 0 \quad o.w.$$

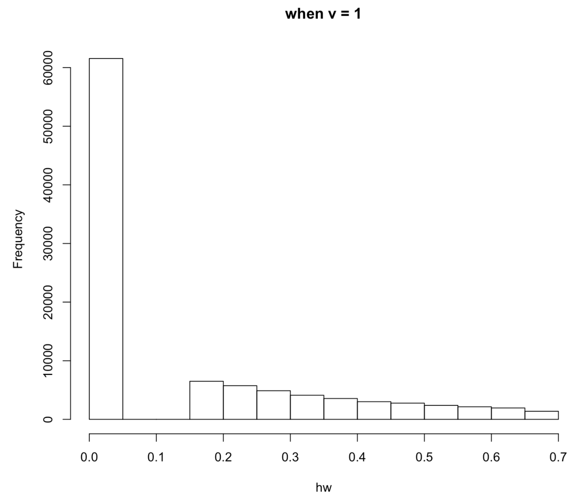
When $\upsilon = 0.1$, the estimated integral is 0.1265772, variance is 0.0001468637

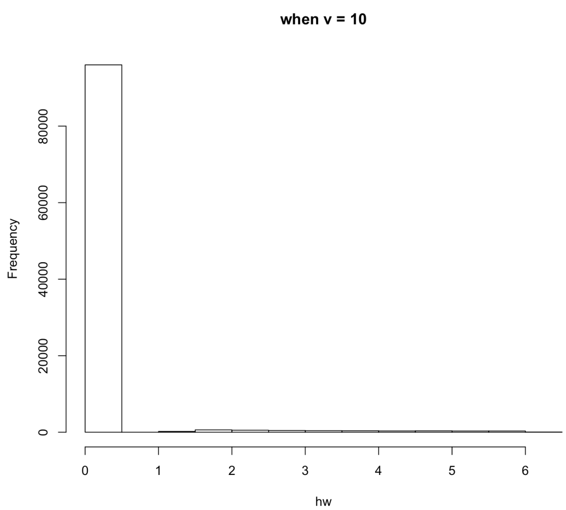
When $\upsilon = 1$, the estimated integral is 0.1366478, variance is 3.844606×10^{-7}

When $\upsilon = 10$, the estimated integral is 0.1357925, variance is 5.201715×10^{-6}

From the histograms, we can observe that there're extreme values when $\upsilon = 0.1$ and 10.


when v = 0.1
Frequency
0e+00
2e+04
4e+04
6e+04
8e+04
1e+05
0
500
1000
1500
2000
hw


when v = 1
Frequency
0
10000
20000
30000
40000
50000
60000
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
hw


when v = 10
Frequency
0
20000
40000
60000
80000
0
1
2
3
4
5
6
hw

Problem 3

(a) 1. Generate $X \sim \text{Uniform}(0,1)$

2. Calculate $h(x)$

3. Calculate $\hat{I}_{MC} = \frac{1}{n}\sum_{i=1}^{n} h(x_i)$

The estimated integral is $\hat{I}_{MC} = 0.6923455$, variance is 1.269863×10^{-5}

(b) $E[c(x)] = 1 + E(x) = 1 + 0.5 = 1.5$

Under X and $h(x)$ generated in part (a)

Calculate $c(x) = 1 + x$

Generate linear regression under $h(x) \sim c(x)$

$h(x) = b_0 + b^* c(x) + e$

the optimal value for b is b^* ($b^* = -0.4772632$)

$\rightarrow$ Calculate $\hat{I}_{CV} = \frac{1}{n}\sum_{i=1}^{n} h(x_i) - b^*\left[\frac{1}{n}\sum_{i=1}^{n} c(x_i) - E[c(x)]\right]$

The estimated integral is $\hat{I}_{CV} = 0.692441$

(c) $\frac{Var(\hat{I}_{CV}^{opt})}{Var(\hat{I}_{MC})} = 1 - \rho^2$, ρ is correlation coefficient between $h(x)$ and $c(x)$

$\rho = -0.9840014$, $Var(\hat{I}_{MC}) = 1.269863 \times 10^{-5}$

So $Var(\hat{I}_{CV}^{opt}) = 4.030708 \times 10^{-7}$,

$Var(\hat{I}_{CV}^{opt}) < Var(\hat{I}_{MC})$

(d) Set $c(x)\_new = 1 + x^{0.5}$

$\rho\_new = -0.99783130$

So $Var(\hat{I}_{CV}^{opt}\_new) = 5.501937 \times 10^{-8}$

With this new control variate, variance is smaller than $Var(\hat{I}_{CV})$.

Problem 4

(a) $f(x) = \frac{1}{2\theta} e^{-\frac{|x|}{\theta}}$, $\theta > 0$ $\quad f(x_n|\theta) = (2\theta)^{-\sum_{i=1}^{k} n_i} e^{-\sum_{i=1}^{k}\sum_{j=1}^{n_i} \frac{|e_{ij}|}{\theta}}$ $e_{ij} \sim$ double exponential $(0,\theta)$

with model : $y_{ij} = \mu + \alpha_i + e_{ij}$ $(i=1,\dots,k$, $j=1,\dots,n_i)$, $y_{ij} \sim$ double exponential $(\mu+\alpha_i, \theta)$

H_0: $\alpha_i = 0$ for all i ; H_1: not all α_i are zero

$$\lambda(x) = \frac{\sup_{\theta\in\Theta_0} f(x_n|\theta)}{\sup_{\theta\in\Theta} f(x_n|\theta)} = \frac{(2\hat{\theta}_0)^{-\sum_{i=1}^{k} n_i} e^{-\sum_{i=1}^{k}\sum_{j=1}^{n_i}|y_{ij}-\mu|/\hat{\theta}_0}}{(2\hat{\theta})^{-\sum_{i=1}^{k} n_i} e^{-\sum_{i=1}^{k}\sum_{j=1}^{n_i}|y_{ij}-(\mu+\alpha_i)|/\hat{\theta}}}$$

since $\sum \alpha_j = 0$, $\lambda(x) = \left(\frac{\hat{\theta}_0}{\hat{\theta}}\right)^{-\sum_{i=1}^{k} n_i}$

the constrained MLE of θ is :

$$\log f(x_n|\theta) = -\sum_{i=1}^{k} n_i \log(2\theta) - \sum_{i=1}^{k}\sum_{j=1}^{n_i}\frac{|y_{ij}-(\mu+\alpha_i)|}{\theta}$$

$$\frac{\partial \log f(x_n|\theta)}{\partial\theta} = -\sum_{i=1}^{k} n_i \frac{1}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}|y_{ij}-(\mu+\alpha_i)|, \quad \hat{\theta}_0 = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}|y_{ij}-\tilde{y}_{ij}|}{\sum_{i=1}^{k} n_i}$$, $\tilde{y}_{ij}$ is median of y_{ij} for each i & j)

the unconstrained MLE of θ is: $\hat{\theta} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}|y_{ij}-\tilde{y}_i|}{\sum_{i=1}^{k} n_i}$, $\tilde{y}_i$ is median of $y_{i1},\dots,y_{in_i}$

$$\text{so } \lambda(x) = \left(\frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}|y_{ij}-\tilde{y}_i|}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}|y_{ij}-\tilde{y}_{ij}|}\right)^{\sum_{i=1}^{k} n_i}$$

$$\Rightarrow \lambda(x)' = \left(\frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}|e_{ij}-\tilde{e}_i|}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}|e_{ij}-\tilde{e}_{ij}|}\right)^{\sum_{i=1}^{k} n_i}$$ $\tilde{e}_i$ is median of $e_{i1},\dots,e_{in_i}$

$\tilde{e}_{ij}$ is median of e_{ij} for each i,j

$$\Rightarrow \lambda(x)'' = \left(\frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}|e_{ij}/\theta-\tilde{e}_i/\theta|}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}|e_{ij}/\theta-\tilde{e}_{ij}/\theta|}\right)^{\sum_{i=1}^{k} n_i}$$

and $e_{ij}/\theta \sim$ double exponential $(0,1)$

$$f(x) = \frac{1}{2} e^{-|x|}$$

So the algorithm can be:

1. Generate e_{ij} $(i=1,\dots,k$, $j=1,\dots,n_i)$ $\sim$ double exponential $(0,1)$
2. calculate $\lambda(x)'$, repeat 999 times.
3. If $\lambda(x)$ computed by sample is among the smallest $\alpha\%$ or largest $\alpha\%$ of these $\lambda(x)'$, we can reject H_0.

(b)

for permutation test:

1. combine y_{ij} ($i=1,\dots,k$, $j=1,\dots,n_i$)
2. for $i=1,\dots,k-1$, draw n_i data points for group i, with the left n_k points in group k automatically. (without replacement).
3. Calculate $\sum_{i>j} |\hat{\mu}_i - \hat{\mu}_j|$, ($\hat{\mu}_i$ is the group mean of group i).
4. repeat 1 to 3 for n (large number) times
5. if test value from sample is outside the middle $(1-\alpha)\%$ of these $\sum_{i>j} |\hat{\mu}_i - \hat{\mu}_j|$'s, reject H_0.

Problem 5

(b) (i) $(\lambda | p, r, x) \sim \text{Gamma}(a + \Sigma x_i,\ b + \Sigma r_i)$.

$$f(\lambda | p, r, x) \propto f(x, r, \lambda, p)$$

$$f(x, r, \lambda, p) = \lambda^{a-1+\Sigma x_i} e^{-\lambda(\Sigma r_i + b)} \frac{b^a}{\Gamma(a)} \left(\prod_{i=1}^{n} \frac{r_i^{x_i}}{x_i!} \right) p^{\Sigma r_i} (1-p)^{n - \Sigma r_i}$$

$$\therefore f(\lambda | p, r, x) \propto \lambda^{a + \Sigma x_i - 1} e^{-\lambda(\Sigma r_i + b)}$$

$$(\lambda | p, r, x) \sim \text{Gamma}(a + \Sigma x_i,\ b + \Sigma r_i)$$

(ii) $(p | \lambda, r, x) \sim \text{Beta}(1 + \Sigma r_i,\ n + 1 - \Sigma r_i)$

$$f(p | \lambda, r, x) \propto f(x, r, \lambda, p)$$

$$f(x, r, \lambda, p) = p^{\Sigma r_i} (1-p)^{n - \Sigma r_i} \frac{b^a \lambda^{a-1} e^{-b\lambda}}{\Gamma(a)} \prod_{i=1}^{n} \frac{e^{-\lambda r_i} (\lambda r_i)^{x_i}}{x_i!}$$

$$\therefore f(p | \lambda, r, x) \propto p^{\Sigma r_i} (1-p)^{n - \Sigma r_i}$$

$$(p | \lambda, r, x) \sim \text{Beta}(1 + \Sigma r_i,\ n + 1 - \Sigma r_i)$$

(iii) $(r_i | \lambda, p, x) \sim \text{Bernoulli}\left(\frac{pe^{-\lambda}}{pe^{-\lambda} + (1-p) 1(x_i = 0)} \right)$

$$f(x, r, \lambda, p) = \frac{b^a \lambda^{a-1} e^{-b\lambda}}{\Gamma(a)} \prod_{j=1}^{n} \frac{e^{-\lambda r_j} (\lambda r_j)^{x_j}}{x_j!} p^{r_j} (1-p)^{1-r_j}$$

if $r_i = 0$, $P(x_i = 0) = 1$

$$f_{r_i=0}(r_i | x, \lambda, p) \propto (1-p) 1\{x_i = 0\}$$

if $r_i = 1$; $f_{r_i=1}(r_i | x, \lambda, p) \propto pe^{-\lambda}$

so $(r_i | \lambda, p, x) \sim \text{Bernoulli}\left(\frac{pe^{-\lambda}}{pe^{-\lambda} + (1-p) 1\{x_i = 0\}} \right)$.

(c) I first sample 10000 data points, initial r_i and sample one λ and p, update to sample new r_i, repeat for 10000 times. Take first 1000 data points as burn-in, and calaulate 95% confidence intervals for lambda and p. Then change a=b=1 to a=b=2/5.

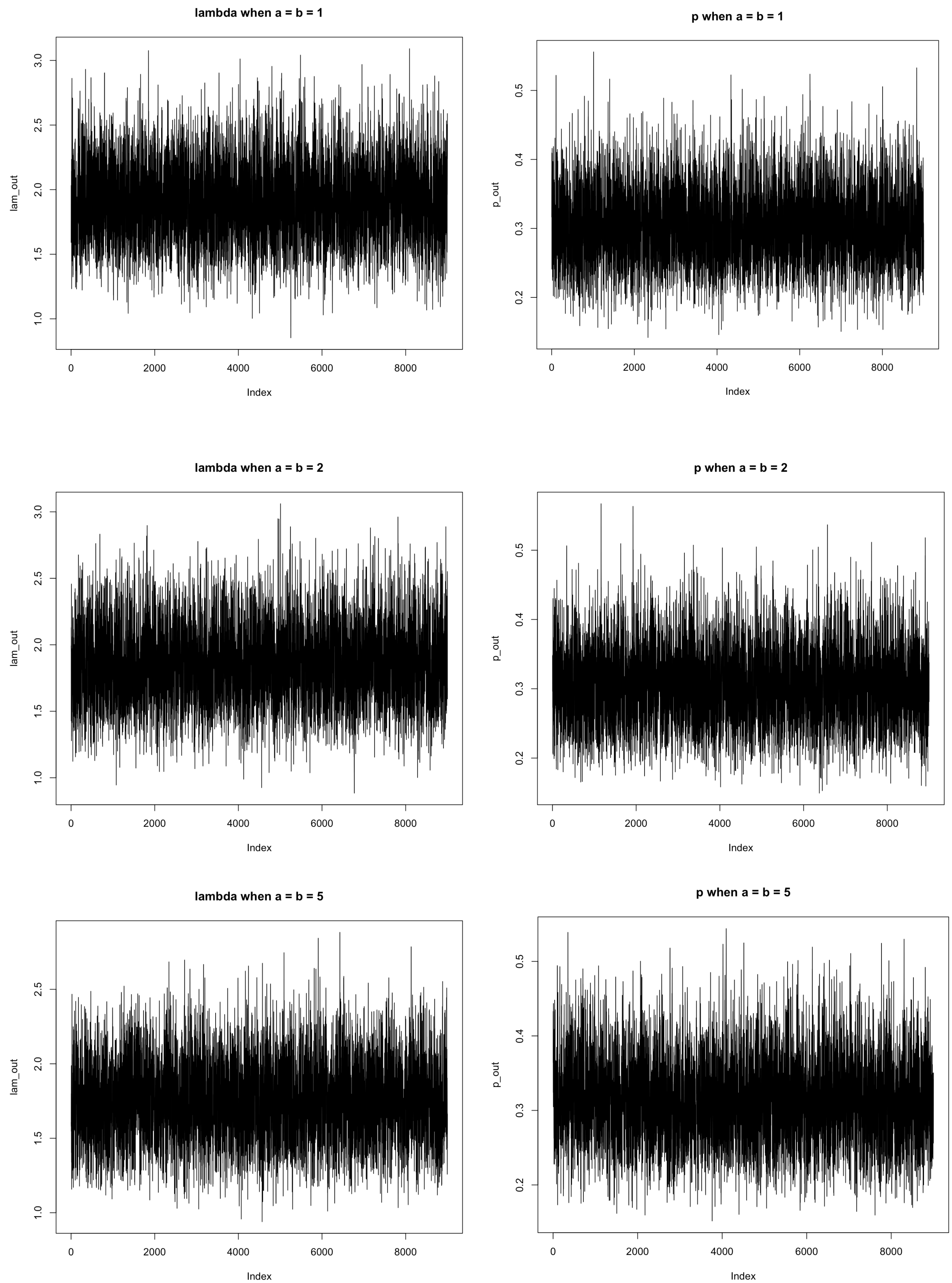
Lambda:	2.5%	97.5%	P:	2.5%	97.5%
a=b=1	1.345102	2.521574	a=b=1	0.2067333	0.471872
a=b=2	1.307687	2.462401	a=b=2	0.2066182	0.4230448
a=b=5	1.259505	2.297989	a=b=5	0.2119599	0.4322999

(for λ, the interval becomes narrower as a=b increase, but not obvious)

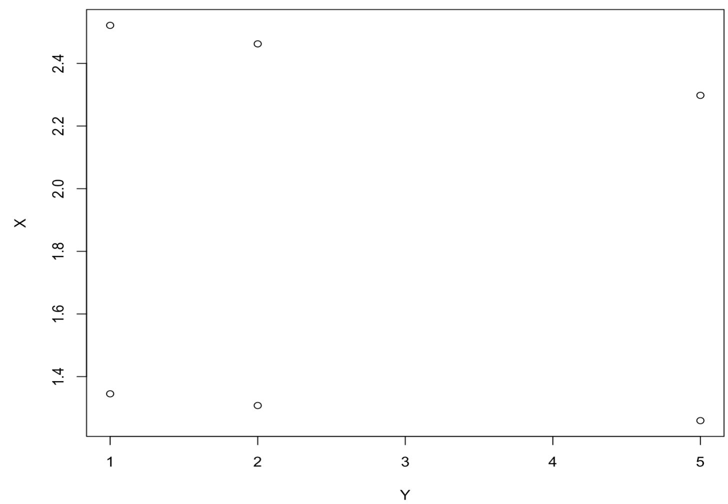
As a=b increases, the differences between intervals are not obvious. $\lambda = 2$, $p = 0.3$ is included in all.

Problem 5

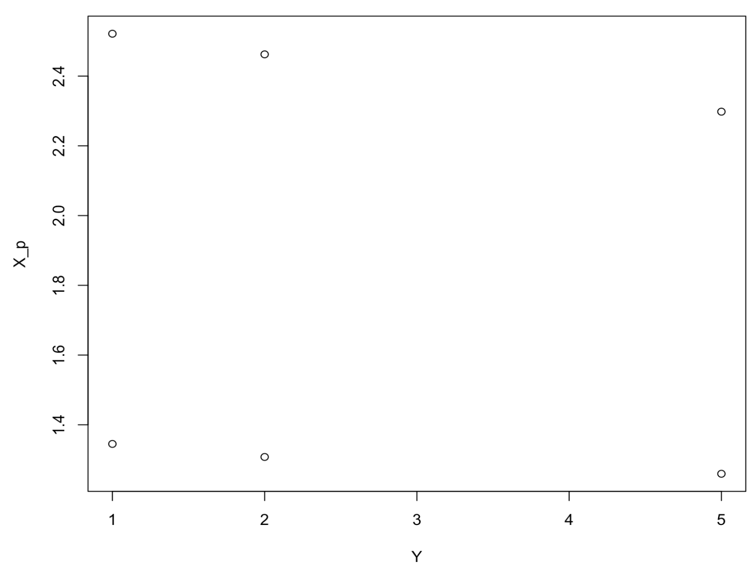
```
# part a

#[1] 0 2 0 0 0 3 0 2 0 0 4 0 0 0 0 0 0 4 0 0 0 1 2 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0
#[41] 0 0 0 1 0 0 0 1 0 0 0 4 0 0 0 2 0 0 0 0 1 3 0 3 0 1 0 0 0 0 3 5 5 0 0 1 1 0 0 0
#[81] 0 1 0 0 0 0 0 2 0 0 1 0 0 0 0 0 0 0 0 2
```


lambda when a = b = 1
lam_out
Index
p when a = b = 1
p_out
Index
lambda when a = b = 2
lam_out
Index
p when a = b = 2
p_out
Index
lambda when a = b = 5
lam_out
Index
p when a = b = 5
p_out
Index

the 95% confidence intervals of lambda



the 95% confidence intervals of p

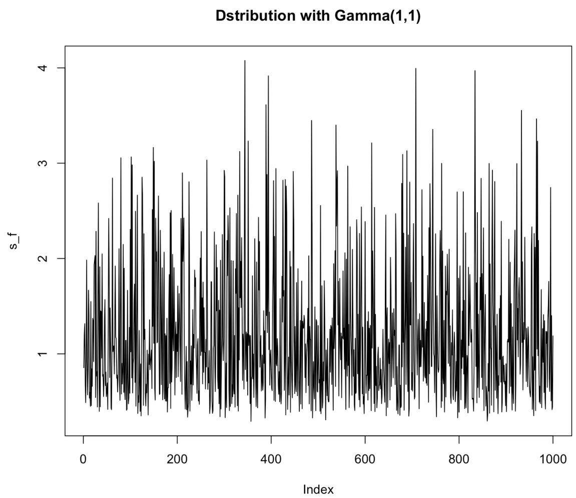


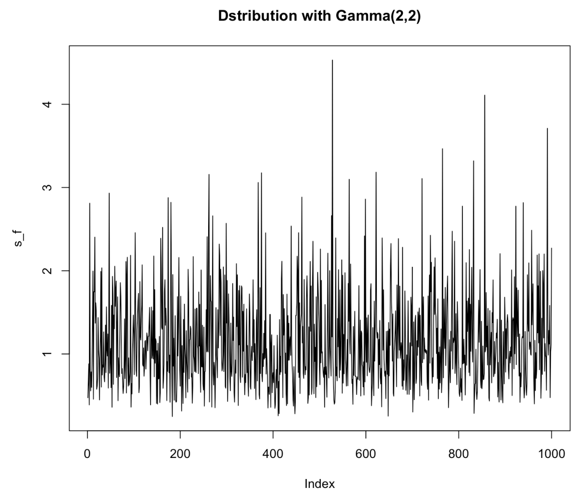
Problem 6

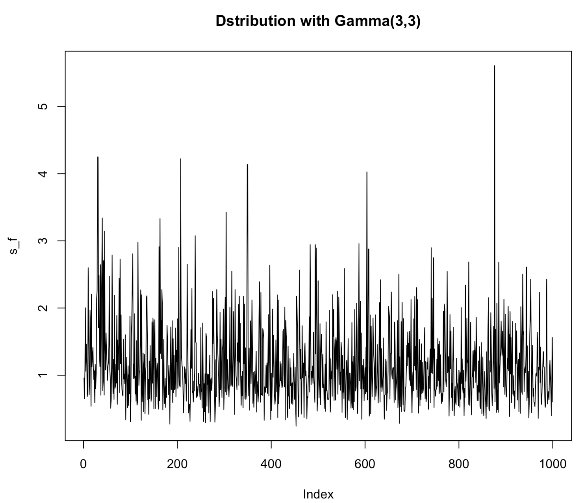
I draw 5000 samples, chose the first 1000 as burn-in, and selected sample every four points.

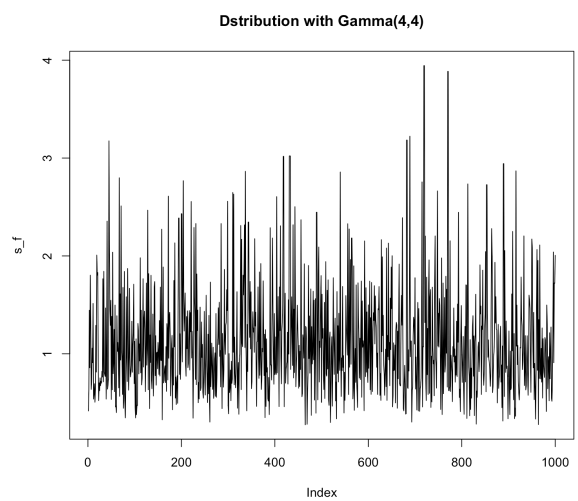
The true mean for z and $\frac{1}{z}$ is 1.154701 and 1.116025. For different Gamma distributions, the outcomes are:

Gamma Distribution	$E(\hat{z})$	$E(\frac{1}{z})$	
Gamma (1, 1)	1.14263	1.112668	
Gamma (2, 2)	1.151036	1.107117	→ the most accurate among the four
Gamma (3, 3)	1.160904	1.112677	
Gamma (4, 4)	1.133949	1.121624	


Dstribution with Gamma(1,1)
s_f
Index
0
200
400
600
800
1000
1
2
3
4


Dstribution with Gamma(2,2)
s_f
Index
0
200
400
600
800
1000
1
2
3
4


Dstribution with Gamma(3,3)
s_f
Index
0
200
400
600
800
1000
1
2
3
4
5


Dstribution with Gamma(4,4)
s_f
Index
0
200
400
600
800
1000
1
2
3
4