STA 243 Assignment 4

Chen Zihao(50%) Junyi Song(50%)

1. Monte Carlo integration

In this part, it is just some basic coding. For example, (a), using runif(n,a,b) function to generate sample $x \in [a, b]$ and evaluate each of the integrals. Details are in the coding file, here is just the results for each question.

(a)

```
## [1] "the integral for (a) is 0.333807882745648"
```

(b)

```
## [1] "the integral for (b) is 3.46411636719222"
```

(c)

```
## [1] "the integral for (c) is 2.27395767095112"
```

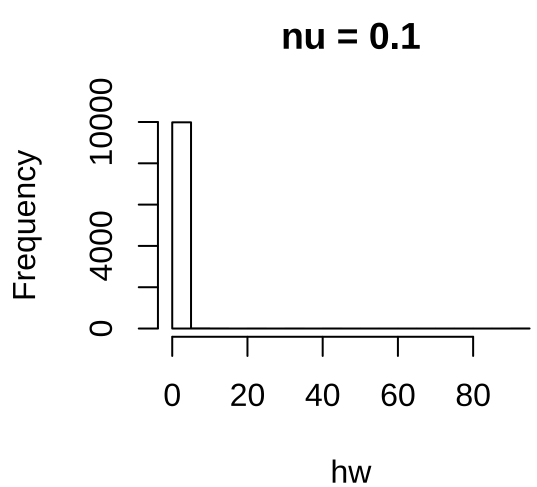
2. Monte Carlo Importance Sampling

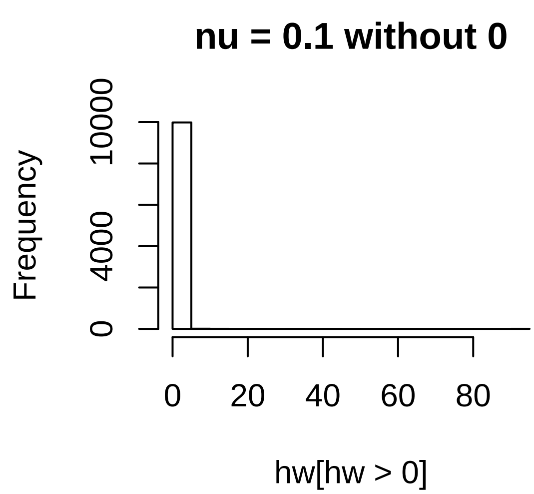
let f(x) be unif(1,2) and use the $N(1.5, \nu^2)$ as g.

$$\begin{aligned} I &= \int_1^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\ &= \int_1^2 \frac{\nu e^{-x^2/2}}{e^{-(x-1.5)^2/2\nu^2}} \frac{1}{\sqrt{2\pi\nu^2}} e^{-(x-1.5)^2/2\nu^2} dx \end{aligned}$$

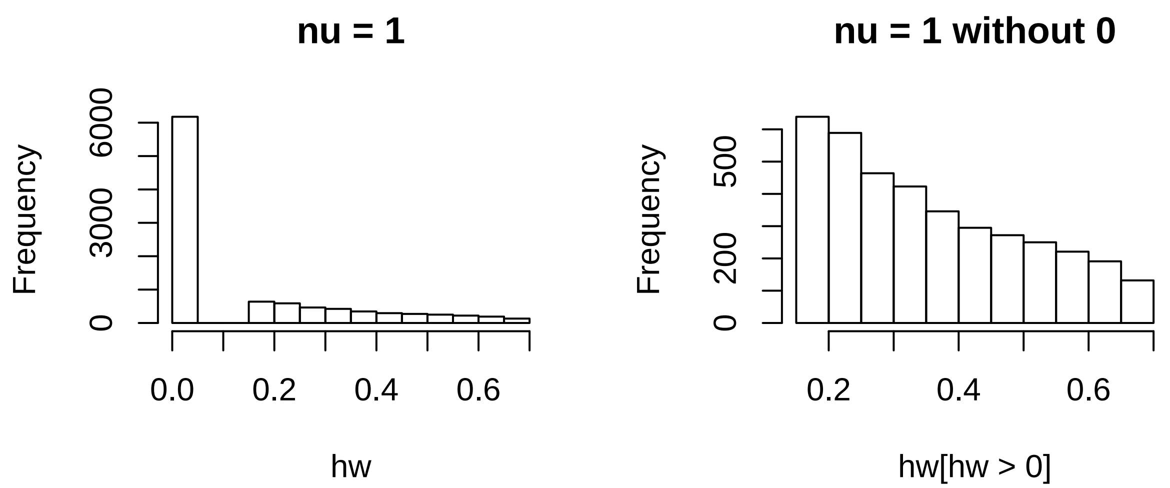
so that we have our $h(x)\frac{f(x)}{g(x)}$ and the $g(x)$

first generate $x_1, x_2, ..., x_m$ iid from $g(x)$

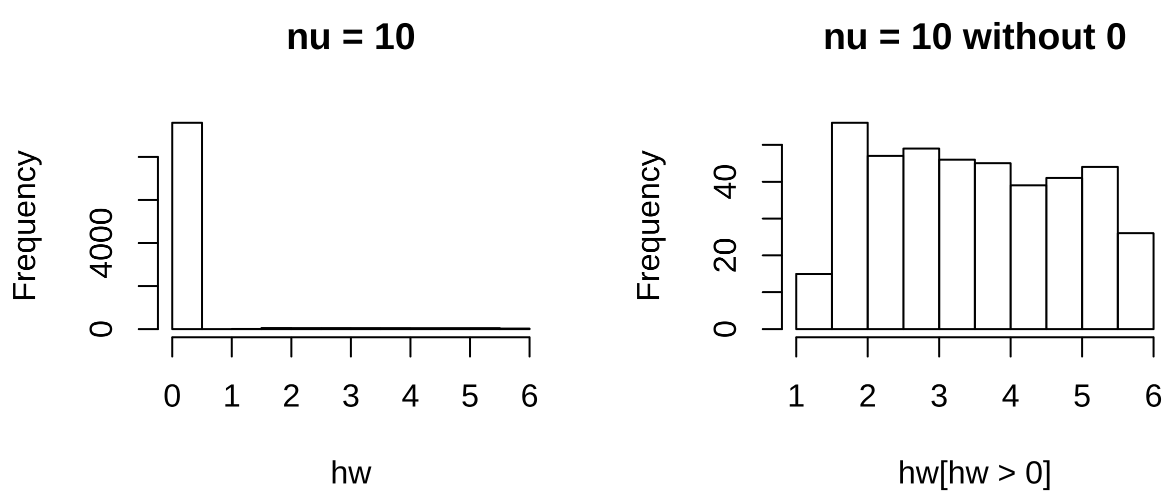




```
## [1] "the estimating integral value is 0.118473920964841"
```




```
## [1] "the estimating integral value is 0.135943427110703"
```




```
## [1] "the estimating integral value is 0.141839666541759"
```

Here we can see that there are outliers for case $\nu = 0.1$ and $\nu = 10$. for $\nu = 0.1$ the dispersion of x is too narrow and for $\nu = 10$, the dispersion is too wide. Both histogram plots on the left shows that about 10000 values ,almost all, are around 0. The hist plot on the right tells much more details. For $\nu = 10$, only a few values lies in [1,2], which need more in sample size to get more non-zero values. For $\nu = 0.1$, most of the x value focus on 1.5 leads that majority of the h(x)w(x) values is close to a single point, in this case, 0. $\nu = 1$ is the best among the three.

3. Control Variate Method

(a) folowing the instruction and here is the result.

```
## [1] "the estimating integral value is 0.699856119320028"
```

The exact value for I is $\ln 2 \approx 0.6931472$

```
## [1] "the sample standard error of the estimating integral value is 0.00372963741051895"
```

(b) and (c)

$E(1+U) = 1 + E(U) = 1 + 0.5 = 1.5$.

$Var[c(U)] = Var(1+U) = \frac{1}{12}$ and $Cov[h(U), c(U)] = 1 - \frac{3\ln 2}{2}$

then we can get the best value for b, $b^* = \frac{Cov[h(U),c(U)]}{Var[c(U)]} \approx -0.4767$

```
## [1] "the estimating integral value is 0.694499660712486"
## [1] "the sample standard error of the estimating integral value is 0.000655148202157561"
```

The variance of the CV version is less which implies the CV version is better.

(d) Here we use $d(x)\sqrt{x}$ instead of $c(x) = 1 + x$.

```
## [1] "the estimating integral value is 0.693055566293582"
## [1] "the sample standard error of the estimating integral value is 0.000230141874533201"
```

the sample stanard error is even smaller than the previous one.

4. ANOVA

(a)

Considering a model $y_{ij} = \mu + \alpha_i + e_{ij}$

e_{ij} have independent and identical double exponential distribution centered on zero with pdf

$$f(x) = \frac{1}{2\theta}\exp(-\frac{|x|}{\theta}), -\infty < x < \infty$$

where $\theta > 0$ and have variance $2\theta^2$

the hypothesis:

- the null hypothesis is $H_0 : \forall\alpha_i = 0$ and $\theta > 0$.
- the alternative hypothesis is $H_1 : \exists\alpha_i \neq 0$ and $\theta > 0$

For this model, suppose we have k groups and for the ith group we have n_i observations,totally n observations. According to the model, $e_{ij} = y_{ij} - \frac{1}{n}\sum_{j=1}^{n_i} y_{ij} = y_{ij} - \mu - \alpha_i$. Under H_0 we have $e_{ij} = y_{ij} - \mu$. use the sample variance $S^2$to estimate the variance $2\theta^2$ and we can get $\hat{\theta} = \sqrt{S^2/2}$

conduct the following test:

1. generate $x_1^*, x_2^*, ..., x_n^*$ from
$$f(x) = \frac{1}{2\hat{\theta}}\exp(-\frac{|x|}{\hat{\theta}}), -\infty < x < \infty$$
and calculate the mean $\bar{x}^*$
2. repeat this for 999 times to get 999 $\bar{x}^*$'s
3. for ith group $(i = 1, 2, ..., k)$, calculate $T_i = \frac{1}{n_i}\sum_{j=1}^{n_i} y_{ij} - \bar{y}$ where $\bar{y}$ is the mean of $\{y_{ij}\}$
4. if any T_i is amongst the smallest 2.5% or is amongst the largest 2.5% of the $\bar{x}^*$'s, we reject the H_0

(b)

Using the Permutation tests.

1. calculate μ and α_i for the model.
2. merge all the y_{ij} to form a sample of n data points.
3. draw without replacement to form k groups with the number of sample as the same as the original groups. In other words, draw n_i sample for the ith group.
4. calculate μ and α_i as μ^* and α_i^*.
5. do the drawing and calculating procedure many times.
6. if any original α_i is outside of the middle 95% of the α_i^*'s, reject H_0

5. Gibbs Sampling

(a) Generate a random sample of size n = 100 for the ZIP model using parameters $p = 0.3$ and $\lambda = 2$.

According to the question, $X_i = R_i Y_i$ where $Y_i's$ have a Poisson(λ) distribution and the $R_i's$ have a Bernoulli(p) distribution, all independent of each other.

first generate 100 $R_i \sim Bernoulli(p)$ and 100 $Y_i \sim Poisson(\lambda)$ and let $X_i = R_i Y_i$

(b)

i. $(\lambda|p, r, x)$ treating p, r, x as fixed in $f(p, \lambda, r, x)$

$$\begin{aligned} f(\lambda|p, r, x) &\propto \lambda^{a-1} e^{-b\lambda} \prod_{i=1}^{n} e^{-\lambda r_i} \lambda^{x_i} \\ &\propto \lambda^{a+\sum_i x_i - 1} e^{-(b+\sum_i r_i)\lambda} \end{aligned}$$

which is the the pdf of Gamma$(a + \sum_i x_i, b + \sum_i r_i)$ with shape and rate parameters.

ii. $(p|\lambda, r, x)$ treating λ, r, x as fixed in $f(p, \lambda, r, x)$

$$\begin{aligned} f(p|\lambda, r, x) &\propto \prod_{i=1}^{n} p^{r_i} (1-p)^{1-r_i} \\ &\propto p^{(\sum_i r_i + 1) - 1} (1-p)^{(n - \sum_i r_i + 1) - 1} \end{aligned}$$

which is the the pdf of Beta$(\sum_i r_i + 1, n - \sum_i r_i + 1)$.

iii.

$(r_i|\lambda, p) \sim Bernoulli(p)$ and $(x_i|r, \lambda, p) \sim Poisson(\lambda r_i)$ and etc.

r_i's are independent.x_i's are independent.

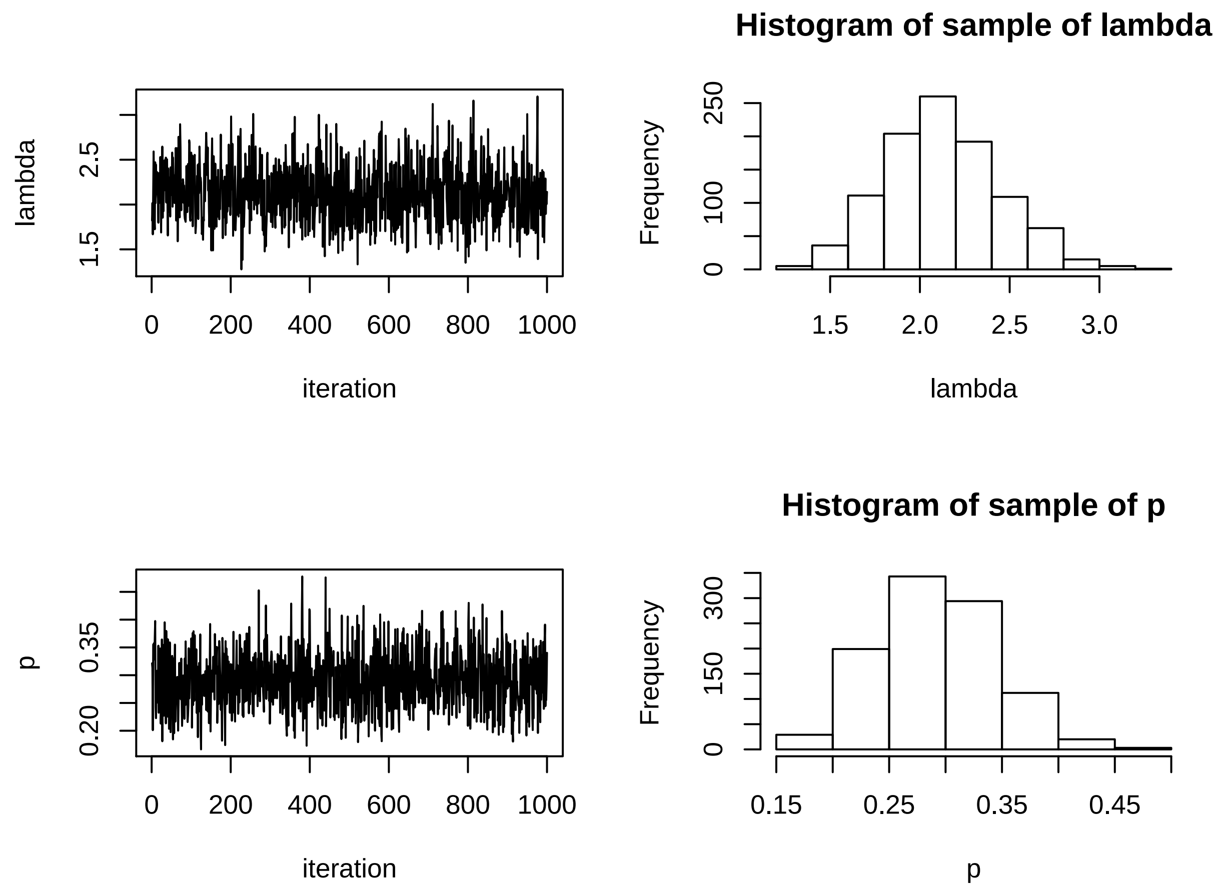
$$\begin{aligned} f(r_i|\lambda, p, x) &= \frac{f(r_i, x|\lambda, p)}{f(x_i|\lambda, p)} \\ &= \frac{f(r_i, x, \lambda|p)}{f(x_i|\lambda, p) f(\lambda|p)} \\ &= \frac{f(r_i, x, \lambda, p)}{f(x_i|\lambda, p) f(\lambda|p) f(p)} \\ &= \frac{e^{-\lambda r_i} (\lambda r_i)^{x_i} p^{r_i} (1-p)^{1-r_i}}{x_i!(f(x|\lambda, p)} \\ &= \frac{e^{-\lambda r_i} (\lambda r_i)^{x_i} p^{r_i} (1-p)^{1-r_i}}{x_i!(f(x|r_i = 1, \lambda, p)p + f(x|r_i = 0, \lambda, p))(1-p)} \end{aligned}$$

so that

$$
\begin{aligned}
P(r_i = 1|\lambda, p, x) &= \frac{e^{-\lambda}\lambda^{x_i}p}{x_i!(\frac{(\lambda r_i)^{x_i}e^{-\lambda r_i}}{x_i!}|_{r_i=0}(1-p) + \frac{(\lambda r_i)^{x_i}e^{-\lambda r_i}}{x_i!}|_{r_i=1}p)} \\
&= \frac{e^{-\lambda}p}{(r_i)^{x_i}e^{-\lambda r_i}|_{r_i=0}(1-p) + (r_i)^{x_i}e^{-\lambda r_i}|_{r_i=1}p)} \\
&= \frac{e^{-\lambda}p}{I_{\{x_i=0\}}(1-p) + e^{-\lambda}p}
\end{aligned}
$$

so that $(r_i|\lambda, p, x) \sim Bernoulli(\frac{e^{-\lambda}p}{I_{\{x_i=0\}}(1-p)+e^{-\lambda}p})$

(c).



bayesian confidence intervals,

for p

```
##              lower     upper
## a=1,b=1 0.2083270 0.3834556
## a=2,b=1 0.2058523 0.3804842
## a=1,b=2 0.2110248 0.3866850
## a=2,b=2 0.2102448 0.3857523
```

for λ

```
##              lower    upper
## a=1,b=1 1.625237 2.668318
## a=2,b=1 1.670546 2.731981
## a=1,b=2 1.556872 2.556077
## a=2,b=2 1.589621 2.599638
```

We can see that the true value for $p = 0.3$ and $\lambda = 2$ are included in the confidence intervals.

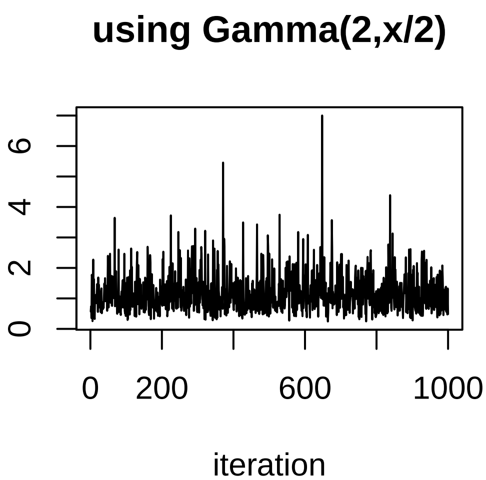
6. Independence-Metropolis-Hastings Algorithm

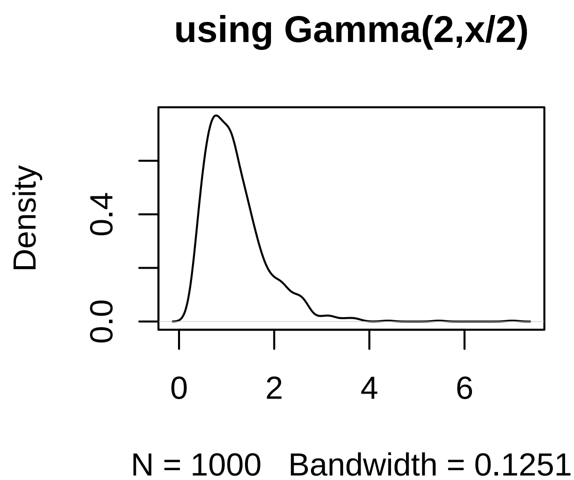
$Gamma(k, \theta)$ has mean $k\theta$

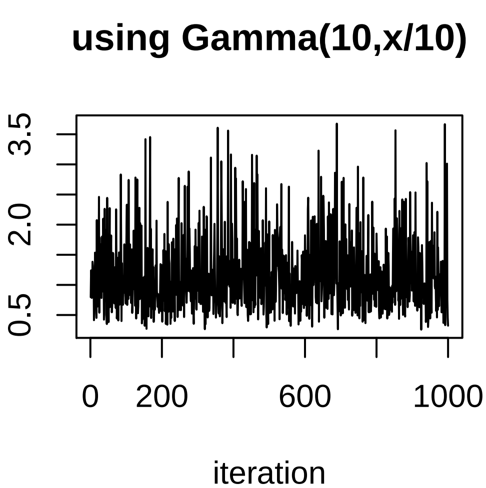
we can draw y_i from $Gamma(k, x_{i-1}/k)$ or $Gamma(x_{i-1}/\theta, \theta)$

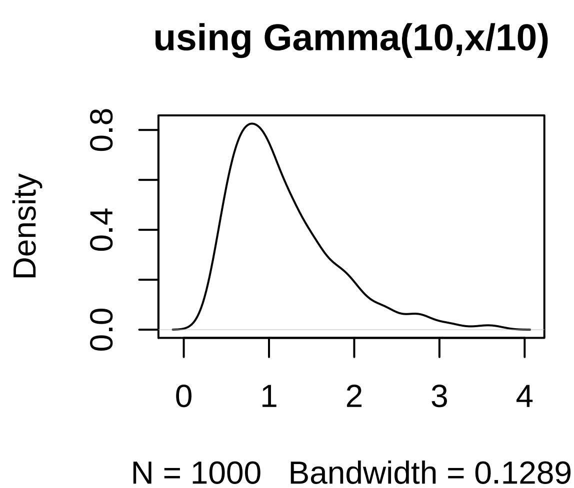
generate y_i from $q(y|x_{i-1})$ which is $gamma(k, x_{i-1}/k))$ or $gamma(x_{i-1}/\theta, \theta))$

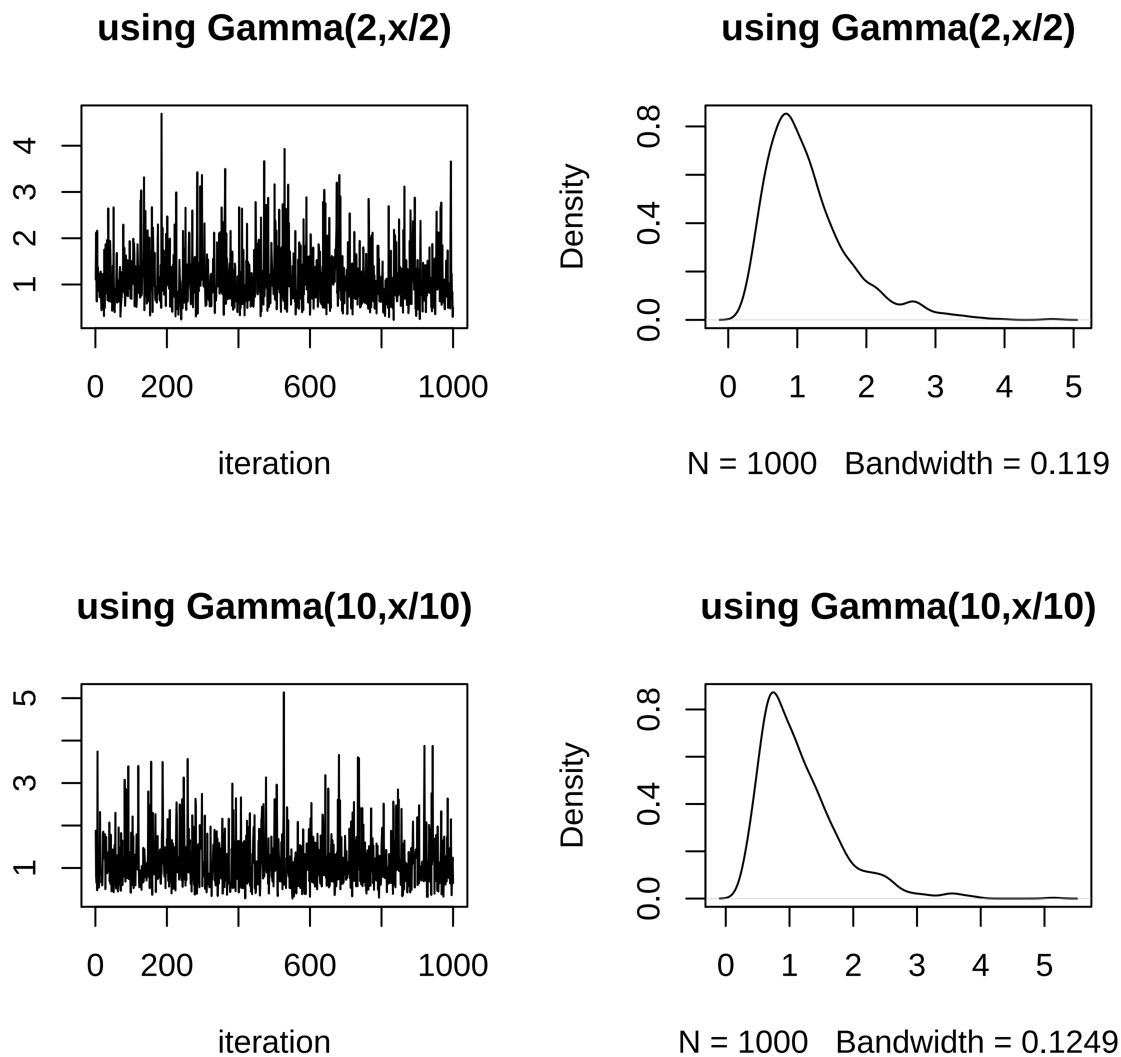
here is some plot of our output.











Here is the table showing the mean of the sampele and the mean of 1/sample

```
##                  sample.mean X1.sample.mean
## True value          1.154701       1.116025
## Gamma(2,x/2)        1.180678       1.110428
## Gamma(10,x/10)      1.164200       1.111339
## Gamma(x/2,2)        1.159860       1.113601
## Gamma(x/10,10)      1.161612       1.104353
```

they are similar enough with each other and it seems all of them provide reasonable estimates.