Zhuoqun Huang (Calvin) 908525
Tutorial information: Anubhav Kaphle, Thur 9-10

October 29, 2019

# MAS Assignment 1 Submission

revision 2

---

# Question1

(a) According to theorem Theorem 1.

The conditional distribution of $\mu|\tau, x_1...x_n$ is

$$\mu|\tau, x_1, .., x_n \sim N(\bar{x}, \frac{1}{n\tau}) \qquad (\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n})$$

and the conditional distribution of $\tau|\mu, x_1...x_n$ is

$$\tau|\mu, x_1, .., x_n \sim Gamma(\frac{n}{2}, \frac{2}{\sum_{i=0}^{n}(x_i - \mu)^2)}) \qquad \text{(shape, scale)}$$

Thus

$$p(\mu|\tau, x_1, .., x_n) = \frac{1}{\sqrt{2\pi\frac{1}{n\tau}}} exp(-\frac{1}{2}\frac{(\mu - \bar{x})^2}{\frac{1}{n\tau}})$$

$$p(\tau|\mu, x_1, .., x_n) = \frac{1}{\Gamma(\frac{n}{2})\frac{2}{\sum_{i=0}^{n}(x_i-\mu)^2)}^{\frac{n}{2}}} x^{\frac{n}{2}-1} exp(-\frac{x}{\frac{2}{\sum_{i=0}^{n}(x_i-\mu)^2)}})$$

**Theorem 1.** *The conditional distribution of $\mu|\tau, x_1...x_n$ is*

$$\mu|\tau, x_1, .., x_n \sim N(\bar{x}, \frac{1}{n\tau}) \qquad (\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n})$$

*and the conditional distribution of $\tau|\mu, x_1...x_n$ is*

$$\tau|\mu, x_1, .., x_n \sim Gamma(\frac{n}{2}, \frac{2}{\sum_{i=0}^{n}(x_i - \mu)^2)})$$

*Proof of Theorem 1.*

$$
\begin{aligned}
p(\mu,\tau,x_1,..,x_n) &= p(x_1,..,x_n|\mu,\tau)p(\mu,\tau) \\
&= \prod_{i=1}^{n}[\sqrt{\frac{\tau}{2\pi}}exp(-\frac{1}{2}(x_i-\mu)^2\tau)]\frac{1}{\tau} \\
&\propto \sqrt{\tau^{n-2}}exp(-\frac{\tau}{2}\sum_{i=0}^{n}(x_i-\mu)^2) \\
p(\mu|\tau,x_1,..,x_n) &\propto exp(-\frac{\tau}{2}\sum_{i=0}^{n}(x_i-\mu)^2) \\
&\propto exp(-\frac{\tau}{2}\sum_{i=0}^{n}(n\mu^2-2\mu\sum_{i=1}^{n}x_i)) \\
&\propto exp(-\frac{n\tau}{2}(\mu-\frac{\sum_{i=1}^{n}x_i}{n})^2) \\
&\propto exp(-\frac{1}{2}\frac{(\mu-\bar{x})^2}{\frac{1}{n\tau}} && (\bar{x}=\tfrac{\sum_{i=1}^{n}x_i}{n}) \\
\mu|\tau,x_1,..,x_n &\sim N(\bar{x},\frac{1}{n\tau}) \\
p(\tau|\mu,x_1,..,x_n) &\propto \tau^{\frac{n}{2}-1}exp(-\frac{\tau}{2}\sum_{i=0}^{n}(x_i-\mu)^2) \\
&\propto \tau^{\frac{n}{2}-1}exp(-\frac{\tau}{\frac{2}{\sum_{i=0}^{n}(x_i-\mu)^2)}} \\
\tau|\mu,x_1,..,x_n &\sim Gamma(\frac{n}{2},\frac{2}{\sum_{i=0}^{n}(x_i-\mu)^2)}) && \text{(shape, scale)}
\end{aligned}
$$

□

(b) see code at the end

(c) $E(\mu) = 5.090795, 90\%\ CI : [4.757753, 5.422978]$
$E(\tau) = 0.2503675, 90\%\ CI : [0.1949002, 0.3116865]$
see code at the end

# Question2

(a) see code at the end

(b) $E(\mu) = 5.093511, 90\%\ CI : [4.756718, 5.425346]$
$E(\tau) = 0.249857, 90\%\ CI : [0.1951379, 0.3107686]$
see code at the end

## Code and result for non-proof

In [1]:

```
library(ggplot2)

set.seed(30027)
x = rnorm(100, 5, 2)

output = function(param, name) {
    print(name)
    print('mean value:')
    print(mean(param))
    print('90%  CI:')
    print(quantile(param, c(0.05,0.95)))
    hist(param, prob=TRUE, main=name)            # prob=TRUE for probabilities n
ot counts
    lines(density(param))             # add a density estimate with defaults
    lines(density(param, adjust=2), lty="dotted")   # add another "smoother" den
sity
}
```

```
Registered S3 methods overwritten by 'ggplot2':
  method         from
  [.quosures     rlang
  c.quosures     rlang
  print.quosures rlang
```

# Q1

## Define algortihm

In [2]:

```
# (b)

gibbs <- function(n, data, mu=1, tau=1) {
  samples = list(mu=numeric(n), tau=numeric(n));
  xbar = mean(data)
  m = length(data)

  for (i in 1:n) {
    mu = rnorm(1, xbar, sqrt(1/(m * tau)))
    tau = rgamma(1, m/2, sum((data-mu)^2)/2)

    samples$mu[i] = mu
    samples$tau[i] = tau
  }
  return(samples)
}
```

### (b) Trace plot

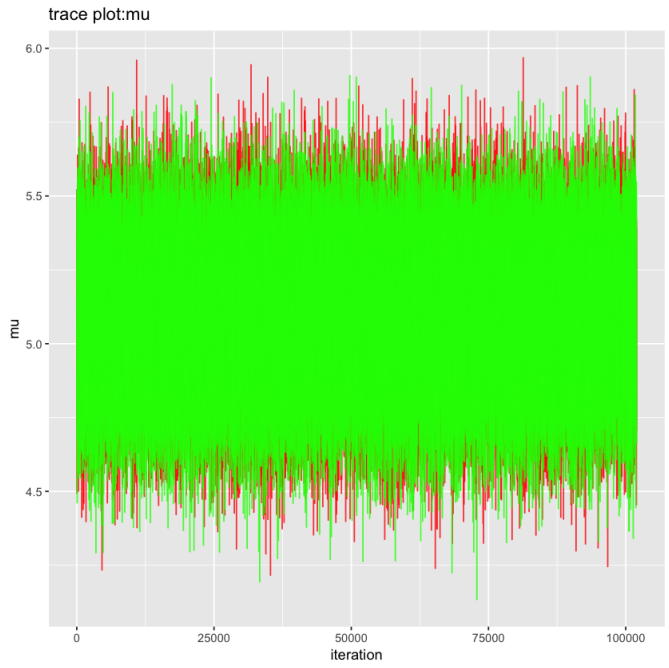
In [3]:

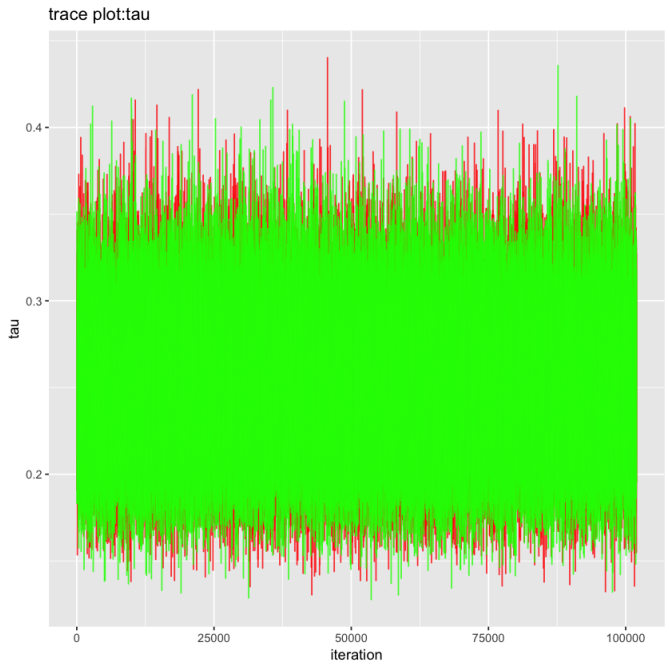
```
set.seed(42)
N = 102000
res1 = gibbs(N, x, 1000, 0.1)
res2 = gibbs(N, x, 0, 999)

df = data.frame(x = 1:N, mu1=res1$mu, tau1=res1$tau, mu2=res2$mu, tau2=res2$tau)

(ggplot(df, aes(x)) + geom_line(aes(y=mu1), colour="red", alpha=0.8) +
  geom_line(aes(y=mu2), colour="green", alpha=0.8) +
  xlab('iteration') + ylab('mu')+ggtitle('trace plot:mu'))


(ggplot(df, aes(x)) + geom_line(aes(y=tau1), colour="red", alpha=0.8) +
  geom_line(aes(y=tau2), colour="green", alpha=0.8) +
  xlab('iteration') + ylab('tau')+ggtitle('trace plot:tau'))
```


trace plot:mu
6.0
5.5
5.0
4.5
mu
0
25000
50000
75000
100000
iteration




(b) They do converge together

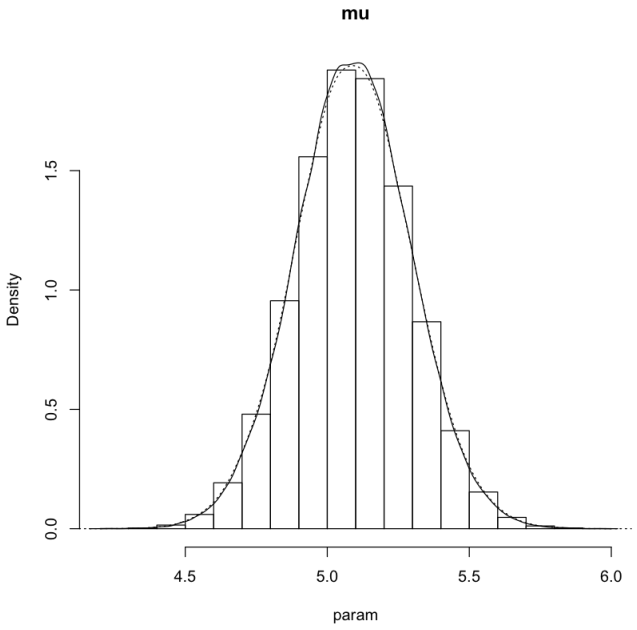
**(c) Distribution/histogram/quantile**

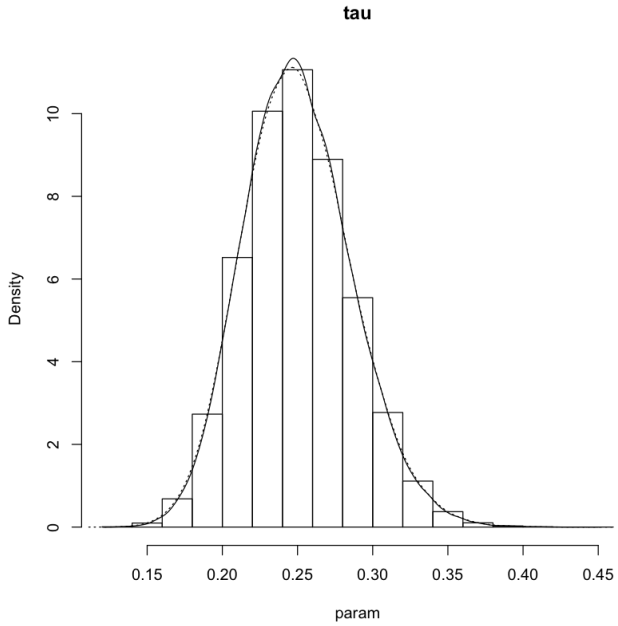
In [4]:

```
mu = sort(df$mu1[c(2000:length(df$mu1))])
tau = sort(df$tau1[c(2000:length(df$tau1))])

output(mu, 'mu')
output(tau, 'tau')
```

```
[1] "mu"
[1] "mean value:"
[1] 5.090795
[1] "90%  CI:"
       5%      95%
4.757753 5.422978
[1] "tau"
[1] "mean value:"
[1] 0.2503675
[1] "90%  CI:"
       5%       95%
0.1949002 0.3116865
```






## Q2

In [5]:

```
logpiparam <- function(mu, tau, data) {
  prior = log(1 / tau)
  likelihood = sum(dnorm(data, mu, sqrt(1/tau),log = TRUE))
  return(prior + likelihood)
}


mh <- function(n, data, mu=1, tau=1) {
  samples = list(mu=numeric(n), tau=numeric(n));
  xbar = mean(data)
  m = length(data)

  for (i in 1:n) {
    tau_ = rgamma(1, 5*tau, 5)
    mu_  = rnorm(1, mu, sqrt(tau_))

    qtop = dnorm(mu, mu_, sqrt(tau), log = TRUE) + dgamma(tau, 5*tau_, 5, log =
TRUE)
    qbot = dnorm(mu_, mu, sqrt(tau_), log = TRUE) + dgamma(tau_, 5*tau, 5, log =
TRUE)
    p.log = qtop + logpiparam(mu_, tau_, data) - (qbot + logpiparam(mu, tau, dat
a))
    if (p.log > log(runif(1))) {
      mu = mu_
      tau = tau_
    }

    samples$mu[i] = mu
    samples$tau[i] = tau
  }
  return(samples)
}
```

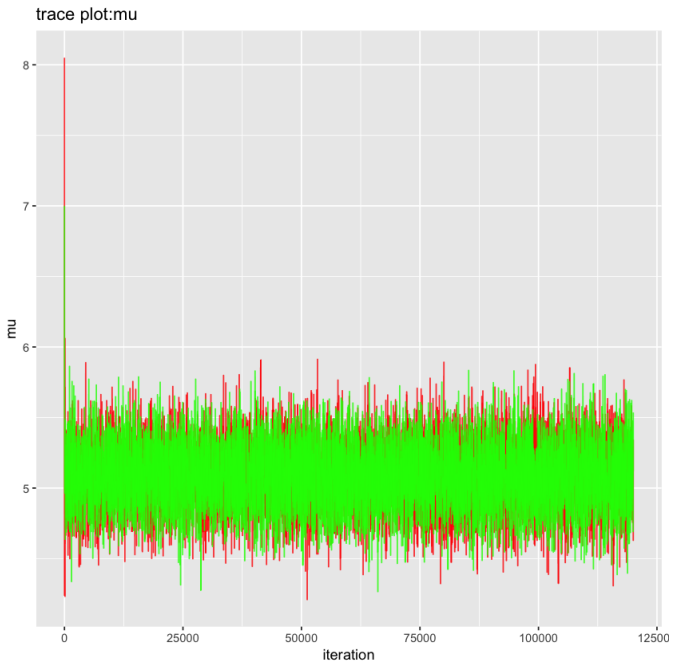
**(b) Trace plot**

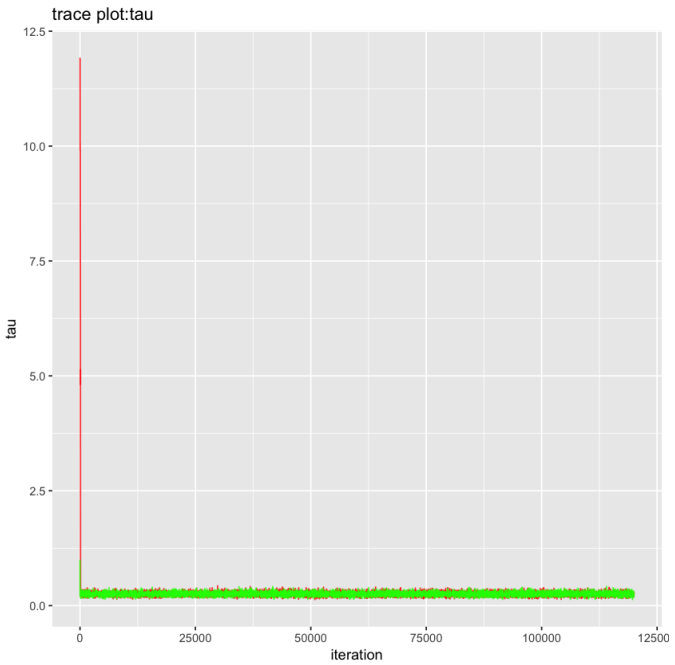
In [6]:

```
set.seed(42)
N = 120000
res1 = mh(N, x, 10, 10)
res2 = mh(N, x, 7, 1)

df = data.frame(x = 1:N, mu1=res1$mu, tau1=res1$tau, mu2=res2$mu, tau2=res2$tau)

(ggplot(df, aes(x)) + geom_line(aes(y=mu1), colour="red", alpha=0.8) +
  geom_line(aes(y=mu2), colour="green", alpha=0.8) +
  xlab('iteration') + ylab('mu')+ggtitle('trace plot:mu'))
(ggplot(df, aes(x)) + geom_line(aes(y=tau1), colour="red", alpha=0.8) +
  geom_line(aes(y=tau2), colour="green", alpha=0.8) +
  xlab('iteration') + ylab('tau')+ggtitle('trace plot:tau'))
```


trace plot:mu
mu
8
7
6
5
0
25000
50000
75000
100000
12500
iteration


trace plot:tau
12.5
10.0
7.5
5.0
2.5
0.0
tau
0
25000
50000
75000
100000
12500
iteration

(b) They do converge together

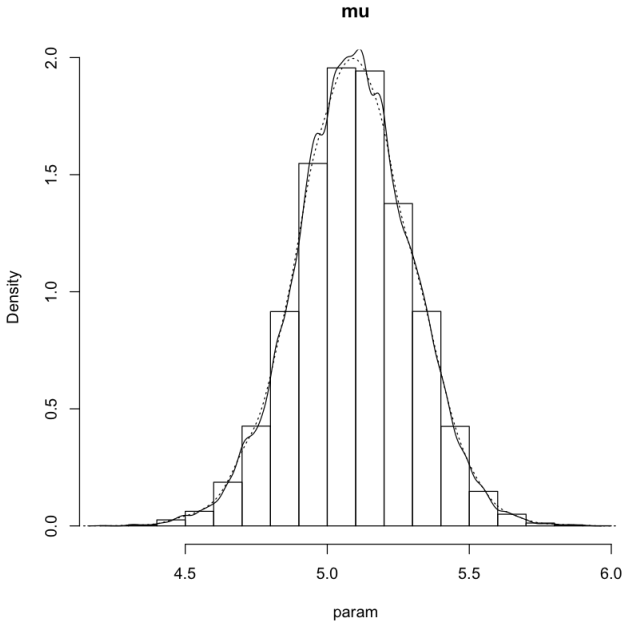
**(c) Distribution/histogram/quantile**

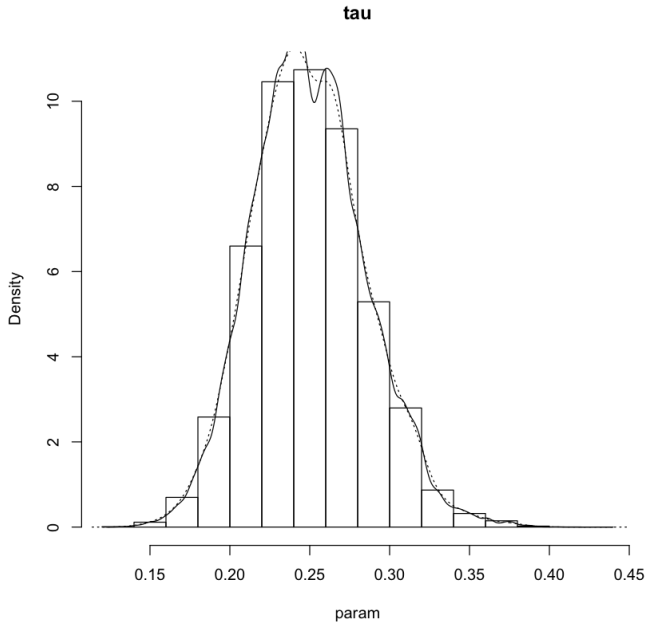
In [7]:

```
mu = sort(df$mu1[c(2000:length(df$mu1))])
tau = sort(df$tau1[c(2000:length(df$tau1))])

output(mu, 'mu')
output(tau, 'tau')
```

```
[1] "mu"
[1] "mean value:"
[1] 5.093511
[1] "90%  CI:"
       5%      95%
4.756718 5.425346
[1] "tau"
[1] "mean value:"
[1] 0.249857
[1] "90%  CI:"
       5%       95%
0.1951379 0.3107686
```




tau
Density
0
2
4
6
8
10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
param