STA 360/602L: Module 3.9

MCMC and Gibbs sampling III

Dr. Olanrewaju Michael Akande

STA 602L

Recap of normal model

- Sampling model:

$$y_i|\mu,\tau \overset{iid}{\sim} \mathcal{N}\left(\mu,\tau^{-1}\right).$$

- Suppose we want to specify our uncertainties about μ and τ independently of each other. That is, we want $\pi(\mu,\tau) = \pi(\mu)\pi(\tau)$.

- For example,

$$\mu \sim \mathcal{N}\left(\mu_0,\sigma_0^2\right).$$
$$\tau \sim \text{Gamma}\left(\frac{\nu_0}{2},\frac{\nu_0}{2\tau_0}\right).$$

- Then in this form, where σ_0^2 is not proportional to $\dfrac{1}{\tau}$, the marginal density of τ is not a gamma density (or a density we can easily sample from).

- We need to do Gibbs sampling.

STA 602L

Full conditionals

- That is, we need

$$\mu | Y, \tau \sim \mathcal{N}(\mu_n, \tau_n^{-1}),$$

where

$$\mu_n = \frac{\frac{\mu_0}{\sigma_0^2} + n\tau\bar{y}}{\frac{1}{\sigma_0^2} + n\tau}$$

$$\tau_n = \frac{1}{\sigma_0^2} + n\tau.$$

STA 602L

Full conditionals

- and

$$\tau|\mu, Y \sim \text{Gamma}\left(\frac{\nu_n}{2}, \frac{\nu_n \sigma_n^2(\mu)}{2}\right),$$

where

$$\nu_n = \nu_0 + n$$

$$\sigma_n^2(\mu) = \frac{1}{\nu_n}\left[\frac{\nu_0}{\tau_0} + \sum_{i=1}^{n}(y_i - \mu)^2\right] = \frac{1}{\nu_n}\left[\frac{\nu_0}{\tau_0} + n s_n^2(\mu)\right]$$

$$\text{with} \quad s_n^2(\mu) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \mu)^2 \quad \Rightarrow \quad n s_n^2(\mu) = (n-1)s^2 + n(\bar{y} - \mu)^2.$$

STA 602L

Recall the Pygmalion data

- Let's implement this Gibbs sampler for the Pygmalion data.
- For now, let's focus only on the accelerated group.
- Data for accelerated group (A): 20, 10, 19, 15, 9, 18.
- Summary statistics: $\bar{y}_A = 15.2$; $s_A = 4.71$.

STA 602L

Recall the Pygmalion data

- Suppose we assume, as we did before, that these improvement scores are normal with mean μ and variance $\frac{1}{\tau}$.

- As a reminder, in the conjugate case, we had

$$
\begin{aligned}
\tau &\sim \text{Gamma}\left(\frac{\nu_0}{2}, \frac{\nu_0}{2\tau_0}\right) \\
\mu|\tau &\sim \mathcal{N}\left(\mu_0, \frac{1}{\kappa_0 \tau}\right).
\end{aligned}
$$

- We set
  - $\mu_0 = 0$, to reflect "no idea whether students would improve IQ on average";
  - $\kappa_0 = 1$, to reflect "little faith in this belief, equivalent to having only 1 prior observation in each group";
  - $\nu_0 = 1$ and $1/\tau_0 = 100$, based on literature, that is, SD of change scores should be around 10.

STA 602L

Recall the Pygmalion data

- Now, in the non-conjugate case, we have

$$\mu \sim \mathcal{N}\left(\mu_0, \sigma_0^2\right).$$
$$\tau \sim \text{Gamma}\left(\frac{\nu_0}{2}, \frac{\nu_0}{2\tau_0}\right).$$

- Suppose for μ, we use a $\mathcal{N}(0, 100)$ prior, and for τ we use a $\text{Ga}(\frac{1}{2}, 50)$ prior.

- Matching with

$$\mu \sim \mathcal{N}\left(\mu_0, \sigma_0^2\right).$$
$$\tau \sim \text{Gamma}\left(\frac{\nu_0}{2}, \frac{\nu_0}{2\tau_0}\right),$$

we have: $\mu_0 = 0$, $\sigma_0^2 = 100$, $\nu_0 = 1$ and $\tau_0 = 1/100$.

STA 602L

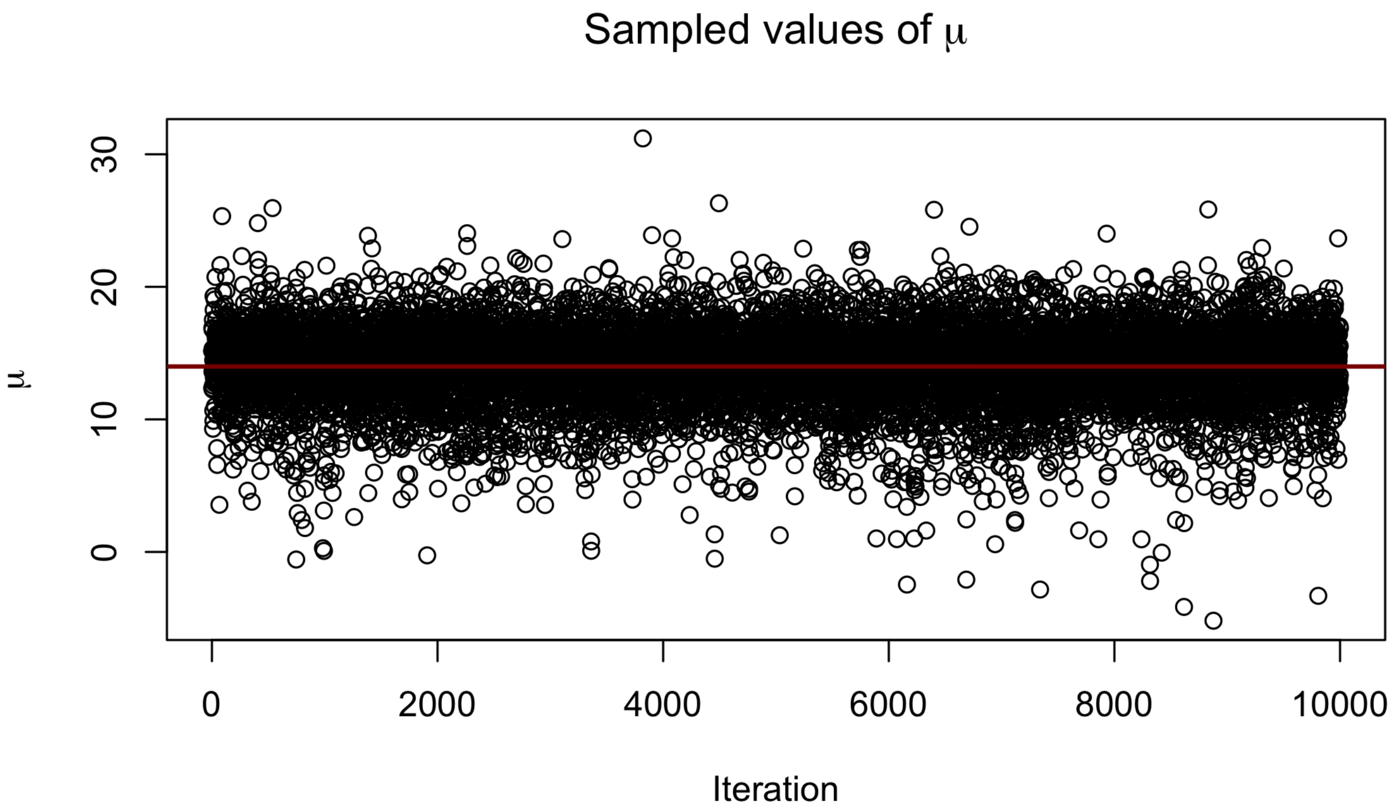
Gibbs sampling for Pygmalion data

```
y <- c(20,10,19,15,9,18) #data
y_bar <- mean(y); s2 <- var(y); n <- length(y) #sample statistics you'll need

S <- 10000 # number of samples to draw

PHI <- matrix(nrow=S,ncol=3); #matrix to save results
colnames(PHI) <- c("mu","tau","sigma2")
PHI[1,] <- phi <- c(y_bar,1/s2,s2) #starting values are MLEs

mu0 <- 0; sigma02 <- 100; nu0 <- 1; tau0 <- 1/100 #hyperparameters


###### Gibbs sampler ######
set.seed(1234) #to replicate results exactly
for(s in 2:S) {
#first, draw new mu
taun <- 1/sigma02 + n*phi[2]
mun <- (mu0/sigma02 + n*y_bar*phi[2])/taun
phi[1] <- rnorm(1,mun,sqrt(1/taun))

#now, draw new tau/sigma2
nun <- nu0+n
#trick to speed up calculation of sum(y_i-\mu)^2
s2nmu <- (nu0/tau0 + (n-1)*s2 + n*(y_bar-phi[1])^2)/nun
phi[2] <- rgamma(1,nun/2,nun*s2nmu/2)
phi[3] <- 1/phi[2] #save sigma2

#save the current joint draws
PHI[s,] <- phi
}
###### End of Gibbs sampler ######
```

STA 602L

Pygmalion data: mean

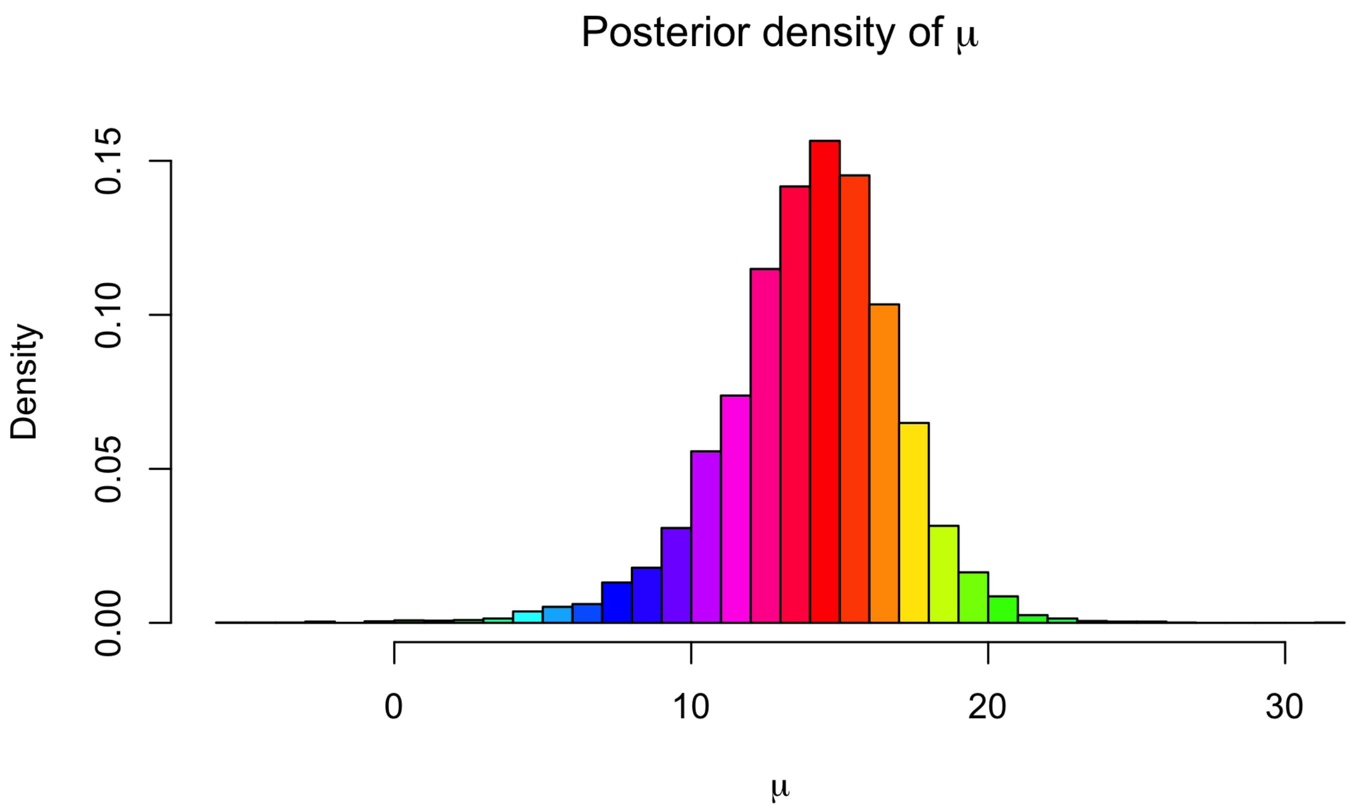
```r
plot(PHI[,1],ylab=expression(mu),xlab="Iteration",
     main=expression(paste("Sampled values of ",mu)))
abline(a=mean(PHI[,1]),b=0,col="red4",lwd=2)
```



STA 602L

Pygmalion data: mean

```r
hist(PHI[,1],col=rainbow(20),xlab=expression(mu),ylab="Density",freq=F,breaks=50,
     main=expression(paste("Posterior density of ",mu)))
```



Pygmalion data: mean

```
round(mean(PHI[,1]),3)
```

```
## [1] 13.99
```

```
round(quantile(PHI[,1],c(0.025,0.5,0.975)),3)
```

```
##   2.5%    50%  97.5%
##  7.520 14.217 19.277
```

Posterior summaries for μ:

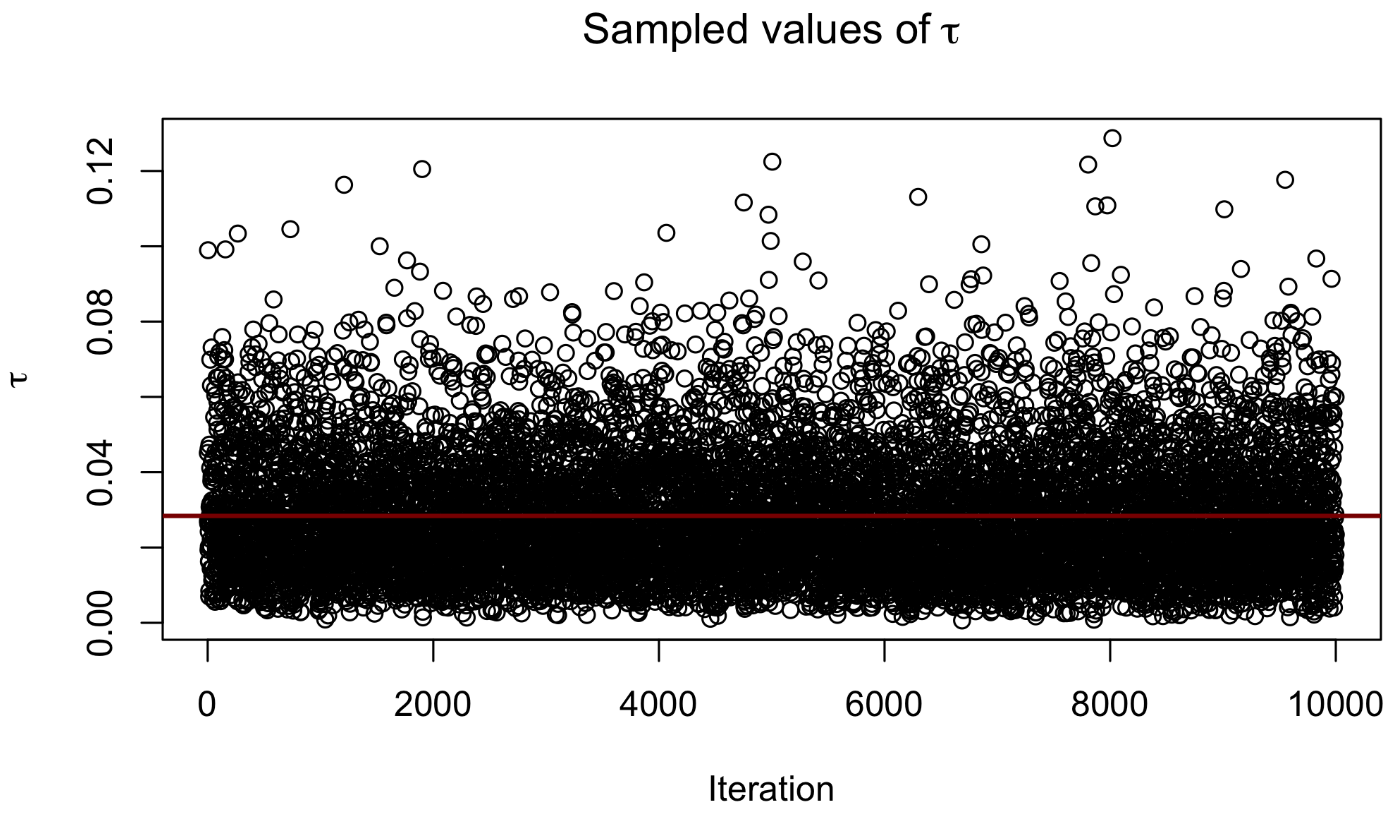
- Posterior mean ≈ 14.
- Posterior median ≈ 14.22.
- 95% credible interval $\approx (7.52, 19.28)$.

For context, $\bar{y}_A = 15.2$, and we used a $\mathcal{N}(0, 100)$ prior for μ.

STA 602L

Pygmalion data: precision

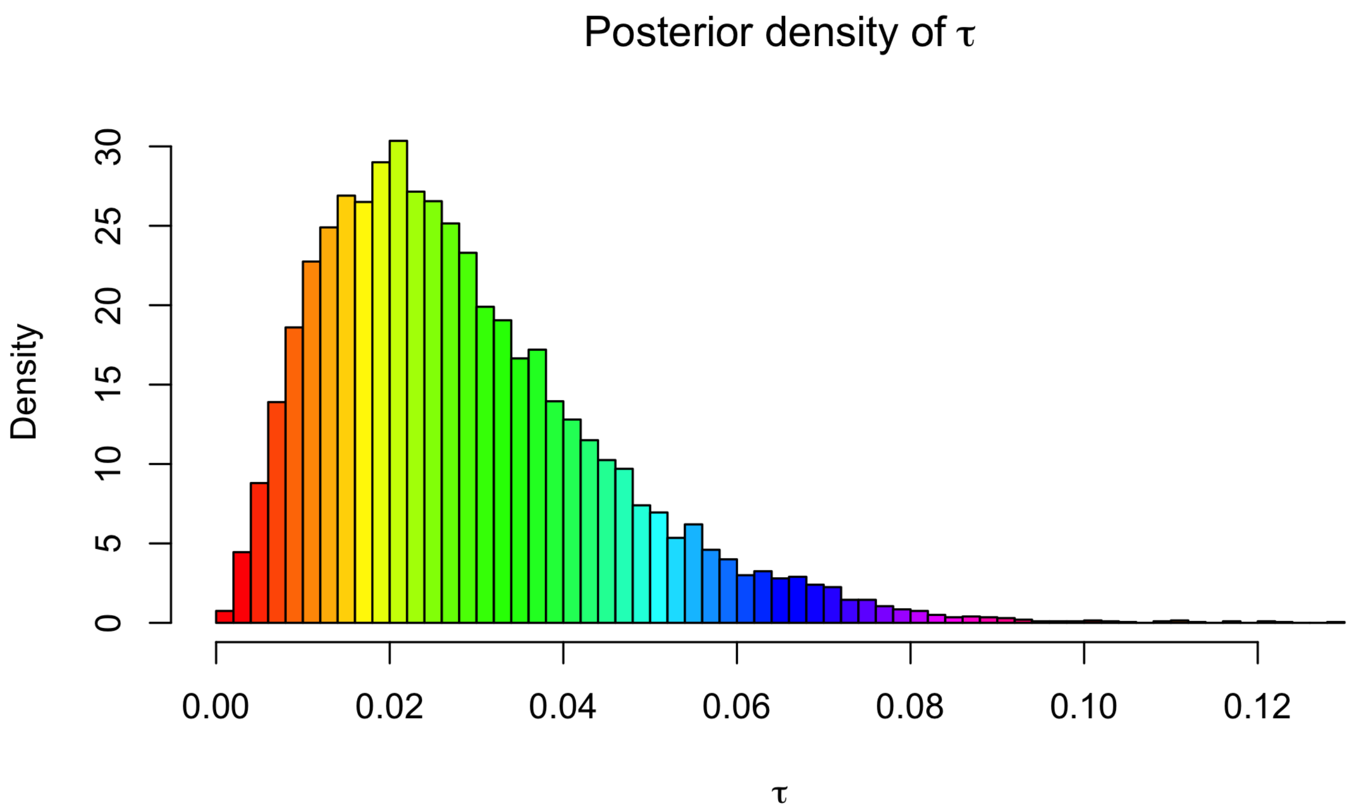
```
plot(PHI[,2],ylab=expression(tau),xlab="Iteration",
     main=expression(paste("Sampled values of ",tau)))
abline(a=mean(PHI[,2]),b=0,col="red4",lwd=2)
```



STA 602L

Pygmalion data: precision

```
hist(PHI[,2],col=rainbow(50),xlab=expression(tau),ylab="Density",freq=F,breaks=50,
     main=expression(paste("Posterior density of ",tau)))
```



Pygmalion data: precision

```
round(mean(PHI[,2]),3)
```

```
## [1] 0.028
```

```
round(quantile(PHI[,2],c(0.025,0.5,0.975)),3)
```

```
##  2.5%   50% 97.5%
## 0.006 0.025 0.069
```

Posterior summaries for τ:

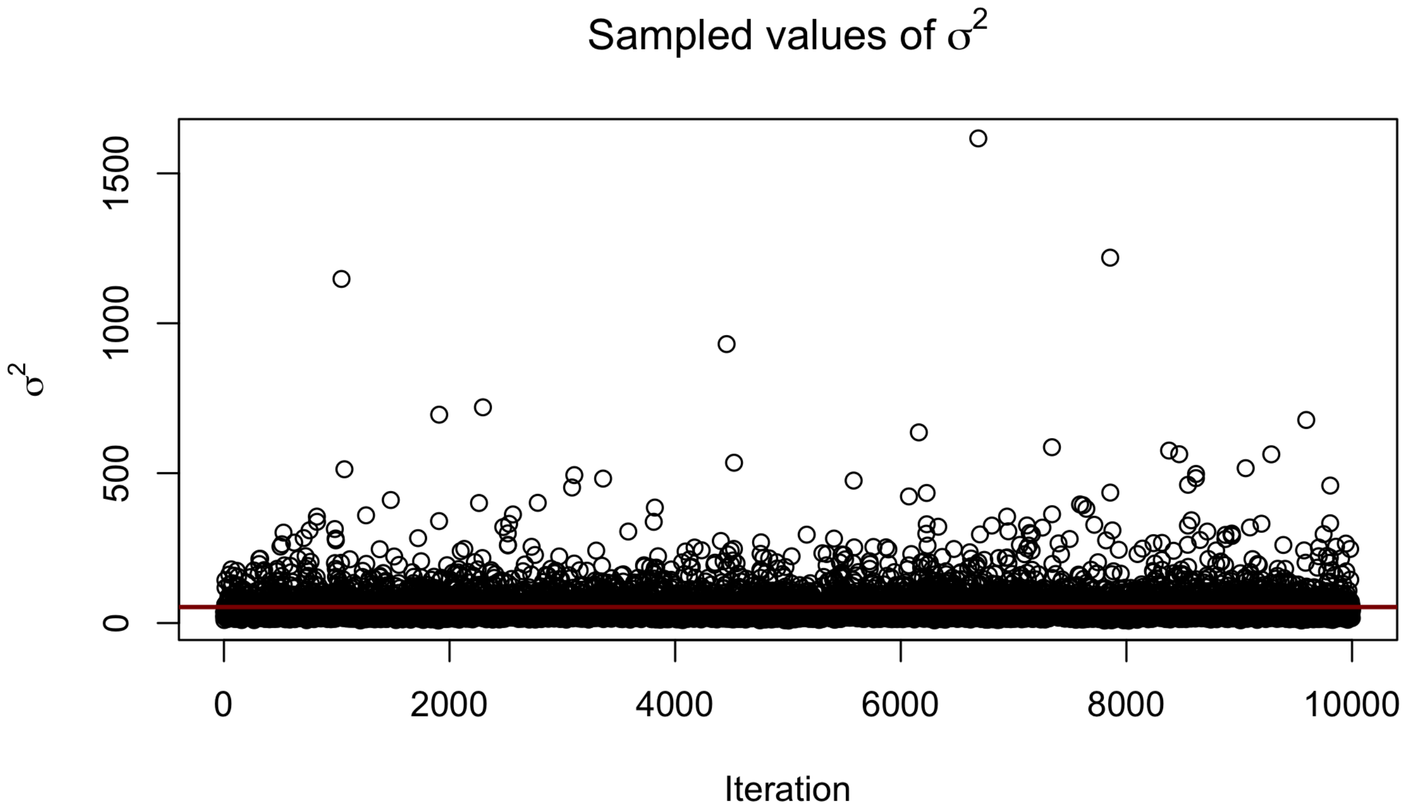
- Posterior mean ≈ 0.028.
- Posterior median ≈ 0.025.
- 95% credible interval $\approx (0.006, 0.069)$.

For context, $s_A = 4.71$, which means sample precision $= 1/4.71^2 = 0.045$. Also, we used a $\text{Ga}(\frac{1}{2}, 50)$ prior for τ.

STA 602L

Pygmalion data: variance

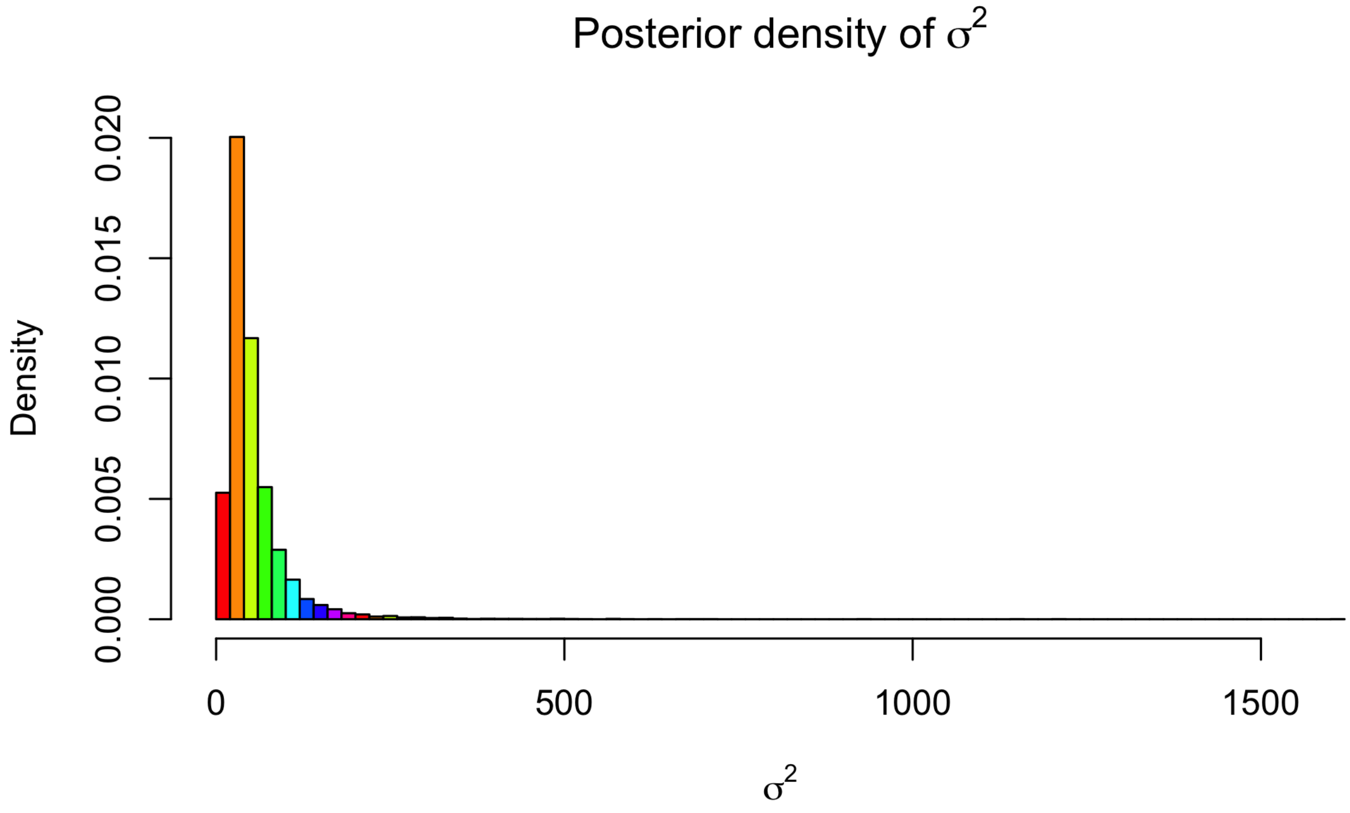
```
plot(PHI[,3],ylab=expression(sigma^2),xlab="Iteration",
     main=expression(paste("Sampled values of ",sigma^2)))
abline(a=mean(PHI[,3]),b=0,col="red4",lwd=2)
```



STA 602L

Pygmalion data: variance

```
hist(PHI[,3],col=rainbow(10),xlab=expression(sigma^2),ylab="Density",freq=F,breaks=100,
     main=expression(paste("Posterior density of ",sigma^2)))
```



Pygmalion data: variance

```
round(mean(PHI[,3]),2)
```

```
## [1] 53.34
```

```
round(quantile(PHI[,3],c(0.025,0.5,0.975)),2)
```

```
##   2.5%    50%  97.5%
##  14.52  39.60 174.11
```

Posterior summaries for σ^2:

- Posterior mean $= 53.34$.
- Posterior median $= 39.60$.
- 95% credible interval $= (14.52, 174.11)$.

For context, $s_A = 4.71$, which means sample variance $4.71^2 = 22.18$. Again, we used a $\text{Ga}(\frac{1}{2}, 50)$ prior for τ.

STA 602L

What's next?

Move on to the readings for the next module!

STA 602L