STA360 HW5

Ken Ye

#1

a) $p(\theta \mid \underset{\sim}{y}) = p(\underset{\sim}{y} \mid \theta)\, p(\theta) / p(\underset{\sim}{y})$

$$\propto p(\underset{\sim}{y} \mid \theta)\, p(\theta)$$

$$= \prod_{i=1}^{n} \frac{1}{y_i!} \theta^{y_i} e^{-\theta} \frac{b^a}{\Gamma(a)} \theta^{a-1} e^{-b\theta}$$

$$\propto \theta^{\Sigma y_i} e^{-n\theta} \theta^{a-1} e^{-b\theta}$$

$$= \theta^{\Sigma y_i + a - 1} e^{-(b+n)\theta}$$

$$\sim \text{gamma}(a + \Sigma y_i,\ b+n)$$

$$E[\theta \mid \underset{\sim}{y}] = \frac{a + \Sigma y_i}{b+n}$$

$$= \frac{a}{b+n} + \frac{\Sigma y_i}{b+n}$$

$$= \frac{a}{b} \cdot \frac{b}{b+n} + \frac{\Sigma y_i}{n} \cdot \frac{n}{b+n}$$

$$= \frac{n}{b+n} \frac{\Sigma y_i}{n} + \frac{b}{b+n} \frac{a}{b}$$

$$= w \bar{y} + (1-w)\mu$$

where $w = \frac{n}{b+n}$ and $\mu = \frac{a}{b}$

b) $\mathbb{E}[\bar{y}|\theta] = \theta$

$$V[\bar{y}|\theta] = V\left[\frac{1}{n}\sum y_i \,|\theta\right]$$
$$= \frac{1}{n^2} V\left[\sum y_i \,|\theta\right]$$
$$= \frac{1}{n^2}\, n\theta \quad (\text{since } Y_1 \ldots Y_n \text{ are independent})$$
$$= \frac{\theta}{n}$$

$$\mathbb{E}[\hat{\theta}_w|\theta] = \mathbb{E}[w\bar{y} + (1-w)\mu], \text{ where } w = \frac{n}{n+b},\ \mu = \frac{a}{b}$$
$$= w\,\mathbb{E}[\bar{y}] + (1-w)\mu$$
$$= w\theta + (1-w)\mu$$

$$V[\hat{\theta}_w|\theta] = V[w\bar{y} + (1-w)\mu]$$
$$= w^2 V[\bar{y}]$$
$$= \frac{w^2\theta}{n}$$

c) $MSE[\hat{\theta}_w] < MSE[\bar{y}]$

$$V[\hat{\theta}_w] + Bias^2[\hat{\theta}_w] < V[\bar{y}] + Bias^2[\bar{y}]$$
$$\frac{w^2\theta}{n} + [w\theta + (1-w)\mu - \theta]^2 < \frac{\theta}{n} + (\theta - \theta)^2$$
$$\frac{w^2\theta}{n} + (\mu-\theta)^2(1-w)^2 < \frac{\theta}{n}$$
$$(\mu-\theta)^2 < \frac{1-w^2}{(1-w)^2}\frac{\theta}{n}$$
$$(\mu-\theta)^2 < \frac{(1+w)(1-w)}{(1-w)^2}\frac{\theta}{n}$$
$$\boxed{(\mu-\theta)^2 < \left(\frac{1+w}{1-w}\right)\frac{\theta}{n}} \Rightarrow \text{Condition s.t. } MSE[\hat{\theta}_w] < MSE[\bar{y}]$$

#2

a) $Pr(Y=y_l) = P_y(y_l) = \sum_{k=1}^{K} Pr(X=x_k, Y=y_l) = \sum_{k=1}^{K} P_{xy}(x_k, y_l)$

$Pr(X=x_k) = P_x(x_k) = \sum_{l=1}^{L} Pr(X=x_k, Y=y_l) = \sum_{l=1}^{L} P_{xy}(x_k, y_l)$

$Pr(X=x_k | Y=y_l) = P_{x|y}(x_k|y_l) = \dfrac{P_{xy}(x_k, y_l)}{P_y(y_l)}$

b) In the simulation, $Pr(X^* = x_k) = Pr(X=x_k | Y=Y^*)$

$= \sum_{l=1}^{L} Pr(X=x_k | Y=y_l) Pr(Y=y_l)$

$= \sum_{l=1}^{L} \dfrac{P_{xy}(x_k, y_l)}{\cancel{P_y(y_l)}} \cancel{P_y(y_l)}$

$= \sum_{l=1}^{L} P_{xy}(x_k, y_l)$

$= P_x(x_k)$

3

b) $p(\theta|\gamma, y_A, y_B) = \dfrac{p(y_A, y_B|\theta,\gamma)\, p(\theta,\gamma)}{p(\gamma, y_A, y_B)}$

$$\underset{\theta}{\propto} p(y_A, y_B|\theta,\gamma)\, p(\theta,\gamma)$$

$$\underset{\theta}{\propto} p(y_A|\theta,\gamma)\, p(y_B|\theta,\gamma)\, p(\theta)$$ since y_A and y_B are independent, and θ and γ are independent

$$= \prod_{i=1}^{n_A} \frac{\theta_A^{y_i}}{y_i!} e^{-\theta_A} \prod_{j=1}^{n_B} \frac{\theta_B^{y_j}}{y_j!} e^{-\theta_B} \frac{b_\theta^{a_\theta}}{\Gamma(a_\theta)} \theta^{a_\theta - 1} e^{-b_\theta \theta}$$

$$\underset{\theta}{\propto} \theta^{\Sigma y_i} e^{-n_A\theta} (\theta\gamma)^{\Sigma y_j} e^{-n_B\theta\gamma} \theta^{a_\theta - 1} e^{-b_\theta\theta}$$ since $\theta_A = \theta$; $\theta_B = \theta\gamma$.

$$\underset{\theta}{\propto} \theta^{\Sigma y_i + \Sigma y_j + a_\theta - 1} e^{-(n_A + n_B\gamma + b_\theta)\theta}$$

$$\sim \text{gamma}\left(\sum_{i=1}^{n_A} y_i + \sum_{j=1}^{n_B} y_j + a_\theta,\ n_A + n_B\gamma + b_\theta\right)$$

Therefore, $\mathbb{E}[\theta|\gamma, y_A, y_B] = \dfrac{\sum_{i=1}^{n_A} y_i + \sum_{j=1}^{n_B} y_j + a_\theta}{n_A + n_B\gamma + b_\theta}$

Scanned with CamScanner

b) continued

$$P(\gamma \mid \theta, y_A, y_B) = \frac{P(y_A, y_B \mid \theta, \gamma)\, P(\theta, \gamma)}{P(\theta, y_A, y_B)}$$

$$\underset{\gamma}{\propto} \; P(y_A \mid \theta, \gamma)\, P(y_B \mid \theta, \gamma)\, P(\gamma)$$ since y_A, y_B indep. and θ, γ indep.

$$= \prod_{i=1}^{n_A} \frac{\theta_A^{y_i}}{y_i!} e^{-\theta_A} \prod_{j=1}^{n_B} \frac{\theta_B^{y_j}}{y_j!} e^{-\theta_B} \; P(\gamma)$$

$$\underset{\gamma}{\propto} \; \theta^{\Sigma y_i} e^{-n_A\theta} (\theta\gamma)^{\Sigma y_j} e^{-n_B\theta\gamma} \gamma^{a_\gamma - 1} e^{-b_\gamma \gamma}$$ since $\theta_A = \theta$, $\theta_B = \theta\gamma$

$$\underset{\gamma}{\propto} \; \gamma^{\Sigma y_j + a_\gamma - 1} e^{-(n_B\theta + b_\gamma)\gamma}$$

$$\sim \text{gamma}\left(\sum_{j=1}^{n_B} y_j + a_\gamma ,\; n_B\theta + b_\gamma \right)$$

Therefore, $\mathbb{E}[\gamma \mid \theta, y_A, y_B] = \dfrac{\sum_{j=1}^{n_B} y_j + a_\gamma}{n_B\theta + b_\gamma}$

STA360 Homework 5 (Ken Ye)

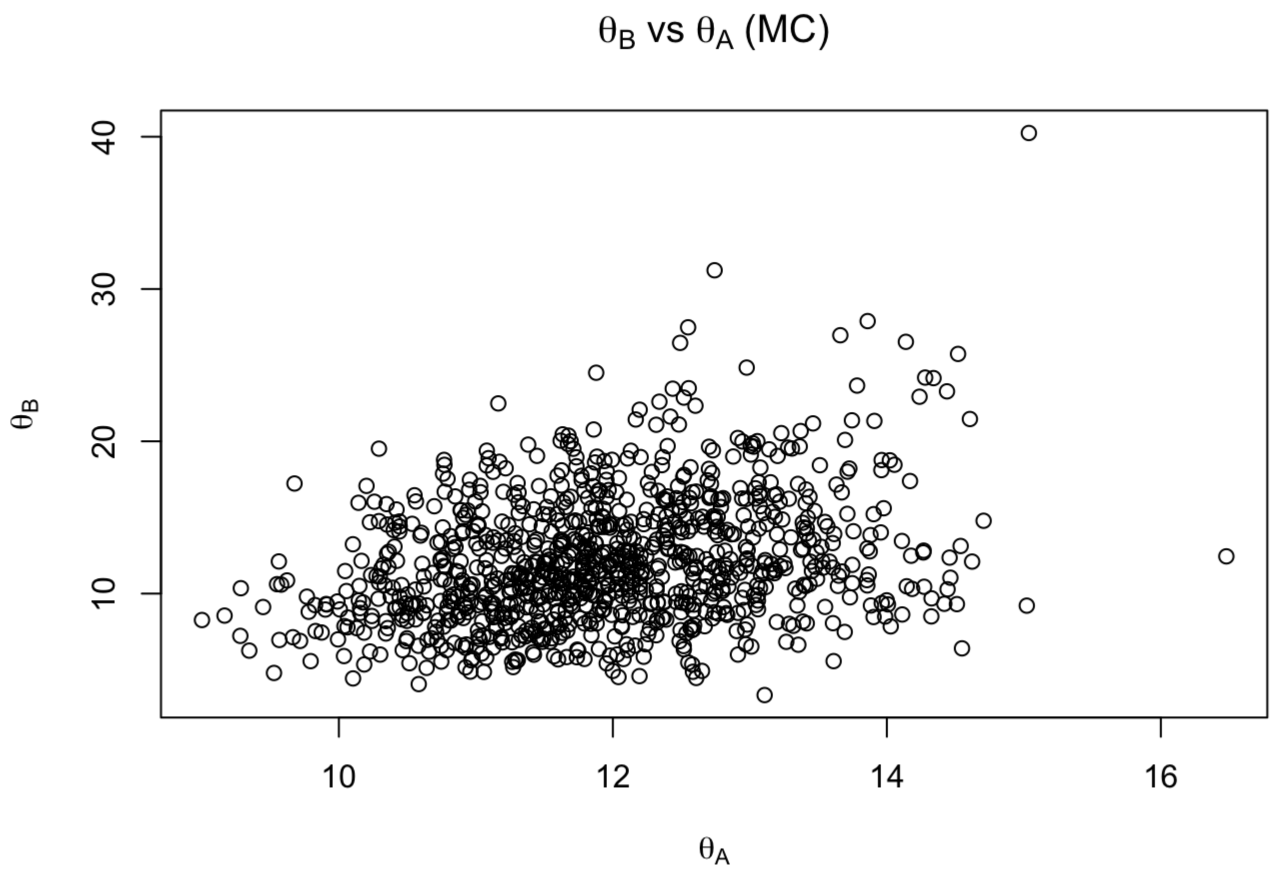
```
library(latex2exp)
set.seed(0)
```

#3

a.

```
# data
ya <- c(12, 9, 12, 14, 13, 13, 15, 8, 15, 6)
na <- length(ya)
yb <- c(11, 11, 10, 9, 9, 8, 7, 10, 6, 8, 8, 9, 7)
nb <- length(yb)
```

```
mc.size <- 1000
thetas <- rgamma(mc.size, 120, 10)
gammas <- rgamma(mc.size, 10, 10)
thetas.a <- thetas
thetas.b <- thetas * gammas
plot(thetas.a, thetas.b,
     main = TeX('\\theta$_B$ vs \\theta$_A$ (MC)'),
     xlab = TeX('\\theta$_A$'),
     ylab = TeX('\\theta$_B$'))
```




```
# correlation
cor(thetas.a,thetas.b)
```

```
## [1] 0.2985973
```

According to the scatter plot and the correlation calculation, it seems that θ_A and θ_B are dependent, as the graph shows a positive relationship between θ_A and θ_B (as θ_A increases, θ_B generally increases), and their correlation is 0.274, which is nonzero and positive.

c.

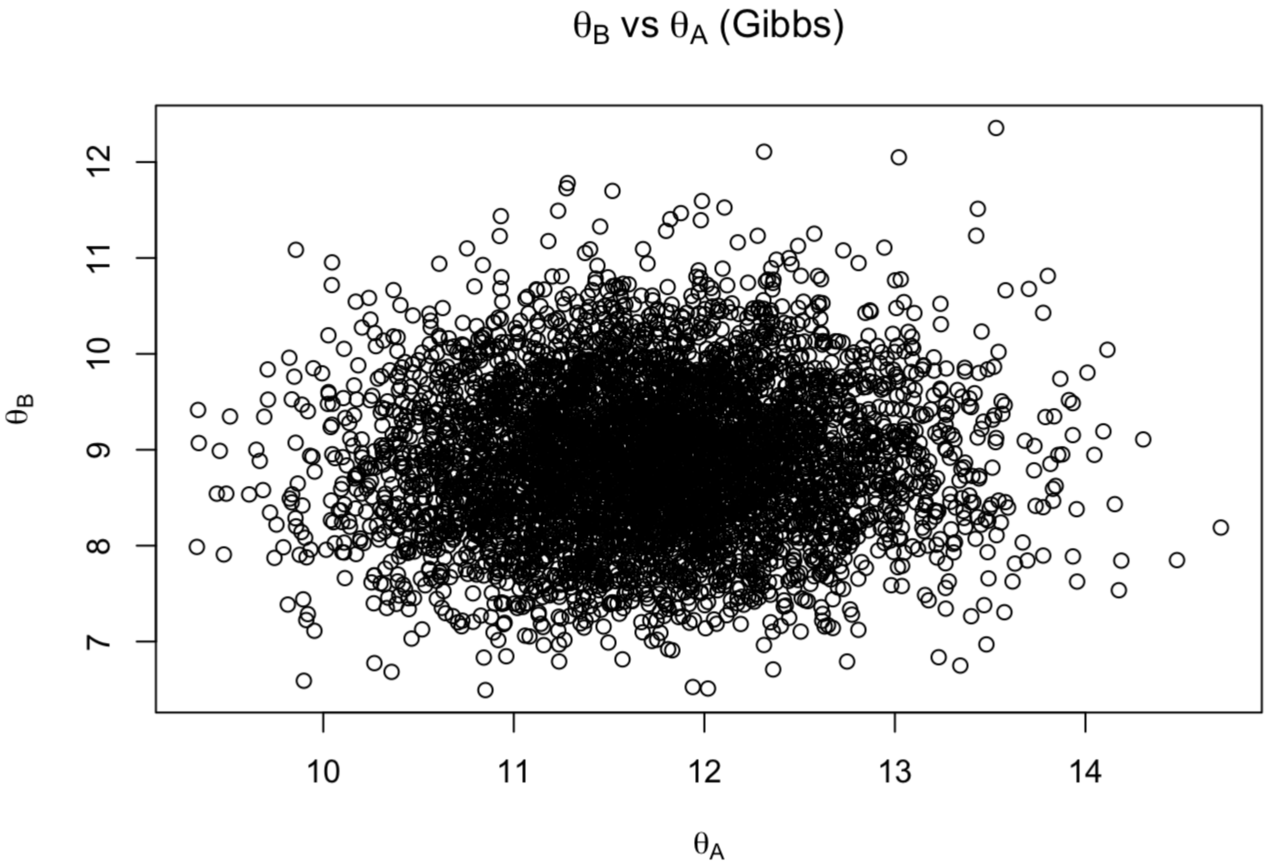
```
# Gibbs sampling
gibbs.size <- 5000
theta0 <- rgamma(1, 120, 10)
gamma0 <- rgamma(1, 10, 10)

THETA <- matrix(nrow = gibbs.size, ncol = 2)
THETA[1,] <- theta <- c(theta0, gamma0)

for (s in 2 : gibbs.size){
theta[1] <- rgamma(1, sum(ya) + sum(yb) + 120, na + nb * theta[2] + 10)
theta[2] <- rgamma(1, sum(yb) + 10, nb * theta[1] + 10)
THETA[s,] <- theta
}

thetas.a2 <- THETA[,1]
thetas.b2 <- THETA[,1] * THETA[,2]
```

```
# plot
plot(thetas.a2, thetas.b2,
     main = TeX('\\theta$_B$ vs \\theta$_A$ (Gibbs)'),
     xlab = TeX('\\theta$_A$'),
     ylab = TeX('\\theta$_B$'))
```




```
# posteior mean
print(mean(thetas.b2 - thetas.a2))
```

```
## [1] -2.81984
```

The MCMC estimate for the posterior mean of $\theta_B - \theta_A$ is -2.81984.

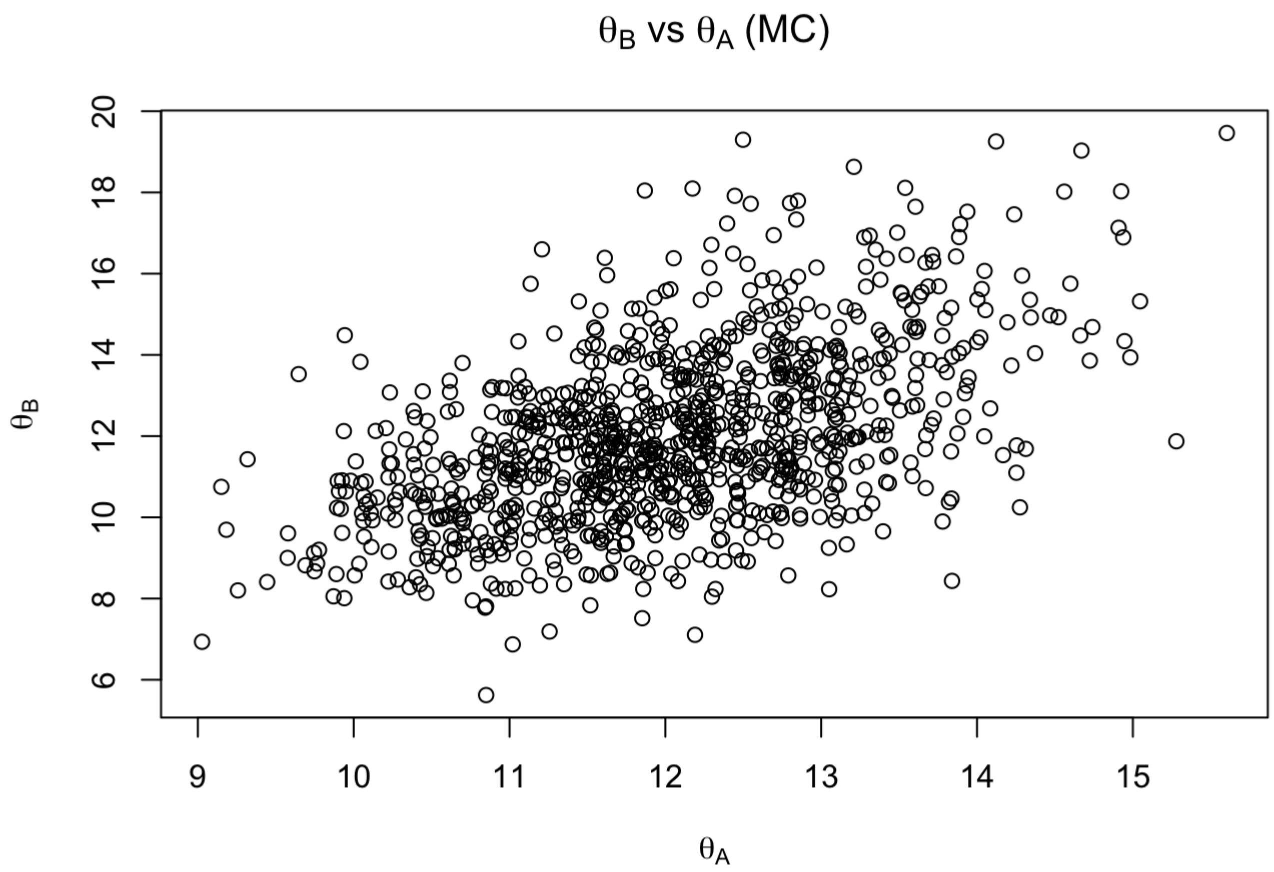
```
# confidence interval
quantile(thetas.b2 - thetas.a2, c(0.025, 0.975))
```

```
##        2.5%      97.5%
## -4.8894486 -0.7468116
```

The MCMC 95% confidence interval for the posterior mean of $\theta_B - \theta_A$ is (-4.8894486, -0.7468116).

d.

```
# mc
mc.size <- 1000
thetas <- rgamma(mc.size, 120, 10)
gammas <- rgamma(mc.size, 45, 45)
thetas.a <- thetas
thetas.b <- thetas * gammas
plot(thetas.a, thetas.b,
     main = TeX('\\theta$_B$ vs \\theta$_A$ (MC)'),
     xlab = TeX('\\theta$_A$'),
     ylab = TeX('\\theta$_B$'))
```




```
# correlation
cor(thetas.a,thetas.b)
```

```
## [1] 0.5207864
```

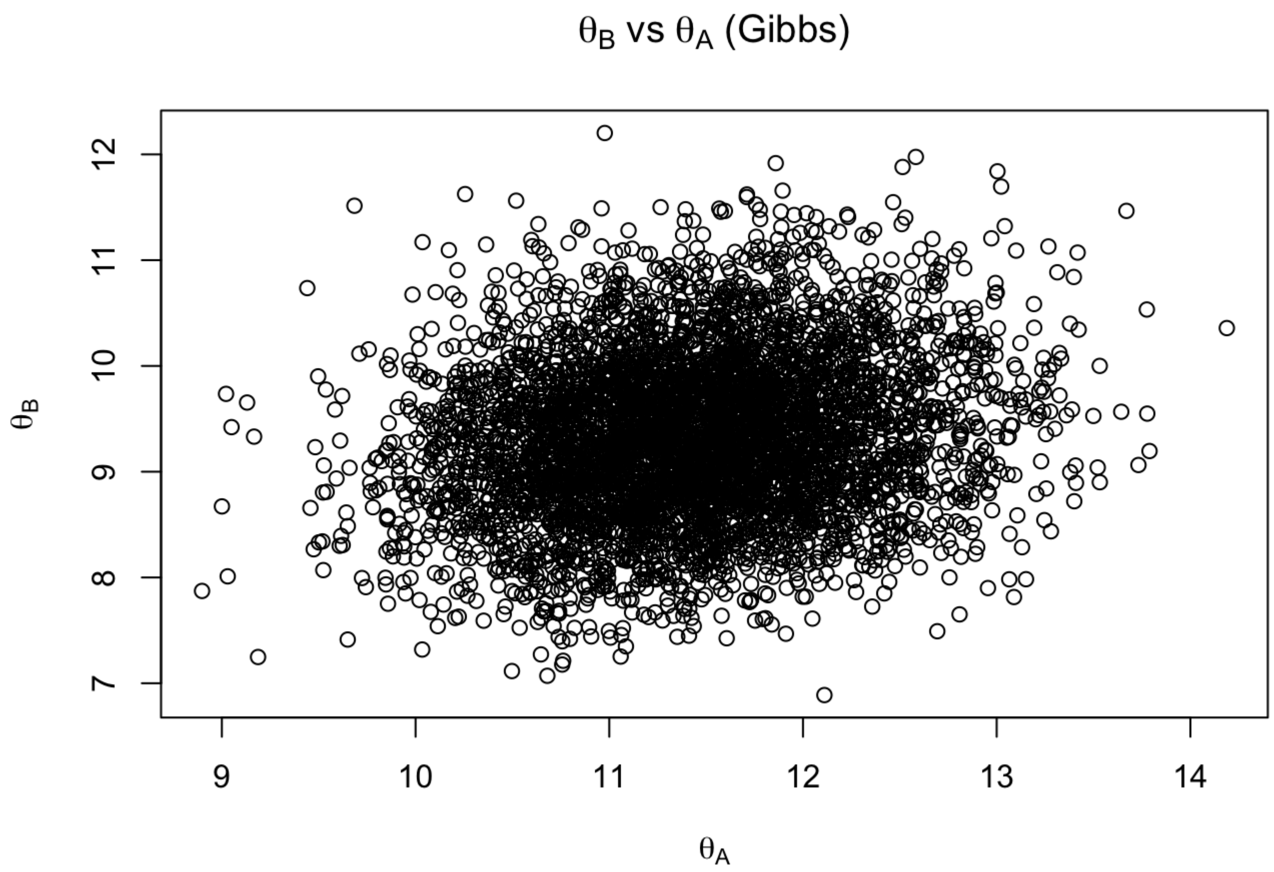
```r
# Gibbs sampling
gibbs.size <- 5000
theta0 <- rgamma(1, 120, 10)
gamma0 <- rgamma(1, 45, 45)

THETA <- matrix(nrow = gibbs.size, ncol = 2)
THETA[1,] <- theta <- c(theta0, gamma0)

for (s in 2 : gibbs.size){
theta[1] <- rgamma(1, sum(ya) + sum(yb) + 120, na + nb * theta[2] + 10)
theta[2] <- rgamma(1, sum(yb) + 45, nb * theta[1] + 45)
THETA[s,] <- theta
}

thetas.a2 <- THETA[,1]
thetas.b2 <- THETA[,1] * THETA[,2]
```

```r
# plot
plot(thetas.a2, thetas.b2,
     main = TeX('\\theta$_B$ vs \\theta$_A$ (Gibbs)'),
     xlab = TeX('\\theta$_A$'),
     ylab = TeX('\\theta$_B$'))
```




```
# posteior mean
print(mean(thetas.b2 - thetas.a2))
```

```
## [1] -2.103156
```

The MCMC estimate for the posterior mean of $\theta_B - \theta_A$ is -2.103156.

```
# confidence interval
quantile(thetas.b2 - thetas.a2, c(0.025, 0.975))
```

```
##       2.5%      97.5% 
## -3.9479005 -0.2856949
```

The MCMC 95% confidence interval for the posterior mean of $\theta_B - \theta_A$ is (-3.9479005, -0.2856949).

After changing the prior of γ to gamma(45,45), in the MC simulation, the graph still shows a stronger positive relationship between θ_A and θ_B (as θ_A increases, θ_B generally increases), and their correlation is larger (0.521). In addition, according to the graph, the distribution is more concentrated, indicating lower variance.

On the other hand, in the Gibbs sampling simulation, the MCMC estimate of $\theta_B - \theta_A$ increases from -2.81984 to -2.103156, and the 95% confidence interval changes from (-4.8894486, -0.7468116) to (-3.9479005, -0.2856949). In addition, according to the graph, similar to what we observe in MC, the distribution is more concentrated, indicating lower variance.