1. (Wasserman, 2003) A random variable Z has an *inverse Gaussian distribution* if it has density

$$f(z \mid \theta_1, \theta_2) \propto z^{-3/2} \exp\left\{-\theta_1 z - \frac{\theta_2}{z} + 2\sqrt{\theta_1\theta_2} + \log(\sqrt{2\theta_2})\right\}, \quad z > 0,$$

where $\theta_1 > 0$ and $\theta_2 > 0$ are parameters. It can be shown that $E(Z) = \sqrt{\theta_2/\theta_1}$ and $E(1/Z) = \sqrt{\theta_1/\theta_2} + 1/(2\theta_2)$.

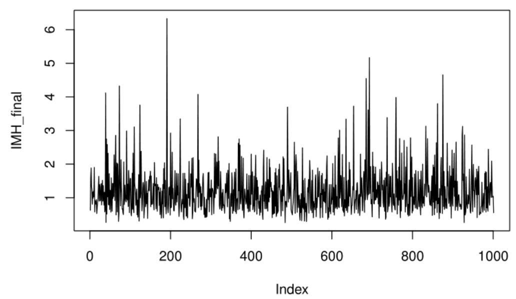
(a) Let $\theta_1 = 1.5$ and $\theta_2 = 2$. Draw a sample of size 1,000 using the independence-Metropolis-Hastings method with a Gamma distribution as the proposal density (note that in an independence-Metropolis-Hastings $q(z^* \mid z) = q(z^*)$). To assess the accuracy of the method, compare the mean of Z and $1/Z$ from the sample to the theoretical means. Try different Gamma distributions to see if you can get an accurate sample.

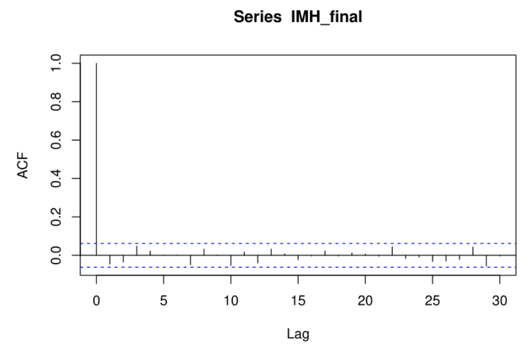
(b) Draw a sample of size 1,000 using the random-walk Metropolis method. Since $z > 0$ we cannot just use a Normal density. Let $W = \log(Z)$. Find the density of W. Use the random-walk Metropolis method to get a sample $W_1, \ldots, W_M$ and let $Z_i = e^{W_i}$. Assess the accuracy of the simulation as in the previous part.

(a) I will calculate the transform probability first:

$$p = \min\left\{1, \frac{f(z^*)\cdot q(z \mid z^*)}{f(z)\cdot q(z^* \mid z)}\right\}$$

Since it's independence MH, $q(z \mid z^*) = q(z)$ and $q(z^* \mid z) = q(z^*)$. After burn-in samples and thinning the chain, I got a trace plot as follows with a ACF plot correspondingly.






```r
mean(IMH_final)
mean(1/IMH_final)
sqrt(theta_2/theta_1)
sqrt(theta_1/theta_2) + 1/(2*theta_2)
```

```
[1] 1.149265
[1] 1.113082
[1] 1.154701
[1] 1.116025
```

It's similar to the true mean, success!

Ps: I tried other distributions like gamma(5,5), but since this is an independent MH, the proposal should be similar to the true distribution. I draw the density of f(x) and approximately used a Gamma(1,1).

Idea: Can we use Laplace approximation in the case of IMH? It's a good way, but I'm not sure, I just came out this idea.

(b) Draw a sample of size 1,000 using the random-walk Metropolis method. Since $z > 0$ we cannot just use a Normal density. Let $W = \log(Z)$. Find the density of W. Use the random-walk Metropolis method to get a sample $W_1, \ldots, W_M$ and let $Z_i = e^{W_i}$. Assess the accuracy of the simulation as in the previous part.

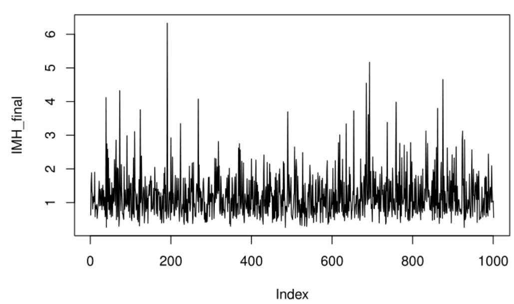
$$W = \log(z) \Rightarrow z = e^w \qquad \frac{dz}{dw} = e^w$$

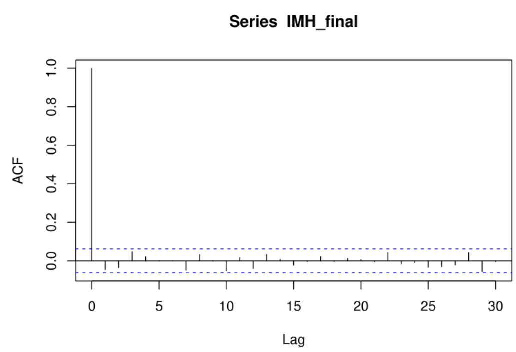
$$f_W \propto (e^w)^{-\frac{3}{2}} \exp\left(-\theta_1 \cdot e^w - \frac{\theta_2}{e^w} + 2\sqrt{\theta_1\theta_2} + \log(\sqrt{2\theta_2})\right) \cdot e^w$$

$$= e^{-\frac{w}{2}} \exp\left(-\theta_1 \cdot e^w - \frac{\theta_2}{e^w} + 2\sqrt{\theta_1\theta_2} + \log(\sqrt{2\theta_2})\right)$$

And $q(z^* | z) \propto (\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{(z^* - z)^2}{2\sigma^2}\right)$

$$\rho = \min\left\{1, \frac{f(z^*)}{f(z)}\right\}$$






```{r}
mean(RMH_final)
mean(1/RMH_final)
sqrt(theta_2/theta_1)
sqrt(theta_1/theta_2) + 1/(2*theta_2)
```

```
[1] 1.594363
[1] 1.454806
[1] 1.154701
[1] 1.116025
```

It seems it gets larger values both in $E(x)$ and $E(\frac{1}{x})$. I think it's because the transform from $\log(z)$ to z is an exponential. so if $\log(z)$ has small outliers, then $\exp(z)$ will be very large or very small, that's why it has both greater $Z(X)$ and $E(\frac{1}{X})$.

Conclusion: If I'm going to sample, I will first plot the pdf and see which distribution does it look like, if it's similar to some familiar distribution which is easy to sample from, I will use IMH. If it's like a Gaussian mixture or not a familiar distribution, I will use RMH.

If proposal in IMH is similar to the torget, then the result will be really nice, but if it's not similar, it will result in a waste of simulations and the result will also be bod.

2. Consider i.i.d. data $x_1, \ldots, x_n$ such that $x_i|\nu, \theta \sim \text{Gamma}(\nu, \theta)$ where $E(x_i) = \nu/\theta$, and assign priors $\nu \sim \text{Gamma}(3, 1)$ and $\theta \sim \text{Gamma}(2, 2)$.

(a) Develop a Metropolis-within-Gibbs algorithm to sample from $p(\nu, \theta|x_1, \ldots, x_n)$ using the full conditional distributions $p(\theta|\nu, x_1, \ldots, x_n)$ and $p(\nu|\theta, x_1, \ldots, x_n)$. For the second full conditional, use a random walk proposal on $\log(\nu)$.

(b) Develop a Metropolis-Hastings algorithm that jointly proposes $\log(\nu)$ and $\log(\theta)$ using a Gaussian random walk centered on the current value of the parameters. Tune the variance-covariance matrix of the proposal using a test run that proposes the parameters independently (but evaluates acceptance jointly). $\Sigma_{12} = \Sigma_{21} = 0$

(c) Develop a Metropolis algorithm that jointly proposes $\log(\nu)$ and $\log(\theta)$ using independent proposals based on the Laplace approximation of the posterior distribution of $\log(\nu)$ and $\log(\theta)$.

(d) Run each of the algorithms for the dataset in `my-data.txt` and compute the effective sample sizes associated with each parameter under each of the samplers. Also, construct trace and autocorrelation plots. Report posterior means for each of the parameters of interest, along with 95% symmetric credible intervals. Discuss.

(0) $f_{x_i|\alpha,\beta} = \dfrac{\beta^{\alpha}}{\Gamma(\alpha)} x_i^{\alpha-1} \exp(-\beta x_i)$

shape $= \alpha$, rate $= \beta$

$\pi(\alpha) = \dfrac{1^3}{\Gamma(3)} \cdot \alpha^{3-1} \exp(-1 \cdot \alpha)$

$\pi(\beta) = \dfrac{2^2}{\Gamma(2)} \beta^{2-1} \exp(-2 \cdot \beta)$

$\pi(\alpha, \beta) \propto \alpha^2 \beta \exp(-\alpha - 2\beta)$

$f(x|\alpha, \beta) = \prod_{i=1}^{n} \dfrac{\beta^{\alpha}}{\Gamma(\alpha)} x_i^{\alpha-1} \exp(-\beta x_i)$

$= \dfrac{\beta^{n\alpha}}{[\Gamma(\alpha)]^n} \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} \exp\left(-\beta \sum_{i=1}^{n} x_i\right)$

$\pi(\alpha, \beta | x) \propto \pi(\alpha, \beta) \cdot f(x|\alpha, \beta)$

$= \dfrac{\alpha^2}{\Gamma(\alpha)^n} \cdot \beta^{n\alpha+1} \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} \exp\left(-\beta \sum_{i=1}^{n} x - \alpha - 2\beta\right)$

$$\pi(\alpha,\beta \mid x) \propto \pi(\alpha,\beta)\cdot f(x\mid\alpha,\beta)$$

$$= \frac{\alpha^2}{\Gamma(\alpha)^n}\cdot \beta^{n\alpha+1}\left(\prod_{i=1}^n x_i\right)^{\alpha-1} \exp\left(-\beta\sum_{i=1}^n x_i-\alpha-2\beta\right)$$

$$\pi(\alpha\mid\beta,x) \propto \pi(\alpha,\beta\mid x)$$

$$\propto \frac{\alpha^2}{\Gamma(\alpha)^n}\cdot \beta^{n\alpha}\left(\prod_{i=1}^n x_i\right)^{\alpha-1}\exp(-\alpha)$$

$$\pi(\beta\mid\alpha,x) \propto \pi(\alpha,\beta\mid x)$$

$$\propto \beta^{n\alpha+1}\exp\left(-\beta\sum_{i=1}^n x_i-2\beta\right)$$

Since $\alpha,\beta>0$, we should do log transformation if we want to use random walk:

$$\pi(w=\log(\alpha)\mid\beta,x) \propto \frac{e^{2w}\cdot e^{w}}{(\Gamma(e^w))^n}\beta^{ne^w}\left(\prod_{i=1}^n x_i\right)^{e^w-1}\exp(-e^w)$$

$$\pi(z=\log(\beta)\mid\alpha,x) \propto e^{z(n\alpha+1)}\exp\left(-e^z\sum_{i=1}^n x_i-2e^z\right)$$

Set initiall value $w^{(1)}, z^{(1)}$. for i in 1 : Max :

Step 1 : $z^* = \text{rnorm}(1, \text{mean}=z^{(i)}, \text{sd}=\text{sd})$

Step 2 : $P_z = \min\left\{\frac{\pi(z^*\mid\alpha,x)}{\pi(z^{(i)}\mid\alpha,x)}, 1\right\}$ $\alpha=e^{w^{(i-1)}}$

$k=\text{runif}(1)$, $\begin{cases} k\le P_z: & z^{(i+1)}\leftarrow z^* \\ k>P_z: & z^{(i+1)}\leftarrow z^{(i)}\end{cases}$

Step 3 : $w^* = \text{rnorm}(1, \text{mean}=w^{(i)}, \text{sd}=\text{sd})$

Step 4: $\quad P_w = \min \left\{ \dfrac{\pi(w^* \mid \beta, x)}{\pi(w^{(i)} \mid \beta, x)} , 1 \right\}$

ps: β here is $\exp(z^{(i+1)})$

Step 5: $\quad k = \text{runif}(1)$, If $\begin{cases} k \le P_w : & w^{(i+1)} \leftarrow w^* \\ k > P_w : & w^{(i+1)} \leftarrow w^{(i)} \end{cases}$

Then, loop.

(b) Get the joint posterior of $\log(\alpha)$ and $\log(\beta)$.

let $w = \log(\alpha)$, $z = \log(\beta)$, $\alpha = e^w$, $\beta = e^z$

$$J = \begin{vmatrix} \partial\alpha/\partial w & \partial\alpha/\partial z \\ \partial\beta/\partial w & \partial\beta/\partial z \end{vmatrix} = \begin{vmatrix} e^w & 0 \\ 0 & e^z \end{vmatrix} = e^w e^z$$

$$\pi(w, z \mid x) \propto \frac{(e^w)^2}{[\Gamma(e^w)]^n} \cdot e^{z(ne^w+1)} \left(\prod_{i=1}^n x_i\right)^{e^w-1} \cdot \exp\left(-e^z \sum_{i=1}^n x_i - e^w - 2e^z\right) \cdot e^w \cdot e^z$$

Set $(w^{(1)}, z^{(1)})$, for i in 1 : Max:

Step 1: $\quad (w^*, z^*) = \text{rMVN}(1, \text{mean} = (w^{(i)}, z^{(i)}), \text{sigma} = \Sigma)$

Step 2: $\quad P = \min \left\{ \dfrac{\pi(w^*, z^*)}{\pi(w^{(i)}, z^{(i)})} , 1 \right\}$

$k = \text{runif}(1)$, If $\begin{cases} k \le p, & (w^{(i+1)}, z^{(i+1)}) \leftarrow (w^*, z^*) \\ k > p, & (w^{(i+1)}, z^{(i+1)}) \leftarrow (w^{(i)}, z^{(i)}) \end{cases}$

loop.

(c) $$\ell(w,z|x) = 2w - n\log(\Gamma(e^w)) + z(ne^w+1) + (e^w-1)\sum_{i=1}^n \log(x_i) - e^z\sum_{i=1}^n x_i - e^w - e^{2z} + w + z$$

$$\partial \ell / \partial w = 3 - n \cdot \frac{\Gamma'(e^w)}{\Gamma(e^w)} \cdot e^w + nze^w + \sum_{i=1}^n \log(x_i) \cdot e^w - e^w$$

$$\partial \ell / \partial z = ne^w + 1 - 2e^{2z} - \sum_{i=1}^n x_i \cdot e^z + 1$$

$$\partial^2 \ell / \partial w^2 = -n\left(e^w \cdot \frac{\Gamma'(e^w)}{\Gamma(e^w)} + e^{2w} \cdot \frac{\Gamma''(e^w)\,\Gamma(e^w) - (\Gamma'(e^w))^2}{[\Gamma(e^w)]^2}\right) + \left(nz + \sum_{i=1}^n \log(x_i) - 1\right)e^w$$

$$\frac{\partial \log(f(x))}{\partial (x)} = \frac{1}{f(x)} \cdot f'(x)$$

$$\frac{\partial^2 \log(f(x))}{\partial x^2} = \frac{f''(x)f(x) - [f'(x)]^2}{[f(x)]^2}$$

$$\partial^2 \ell / \partial w \partial z = ne^w$$

$$\partial^2 \ell / \partial z \partial w = ne^w$$

$$\partial^2 \ell / \partial z^2 = -e^z \sum_{i=1}^n x_i - 4e^{2z}$$

We get mode by optim function, and calculate the negative Hessian matrix.

Then, $(w, z) \mid x$ Approximately follows:

$$MVN\left(\begin{pmatrix}\mu_0 \\ z_0\end{pmatrix}, A^{-1}\right)$$

in which, (μ_0, z_0) is the mode, A is $-$Hessian of $\log(\pi(w, z \mid x))$.

Then do the MH:

Initialize $(w^{(1)}, z^{(1)})$, for i in $1:Max$:

Step 1: $(w^*, z^*) = \text{rMVnorm}:(1, \text{mean} = \begin{pmatrix}\mu_0 \\ z_0\end{pmatrix}, \text{sigma} = A^{-1})$.

Step 2: $p = \min\left\{\dfrac{\pi(w^*, z^* \mid x)\, q(w^{(i)}, z^{(i)})}{\pi(w^{(i)}, z^{(i)} \mid x) \cdot q(w^*, z^*)}, 1\right\}$

$k = \text{runif}(1)$. If $\begin{cases} k \le p, & (w^*, z^*) \to (w^{(i+1)}, z^{(i+1)}) \\ k > p. & (w^{(i)}, z^{(i)}) \to (w^{(i+1)}, z^{(i+1)}) \end{cases}$

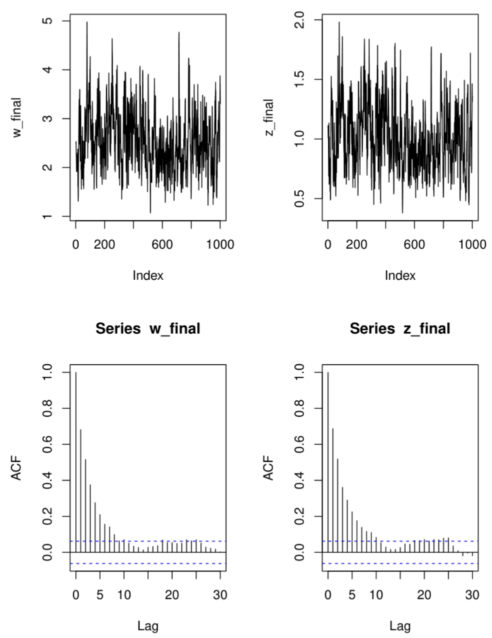
loop!

2.d:

The results are as follows:

MH-Gibbs:

Question 2(a): MH-within-Gibbs

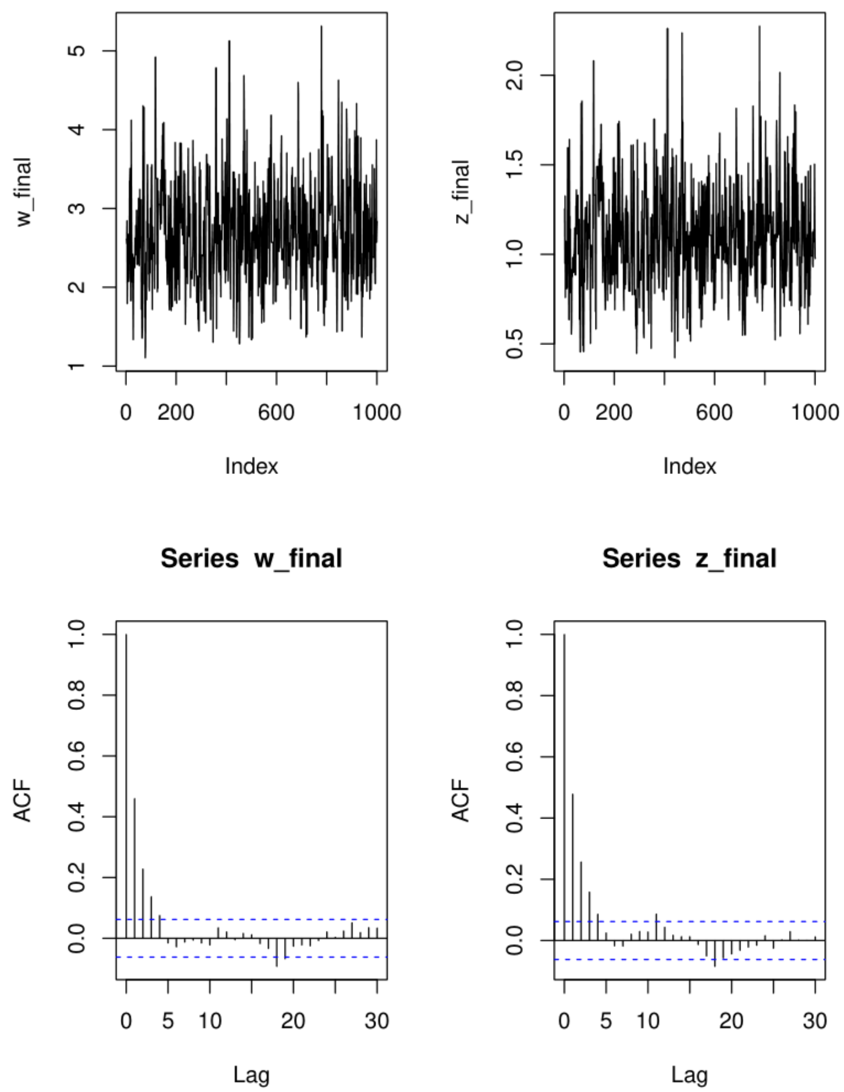


	Alpha	Beta
Mean	2.546372	1.0146160
2.5% Quantile	1.523755	0.5604799
97.5% Quantile	3.921112	1.6057523
Effective Sample Size	136.494488	136.6779558

There are strong correlation even after thinning and burn-in samples, so the effective sample size is small. but the chain converges

Jointly random walk:

Question 2(b): R-MH

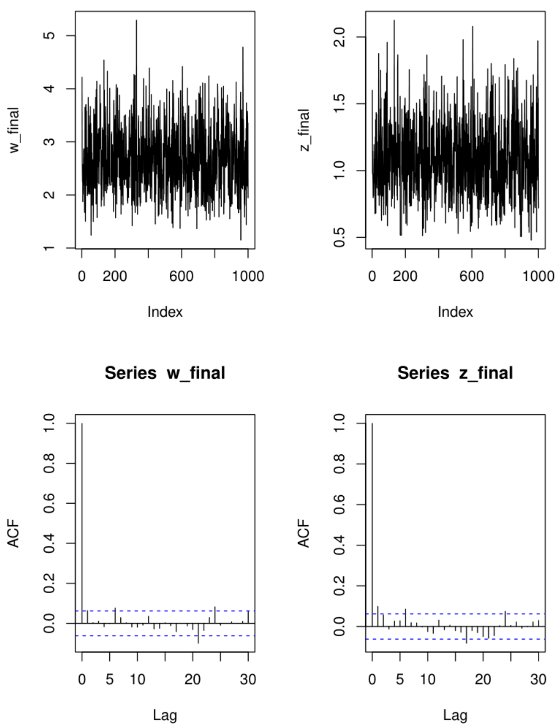


	Alpha	Beta
Mean	2.690364	1.0965319
2.5% Quantile	1.475518	0.5674203
97.5% Quantile	4.174733	1.7291608
Effective Sample Size	368.919070	297.7480121

Effecive sample size increases a lot, but it reports a wider credible interval. The posterior mean is approximately same as MH-Gibbs.

Independent MH with Laplace:

Question 2(c): I-MH_Laplace



	Alpha	Beta
Mean	2.677343	1.0878133
2.5% Quantile	1.633716	0.5795919
97.5% Quantile	4.011857	1.6864096
Effective Sample Size	966.094959	840.3655899

9

This one has the largest ess, so I think this one may be the most accurate, same, the mean is similar to the other two, but I prefer this one.

3. (Robert and Casella) Consider a random effects model,

$$y_{i,j} = \beta + u_i + \epsilon_{i,j}, \quad i = 1 : I, j = 1 : J,$$

where $u_i \sim N(0, \sigma^2)$ and $\epsilon_{i,j} \sim N(0, \tau^2)$. Assume a prior of the form

$$\pi(\beta, \sigma^2, \tau^2) \propto \frac{1}{\sigma^2 \tau^2}.$$

(a) Find the full conditional distributions:

i. $\pi(u_i \mid \boldsymbol{y}, \beta, \tau^2, \sigma^2)$; is $u_{(-i)}$ given?

ii. $\pi(\beta \mid \boldsymbol{y}, \boldsymbol{u}, \tau^2, \sigma^2)$;

iii. $\pi(\sigma^2 \mid \boldsymbol{y}, \boldsymbol{u}, \beta, \tau^2)$;

iv. $\pi(\tau^2 \mid \boldsymbol{y}, \boldsymbol{u}, \beta, \sigma^2)$.

(b) Find $\pi(\beta, \tau^2, \sigma^2 \mid \boldsymbol{y})$ up to a proportionality constant.

(c) Find $\pi(\sigma^2, \tau^2 \mid \boldsymbol{y})$ up to a proportionality constant and show that this posterior is not integrable since, for $\tau \neq 0$, it behaves like σ^{-2} in a neighborhood of 0.

(a) Likelihood:

$$f(y \mid u, \beta, \tau^2, \sigma^2) = \prod_{i=1}^{I} \prod_{j=1}^{J} (2\pi\tau^2)^{-\frac{1}{2}} \exp\left(-\frac{(y_{ij} - \beta - u_i)^2}{2\tau^2}\right)$$

$$\propto (\tau^2)^{-\frac{IJ}{2}} \exp\left(-\frac{1}{2\tau^2} \sum_{i=1}^{I} \sum_{j=1}^{J} (y_{ij} - \beta - u_i)^2\right)$$

$$\pi(u, \beta, \tau^2, \sigma^2 \mid y) \propto (\tau^2)^{-IJ/2} (\sigma^2)^{-I/2} \exp\left(-\frac{1}{2\tau^2} \sum_{i=1}^{I} \sum_{j=1}^{J} (y_{ij} - \beta - u_i)^2\right) \cdot \exp\left(-\frac{\sum_{i=1}^{I} u_i^2}{2\sigma^2}\right)(\sigma^2\tau^2)^{-1}$$

i)

$$\pi(u_i \mid \beta, \tau^2, \sigma^2, y) \propto \exp\left(-\frac{u_i^2}{2\sigma^2}\right) \exp\left(-\frac{1}{2\tau^2} \sum_{j=1}^{J} (y_{ij} - \beta - u_i)^2\right)$$

$$\propto \exp\left(-\frac{1}{2}\left[\left(\frac{1}{\sigma^2} + \frac{J}{\tau^2}\right) u_i^2 + 2 \cdot \frac{1}{\tau^2} \sum_{j=1}^{J} (y_{ij} - \beta) u_i\right]\right)$$

$$\propto \exp\left(-\frac{1}{2}\left(\frac{1}{\sigma^2} + \frac{J}{\tau^2}\right)^{-1} \left(u_i - \frac{\frac{\sum_{j=1}^{J}(y_{ij} - \beta)}{\tau^2}}{1/\sigma^2 + J/\tau^2}\right)^2\right)$$

Kernel of: $N\left(\frac{\sum_{j=1}^{J}(y_{ij} - \beta)/\tau^2}{1/\sigma^2 + J/\tau^2}, \left(1/\sigma^2 + J/\tau^2\right)^{-1}\right)$

$$\pi(u,\beta,\tau^2,\sigma^2 \mid y) \propto (\tau^2)^{-IJ/2} (\sigma^2)^{-I/2} \exp\left(-\frac{1}{2\tau^2}\sum_{i=1}^{I}\sum_{j=1}^{J}(y_{ij}-\beta-u_i)^2\right)$$

$$\cdot \exp\left(-\frac{\sum_{i=1}^{I}u_i^2}{2\sigma^2}\right)\cdot\frac{1}{\sigma^2\tau^2}$$

ii) $\pi(\beta \mid y, u, \tau^2, \sigma^2) \propto \exp\left(-\frac{1}{2\tau^2}\sum_{i=1}^{I}\sum_{j=1}^{J}\left(\beta-(y_{ij}-u_i)\right)^2\right)$

$$\propto \exp\left(-\frac{1}{2\tau^2}\left(IJ\beta^2 - 2\sum_{i=1}^{I}\sum_{j=1}^{J}(y_{ij}-u_i)\beta\right)\right)$$

$$\propto \exp\left(-\frac{IJ}{2\tau^2}\left(\beta-\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(y_{ij}-u_i)}{IJ}\right)^2\right)$$

It's $N\left(\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(y_{ij}-u_i)}{IJ}, \frac{\tau^2}{IJ}\right)$

iii) $\pi(\sigma^2 \mid \beta, u, \tau^2, y) \propto (\sigma^2)^{-I/2-1}\exp\left(-\frac{1}{\sigma^2}\frac{\sum_{i=1}^{I}u_i^2}{2}\right)$

It's $IG\left(\frac{I}{2}, \frac{\sum_{i=1}^{I}u_i^2}{2}\right)$

iv) $\pi(\tau^2 \mid \beta, u, \sigma^2, y) \propto (\tau^2)^{-\frac{IJ}{2}-1}\exp\left(-\frac{1}{\tau^2}\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(y_{ij}-\beta-u_i)^2}{2}\right)$

It's $IG\left(IJ/2, \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}(y_{ij}-u_i-\beta)^2}{2}\right)$

(b) $\pi(\beta,\tau,\sigma^2 \mid y) = \int_u \pi(\beta,\tau,\sigma^2,u \mid y)\,du$

$$\propto \int_u (\tau^2)^{-\frac{IJ}{2}-1}(\sigma^2)^{-I-1}\exp\left(-\frac{1}{2\tau^2}\sum_{i=1}^{I}\sum_{j=1}^{J}(y_{ij}-\beta-u_i)^2-\frac{\sum_{i=1}^{n}u_i^2}{2\sigma^2}\right)du$$

$$\propto (\tau^2)^{-\frac{IJ}{2}-1}(\sigma^2)^{-\frac{I}{2}-1}\int_u \exp\left(\sum_{i=1}^{I}\underbrace{\left(-\frac{\sum_{j=1}^{J}(y_{ij}-\beta-u_i)^2}{2\tau^2}-\frac{u_i^2}{2\sigma^2}\right)}_{\text{Denote to }\star}\right)du$$

$$\star = -\frac{1}{2}\left(\left(\frac{J}{\tau^2}+\frac{1}{\sigma^2}\right)u_i^2 - 2\frac{\sum_{j=1}^{J}(y_{ij}-\beta)\cdot u_i}{\tau^2} + \sum_{j=1}^{J}(y_{ij}-\beta)^2/\tau^2\right)$$

$$= -\frac{1}{2}\left(J/\tau^2+1/\sigma^2\right)\left(u_i - \frac{\frac{1}{\tau^2}\sum_{j=1}^{J}(y_{ij}-\beta)}{J/\tau^2+1/\sigma^2}\right)^2 + \frac{\left(\sum_{j=1}^{J}(y_{ij}-\beta)\right)^2/\tau^2}{J/\tau^2+1/\sigma^2} -$$

$$= -\frac{1}{2}\left(\frac{J}{\tau^2}+\frac{1}{\sigma^2}\right)\left[\left(u_i - \frac{\frac{1}{\tau^2}\sum_{j=1}^{J}(y_{ij}-\beta)}{J/\tau^2+1/\sigma^2}\right)^2 + \frac{\sum_{j=1}^{J}(y_{ij}-\beta)^2\cdot\frac{1}{\tau^2}}{J/\tau^2+1/\sigma^2} - \left[\frac{\frac{1}{\tau^2}\sum_{j=1}^{J}(y_{ij}-\beta)}{J/\tau^2+1/\sigma^2}\right]^2\right]$$

Let $\left(J/\tau^2+1/\sigma^2\right)^{-1} = \tau_1^2$

$$\frac{1/\tau^2\cdot\sum_{j=1}^{J}(y_{ij}-\beta)}{J/\tau^2+1/\sigma^2} = \mu_{u_i}$$

$$\Rightarrow \pi(\beta,\sigma,\tau\,|\,y) \propto (\tau^2)^{-\frac{IJ}{2}-1}(\sigma^2)^{-\frac{I}{2}-1}\cdot\exp\left(\frac{-\sum_{i=1}^{I}\sum_{j=1}^{J}(y_{ij}-\beta)^2}{2\tau^2}\right)$$

$$\exp\left(\sum_{i=1}^{I}\frac{\mu_{u_i}^2}{2\tau_1^2}\right)\cdot$$

$$\prod_{i=1}^{I}\int_{u_i}\underbrace{\exp\left(-\frac{1}{2\tau_1^2}(u_i-\mu_{u_i})^2\right)}_{\text{Normal Kernel}}du_i$$

$$\propto (\tau^2)^{-\frac{IJ}{2}-1}(\tau_1^2)^{\frac{J}{2}}(\sigma^2)^{-\frac{I}{2}-1}\exp\left(\sum_{i=1}^{I}\sum_{j=1}^{J}(y_{ij}-\beta)^2/2\tau^2\right)$$

$$\exp\left(\sum_{i=1}^{I}\frac{\mu_{u_i}^2}{2\tau_1^2}\right)$$

d) $\pi(\sigma^2,\tau^2\,|\,y) = \int_\beta \pi(\beta,\sigma,\tau\,|\,y)\,d\beta$

$$\propto (\tau^2)^{-\frac{IJ}{2}-1}(\tau_1^2)^{\frac{I}{2}}(\sigma^2)^{-\frac{I}{2}-1}\int_\beta \exp\left(\sum_{i=1}^{I}\sum_{j=1}^{J}(y_{ij}-\beta)^2/2\tau^2\right)$$

$$\exp\left(-\frac{1/\tau^2}{2(J+\tau^2/\sigma^2)}\sum_{i=1}^{I}\left[\sum_{j=1}^{J}(y_{ij}-\beta)\right]^2\right)d\beta$$

$$\phi = \exp\left(-\sum_{i=1}^{I}\sum_{j=1}^{J}(y_{ij}-\beta)^2/2\tau^2\right)\cdot$$

$$\exp\left(\frac{1/\tau^2}{2(J+\tau^2/\sigma^2)}\sum_{i=1}^{I}\left[\sum_{j=1}^{J}(y_{ij}-\beta)\right]^2\right)d\beta$$

$$=\exp\left(-\frac{1}{2}\left(-\frac{IJ\beta^2-2\sum_i\sum_j y_{ij}\beta+\sum_i\sum_j y_{ij}^2}{\tau^2}+\frac{\sum_{i=1}^{I}\left(\sum_{j=1}^{J}y_{ij}-J\beta\right)^2}{2(J+\tau^2/\sigma^2)\tau^2}\right)\right)d\beta$$

$$\sum_{i=1}^{I}\left(J^2\beta^2+\left[\sum_{j=1}^{J}y_{ij}\right]^2-2J\beta\sum_{j=1}^{J}y_{ij}\right)$$

$$=\frac{IJ^2\beta^2-2J\beta\sum_{i=1}^{I}\sum_{j=1}^{J}y_{ij}+\sum_{i=1}^{I}\left[\sum_{j=1}^{J}y_{ij}\right]^2}{2(J/\tau^2+1/\sigma^2)\tau^4}$$

$$=\exp\left\{-\frac{1}{2}\left[\beta^2\left(\frac{IJ^2}{(J/\tau^2+1/\sigma^2)\tau^4}-\frac{IJ}{\tau^2}\right)-2\beta\cdot\right.\right.$$

$$= (\tilde\tau^2)^{-1}$$

$$\tilde\mu_\beta = \tilde\mu\cdot\tilde\tau^2$$

$$\left(\frac{2J\sum_{i=1}^{I}\sum_{j=1}^{J}y_{ij}}{(J/\tau^2+1/\sigma^2)\tau^4}-\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}y_{ij}}{\tau^2}\right) = \tilde\mu \quad +$$

$$\tilde s \leftarrow \left.\left.\left(\frac{\sum_{i=1}^{I}\left[\sum_{j=1}^{J}(y_{ij})\right]^2}{(J/\tau^2+1/\sigma^2)\tau^4}-\frac{\sum_i\sum_j y_{ij}^2}{\tau^2}\right)\right]\right\}$$

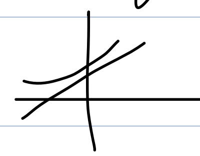
$$=\exp\left\{-\frac{1}{2}(\tilde\tau)^{-1}\left[(\beta-\tilde\mu_\beta)^2-\tilde\mu_\beta^2+\tilde s\,\tilde\tau\right]\right\}$$

$$\Rightarrow \pi(\sigma^2, \tau^2 \mid y) \propto (\sigma^2)^{-\frac{J}{2}-1} (\tau^2)^{-\frac{2J}{2}-1} (\tau_1^2)^{\frac{J}{2}} \cdot (\tilde{\tau}^2)^{\frac{1}{2}} \exp\left(-\frac{\tilde{S}}{2}\right)$$

$$\exp\left(\frac{\hat{\mu}_\beta \tilde{\tau}^2}{2} \right)$$

$\tau_1 = o(\sigma^2)$ $\quad$ $\hat{\mu}_\beta = o(\sigma^2)$ $\quad$ $\tilde{\tau}^2 = o(\sigma^2)$

$\tilde{S} = o(\sigma^2)$



Since around 0, $\exp(x) = x+1$

Ignore τ^2, $\pi(\sigma^2, \tau^2 \mid y) \propto (\sigma^2)^{-\frac{J}{2}-1} (\sigma^2)^{J/2} \cdot (\sigma^2)^{\frac{1}{2}} \cdot h(\tau^2, \sigma^2)$

inwhich, $h(\tau^2)$ is finite . and around 0.

$\Rightarrow \pi(\sigma^2, \tau^2 \mid y) \propto (\sigma^2)^{-\frac{1}{2}}$

$\Rightarrow$ Around 0, $\pi(\sigma^2, \tau^2 \mid y)$ is not defined well

around 0. not integrable.

4. (Carlin, Gelfand and Smith, 1992) Let $y_1, \ldots, y_n$ be a sample from a Poisson distribution for which there is a suspicion of a change point m along the observation process where the means change, $m = 1, \ldots, n$. Given m, $y_i \sim \text{Poi}(\theta)$, for $i = 1, \ldots, m$ and $y_i \sim \text{Poi}(\phi)$, for $i = m+1, \ldots, n$. The model is completed with independent prior distributions $\theta \sim \text{Gamma}(\alpha, \beta)$, $\phi \sim \text{Gamma}(\gamma, \delta)$ and m uniformly distributed over $\{1, \ldots, n\}$ where α, β, γ and δ are known constants. The data in file `mining-data.r` consists of counts of coal mining disasters in Great Britain by year from 1851 to 1962.

(a) Describe a Gibbs sampling algorithm to obtain samples from the joint posterior distribution.

(b) Specify values of the fixed hyperparameters, α, β, γ and δ and justify your choices.

(c) Implement your sampling algorithm in part (a) and summarize the posterior distribution using point estimates and interval estimates of the parameters. Provide interpretations of the inference in layman's word.

(d) Perform prior sensitivity examination. Refit the model with different sets of the fixed hyperparameter values and examine if the prior distributions have an undesired influence on the posterior inference.

(a). $\pi(\theta) \propto \theta^{\alpha-1} \exp(-\beta\theta)$

$\pi(\phi) \propto \phi^{\gamma-1} \exp(-\delta\phi)$

$\pi(m) = 1/n$

$$\pi(y \mid m, \theta, \phi) = \prod_{i=1}^{m} \frac{\theta^{y_i}}{y_i!} e^{-\theta} \cdot \prod_{i=m+1}^{n} \frac{\phi^{y_i}}{y_i!} e^{-\phi}$$

$$\propto \theta^{\sum_{i=1}^{m} y_i} \exp(-m\theta)\, \phi^{\sum_{i=m+1}^{n} y_i} \exp(-(n-m)\phi)$$

$$\pi(m, \phi, \theta \mid y) \propto \pi(y \mid m, \theta, \phi) \cdot \pi(m) \cdot \pi(\theta) \cdot \pi(\phi)$$

$$\propto \theta^{\sum_{i=1}^{m} y_i + \alpha - 1} \exp(-(m+\beta)\theta) \cdot \phi^{\sum_{j=m+1}^{n} y_j + \gamma - 1} \exp(-(n-m+\delta)\phi)$$

$$\pi(m \mid \phi, \theta, y) \propto \theta^{\sum_{i=1}^{m} y_i} \exp(-m\theta)\, \phi^{\sum_{j=m+1}^{n} y_j} \exp(m\phi)$$

Given the others, m can be calculated in each case.

$P(m=k) = ?$ can be calculated.

$$\theta^{\sum_{i=1}^{m} y_i + \alpha - 1} \exp(-(m+\beta)\theta) \cdot \phi^{\sum_{j=m+1}^{n} y_j + \gamma - 1} \exp(-(n-m+\delta)\phi)$$

$$\pi(\theta \mid y, m, \phi) \propto \theta^{\sum_{i=1}^{m} y_i + \alpha - 1} \exp(-(m+\beta)\theta)$$

$$\pi(\phi \mid y, m, \theta) \propto \phi^{\sum_{j=m+1}^{n} y_j + \gamma - 1} \exp(-(n-m+\delta)\phi)$$

$$\pi(\theta \mid y, m, \phi) \sim \text{Gamma}\left(\sum_{i=1}^{m} y_i + \alpha,\ m+\beta\right)$$

$$\pi(\phi \mid y, m, \theta) \sim \text{Gamma}\left(\sum_{j=m+1}^{n} y_j + \gamma,\ n-m+\delta\right)$$

In Gibbs sampling,

Step 1: Initialize:

$m^{(1)}$, $\theta^{(1)}$, $\phi^{(1)}$

Step 2: r Gamma (1) generate $\theta^{(2)}$, based on $m^{(1)}$, $\phi^{(1)}$

Step 3: r Gamma (1) generate $\phi^{(2)}$, based on $\theta^{(2)}$, $m^{(1)}$

Step 4: r discrete (1), generate $m^{(2)}$ based on $\theta^{(2)}$, $\phi^{(2)}$

(b) I got $\alpha, \beta, \gamma, \delta$ by specifying a weak prior: rate $= \beta$, rate $= \delta$,

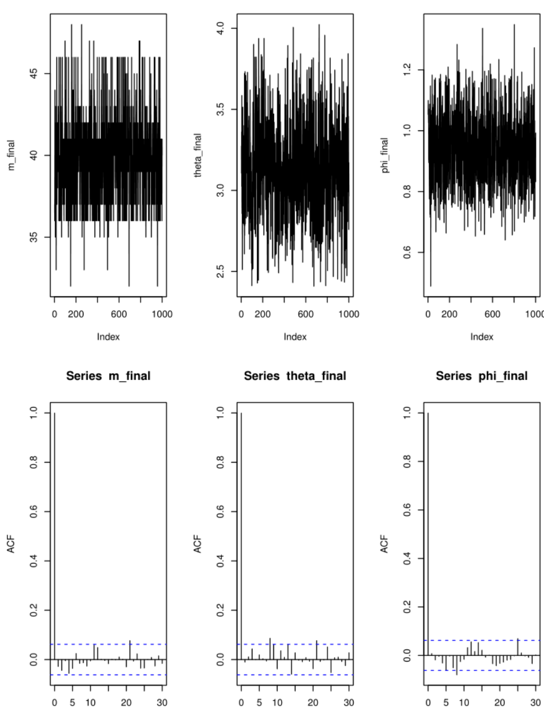
$E(\theta) = \alpha/\beta$, $E(\phi) = \gamma/\delta$

$Var(\theta) = \alpha/\beta^2$, $Var(\phi)$ γ/δ^2

I first set $\alpha = \gamma = 1$, $\beta = \delta = 0.1$.

(c). The results are as follows

Question 4:



	M	Theta	Phi
Mean	39.87413	3.140417	0.9409816
2.5% Quantile	36.00000	2.550139	0.7269625
97.5% Quantile	46.00000	3.752849	1.1765244
Effective Sample Size	1001.00000	1001.000000	1001.0000000

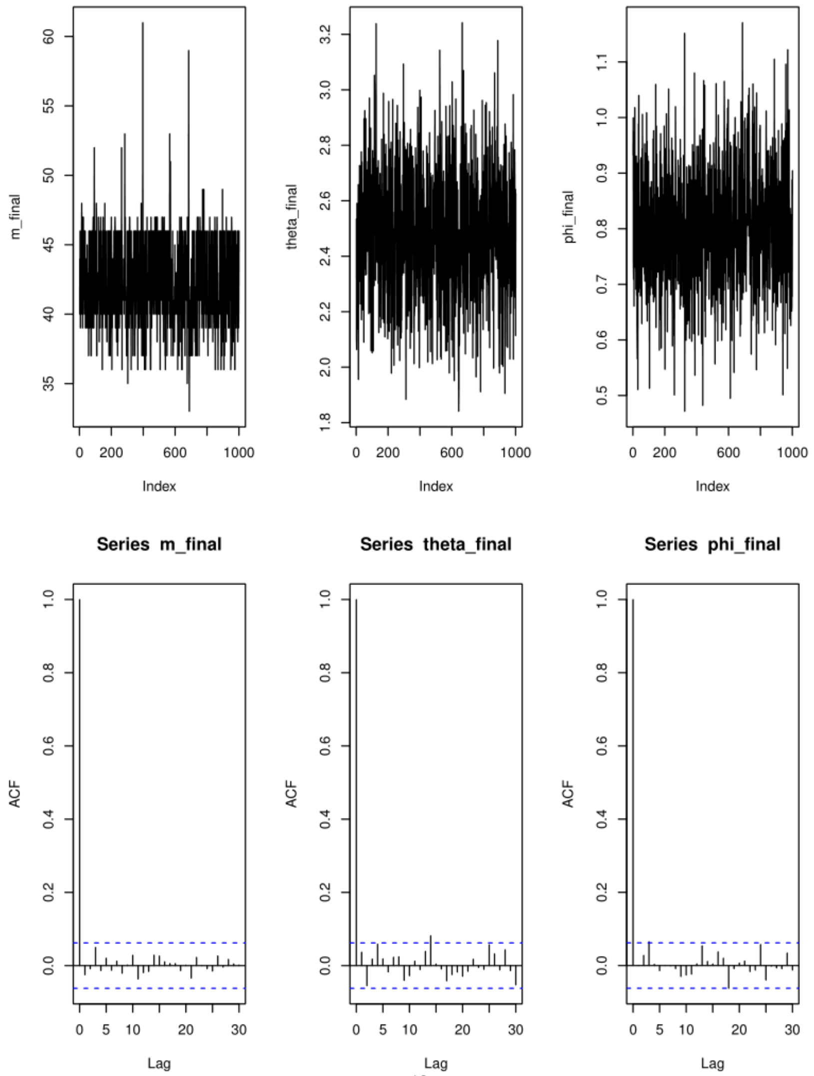
Interpretation:

There is a 95% probability that the true parameter falls in this interval.

(d) If I use stronger prior:

$\alpha = \gamma = 1$, $\beta = \delta = 10$:

Question 4: (alpha = gamma = 1, beta = delta = 10)

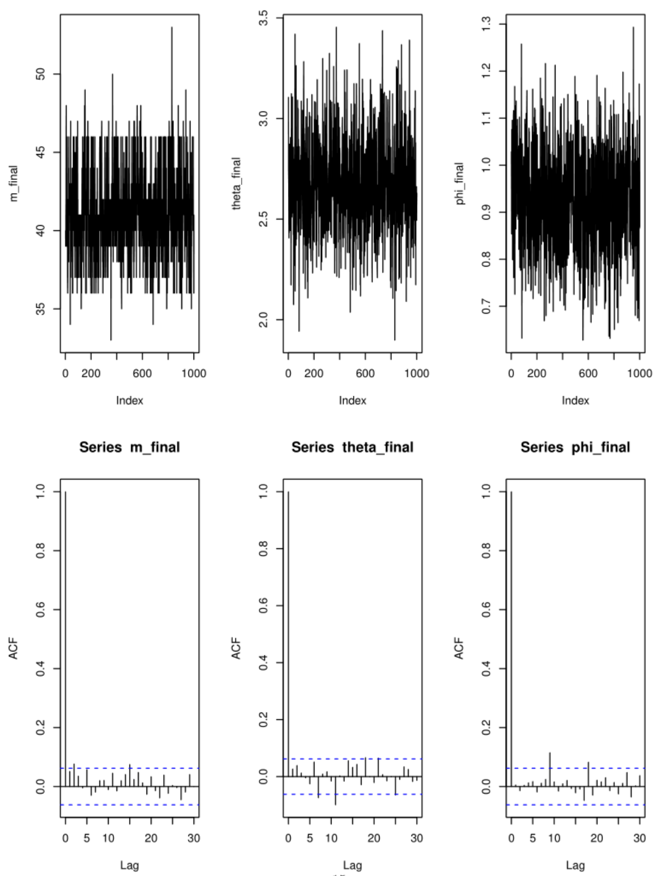


	M	Theta	Phi
Mean	42.19281	2.480381	0.8010412
2.5% Quantile	37.00000	2.029126	0.5991549
97.5% Quantile	47.00000	2.937817	1.0121270
Effective Sample Size	1001.00000	1001.000000	1001.0000000

If prior specified some wrong information:

$\alpha = \gamma = 10$, $\beta = \delta = 10$

Question 4: (alpha = gamma = 10, beta = delta = 10)

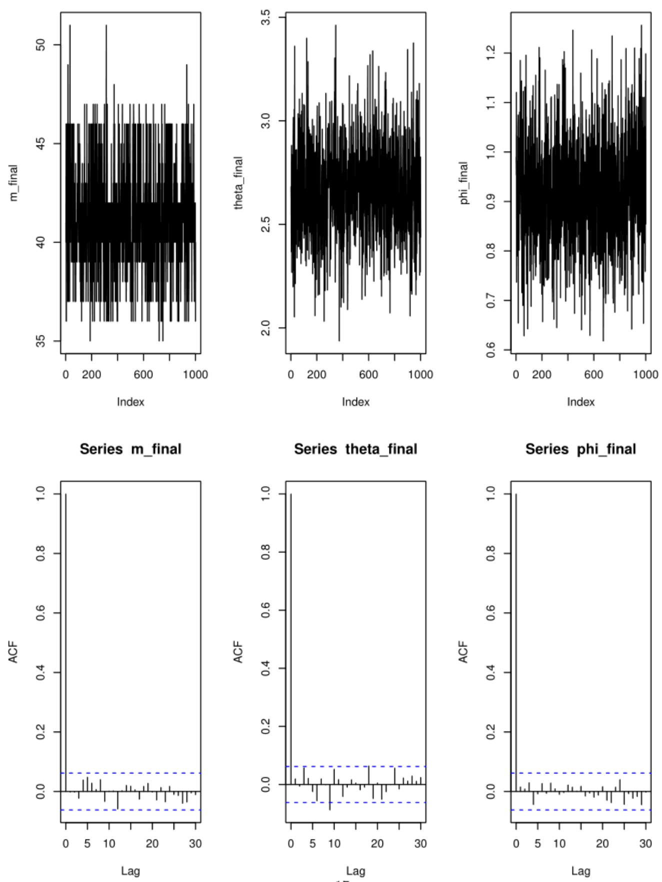


15

	M	Theta	Phi
Mean	41.05395	2.671482	0.9235956
2.5% Quantile	36.00000	2.210222	0.7106051
97.5% Quantile	46.00000	3.147084	1.1312292
Effective Sample Size	1001.00000	1001.000000	1001.0000000

or $\alpha = \gamma = 100$, $\beta = \delta = 10$

Question 4: (alpha = gamma = 100, beta = delta = 10)



	M	Theta	Phi
Mean	41.12687	2.664721	0.9170876
2.5% Quantile	36.00000	2.213647	0.7073331
97.5% Quantile	46.00000	3.106189	1.1443681
Effective Sample Size	1001.00000	1001.000000	1001.0000000

Conclusion: This prior seems not to have a very strong influence on the posterior. I think it's about the sample size, if sample size $n \uparrow$, the influence about the prior is decreasing all the time.

5. Souza (1999) considers a number of hierarchical models to describe the nutritional pattern of pregnant women. One of the models adopted was a hierarchical regression model where

$$\begin{aligned} y_{i,j} &\sim \mathrm{N}(\alpha_i + \beta_i t_{i,j}, \sigma^2), \\ (\alpha_i, \beta_i)' \mid \alpha, \beta &\sim \mathrm{MVN}_2((\alpha, \beta)', diag(\tau_\alpha^2, \tau_\beta^2)), \\ (\alpha, \beta)' &\sim \mathrm{MVN}_2((0,0)', diag(P_\alpha^2, P_\beta^2). \end{aligned}$$

Here $y_{i,j}$ and $t_{i,j}$ are the jth weight measurement and visit time of the ith woman with $j = 1 : n_i$ and $i = 1 : I$ for $I = 68$ pregnant women. Here $n = \sum_{i=1}^{I} n_i = 427$. For unknown scale parameters, we assume a priori independence and place a Gamma prior,

$$\sigma^2 \sim \mathrm{IG}(a_\sigma, b_\sigma), \quad \tau_\alpha^2 \sim \mathrm{IG}(a_\alpha, b_\alpha), \text{ and } \tau_\beta^2 \sim \mathrm{IG}(a_\beta, b_\beta).$$

Hyperparameters, $a_\sigma, b_\sigma, a_\alpha, b_\alpha, a_\beta, b_\beta$ P_α^2 P_β^2 are fixed.

(a) Find the joint posterior distribution of all random parameters (up to proportionality)

(b) Find the full conditional distributions of $\alpha, \beta, \tau_\alpha, \tau_\beta$, $\sigma^{-2}, \alpha_i, \beta_i$, and (α_i, β_i).

(c) Describe a Gibbs sampling algorithm to obtain samples from the joint posterior distribution.

(d) Specify values of the fixed hyperparameters, and justify your choices.

(e) Implement your sampling algorithm in part (c) and summarize the posterior distribution using point estimates and interval estimates of the parameters. Provide interpretations of the inference in layman's word.

(a) Since the MVN distribution has a diagonal form, they are independent and separable.

$$f(y \mid \alpha_i, \beta_i, \sigma^2) = \prod_{i=1}^{I} \prod_{j=1}^{n_i} (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{\sum_{i=1}^{I}\sum_{j=1}^{n_i}(y_i - \alpha_i - \beta_i t_{ij})^2}{2\sigma^2}\right)$$

$$\pi(\alpha_i, \beta_i) = (2\pi \tau_\alpha^2)^{-\frac{1}{2}} (2\pi \tau_\beta^2)^{-\frac{1}{2}} \exp\left(-\frac{(\alpha_i - \alpha)^2}{2\tau_\alpha^2} - \frac{(\beta_i - \beta)^2}{2\tau_\beta^2}\right)$$

$$\pi(\alpha, \beta) = (2\pi P_\alpha^2)^{-\frac{1}{2}} (2\pi P_\beta^2)^{-\frac{1}{2}} \exp\left(-\frac{\alpha^2}{2P_\alpha^2} - \frac{\beta^2}{2P_\beta^2}\right)$$

$$\pi(\sigma^2) \propto (\sigma^2)^{-a_\sigma - 1} \exp\left(-\frac{1}{\sigma^2} b_\sigma\right)$$

$$\pi(\tau_\alpha^2) \propto (\tau_\alpha^2)^{-a_\alpha - 1} \exp\left(-\frac{1}{\tau_\alpha^2} b_\alpha\right)$$

$$\pi(\tau^2_\beta) \propto (\tau^2_\beta)^{-a_\beta - 1} \exp\left(-\frac{1}{\tau^2_\beta} b_\beta\right)$$

$$\pi(\alpha_i, \beta_i, \alpha, \beta, \tau^2_\alpha, \tau^2_\beta, \sigma^2 \mid y)$$

$$\propto \prod_{i=1}^{7}\prod_{j=1}^{n_i} (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{\sum_{i=1}^{7}\sum_{j=1}^{n_i}(y_{ij}-\alpha_i-\beta_i t_{ij})^2}{2\sigma^2}\right)$$

$$\prod_{i=1}^{7} (2\pi\tau^2_\alpha)^{-\frac{1}{2}} (2\pi\tau^2_\beta)^{-\frac{1}{2}} \exp\left(-\frac{(\alpha_i-\alpha)^2}{2\tau^2_\alpha} - \frac{(\beta_i-\beta)^2}{2\tau^2_\beta}\right)$$

$$(2\pi p^2_\alpha)^{-\frac{1}{2}} (2\pi p^2_\beta)^{-\frac{1}{2}} \exp\left(-\frac{\alpha^2}{2p^2_\alpha} - \frac{\beta^2}{2p^2_\beta}\right)$$

$$(\sigma^2)^{-a_\sigma-1} \exp\left(-\frac{1}{\sigma^2} b_\sigma\right) (\tau^2_\alpha)^{-a_\alpha-1} \exp\left(-\frac{1}{\tau^2_\alpha} b_\alpha\right)$$

$$(\tau^2_\beta)^{-a_\beta-1} \exp\left(-\frac{1}{\tau^2_\beta} b_\beta\right)$$

$$\propto (\sigma^2)^{-\frac{N}{2}-a_\sigma-1} \exp\left(-\frac{\sum_{i=1}^{2}\sum_{j=1}^{n_i}(y_{ij}-\alpha_i-\beta_i t_{ij})^2}{2\sigma^2} - \frac{b_\sigma}{\sigma^2}\right) \cdot$$

$$(\tau^2_\alpha)^{-\frac{7}{2}-a_\alpha-1} \exp\left(-\frac{\sum_{i=1}^{7}(\alpha_i-\alpha)^2}{2\tau^2_\alpha} - \frac{b_\alpha}{\tau^2_\alpha}\right) \cdot$$

$$(\tau^2_\beta)^{-\frac{7}{2}-a_\beta-1} \exp\left(-\frac{\sum_{i=1}^{2}(\beta_i-\beta)^2}{2\tau^2_\beta} - \frac{b_\beta}{\tau^2_\beta}\right) \cdot$$

$$(p^2_\alpha)^{-\frac{1}{2}} (p_\beta)^{-\frac{1}{2}} \exp\left(-\frac{\alpha^2}{2p^2_\alpha} - \frac{\beta^2}{2p^2_\beta}\right)$$

(b) $\pi(\sigma^2 \mid \text{others})$ is a $IG: \left(\frac{N}{2}+a_\sigma, \frac{\sum_{i=1}^{2}\sum_{j=1}^{n_i}(y_{ij}-\alpha_i-\beta_i t_{ij})^2}{2} + b_\sigma\right)$

$\pi(\tau^2_\alpha \mid \text{others})$ is : $IG: \left(\frac{7}{2}+a_\alpha, \frac{\sum_{i=1}^{2}(\alpha_i-\alpha)^2}{2} + b_\alpha\right)$

$\pi(\tau^2_\beta \mid \text{others})$ is: $IG: \left(\frac{7}{2}+a_\beta, \frac{\sum_{i=1}^{2}(\beta_i-\beta)^2}{2} + b_\beta\right)$

$$\pi(\alpha_i \mid \text{others}) \propto \exp\left(-\frac{\sum_{j=1}^{n_i}(y_{ij}-\alpha_i-\beta_i t_{ij})^2}{2\sigma^2} - \frac{(\alpha_i-\alpha)^2}{2\tau^2_\alpha}\right)$$

$$\sim N\left(\frac{\sum_{j=1}^{n_i}(-\beta_i t_{ij}+y_{ij})/\sigma^2 + \alpha/\tau^2_\alpha}{n_i/\sigma^2 + 1/\tau^2_\alpha}, \left(n_i/\sigma^2 + 1/\tau^2_\alpha\right)^{-1}\right)$$

$$\pi(\beta_i \mid \text{others}) \propto \exp\left(- \frac{\sum_{j=1}^{n_i} (y_{ij} - \alpha_i - \beta_i t_{ij})^2}{2\sigma^2} - \frac{(\beta_i - \beta)^2}{2\tau_\beta^2} \right)$$

$$\sum_{j=1}^{n_i} \left[\beta_i t_{ij} + (\alpha_i - y_{ij}) \right]^2$$

$$= \sum_{j=1}^{n_i} t_{ij}^2 \beta_i^2 - 2 \sum_{j=1}^{n_i} t_{ij} (-\alpha_i + y_{ij}) \beta_i + C$$

$$\propto \exp\left\{ -\frac{1}{2} \left[\beta_i^2 \left(\frac{\sum_{j=1}^{n_i} t_{ij}^2}{\sigma^2} + \frac{1}{\tau_\beta^2} \right) - 2\beta_i \left(\frac{\sum_{j=1}^{n_i} (\alpha_i - y_{ij}) t_{ij}}{\sigma^2} + \frac{\beta}{\tau_\beta^2} \right) \right] \right\}$$

$$\sim N\left(\frac{\frac{\sum_{j=1}^{n_i} (-\alpha_i + y_{ij}) t_{ij}}{\sigma^2} + \frac{\beta}{\tau_\beta^2}}{\frac{\sum_{j=1}^{n_i} t_{ij}^2}{\sigma^2} + \frac{1}{\tau_\beta^2}}, \left(\frac{\sum_{j=1}^{n_i} t_{ij}^2}{\sigma^2} + \frac{1}{\tau_\beta^2} \right)^{-1} \right)$$

$$\pi(\alpha \mid \text{others}) \propto \exp\left(- \frac{\sum_{i=1}^{I} (\alpha_i - \alpha)^2}{2\tau_\alpha^2} - \frac{\alpha^2}{2p_\alpha^2} \right)$$

$$\propto \exp\left(-\frac{1}{2} \left[\alpha^2 \left(\frac{I}{\tau_\alpha^2} + \frac{1}{p_\alpha^2} \right) - 2\alpha \frac{\sum_{i=1}^{I} (\alpha_i)}{\tau_\alpha^2} + C \right] \right)$$

$$\sim N\left(\frac{\frac{\sum_{i=1}^{I} \alpha_i}{\tau_\alpha^2}}{I/\tau_\alpha^2 + 1/p_\alpha^2}, \left(I/\tau_\alpha^2 + 1/p_\alpha^2 \right)^{-1} \right)$$

$$\pi(\alpha_i, \beta_i \mid \text{others}) \propto \exp\left(\frac{\sum_{i=1}^{2} \sum_{j=1}^{n_i} (y_{ij} - \alpha_i - \beta_i t_{ij})^2}{2\sigma^2} \right)$$

$$\exp\left(- \frac{\sum_{i=1}^{2} (\alpha_i - \alpha)^2}{2\tau^2_\alpha} \right) \cdot \exp\left(- \frac{\sum_{i=1}^{2} (\beta_i - \beta)^2}{2\tau^2_\beta} \right)$$

$$\propto \exp\left(- \frac{\sum_{j=1}^{n_i} (y_{ij} - \alpha_i - \beta_i t_{ij})^2}{2\sigma^2} \right)$$

$$\exp\left(- \frac{(\alpha_i - \alpha)^2}{2\tau^2_\alpha} - \frac{(\beta_i - \beta)^2}{2\tau^2_\beta} \right)$$

$$\propto \exp\left(- \frac{\sum_{j=1}^{n_i} (\beta_i t_{ij} - \alpha_i)^2}{2\sigma^2} - \frac{-2 \sum_{j=1}^{n_i} y_{ij} (\beta_i t_{ij} - \alpha_i)}{2\sigma^2} \right)$$

$2\alpha_i \sum_{j=1}^{n_i} y_{ij}$

$$\propto \exp\left(- \frac{n_i \alpha_i^2 - 2\alpha_i \beta_i \sum_{j=1}^{n_i} t_{ij} + \beta_i^2 \sum_{j=1}^{n_i} t_{ij}^2 - 2\beta_i \sum_{j=1}^{n_i} y_{ij} t_{ij} +}{2\sigma^2} \right.$$

$$\pi(\alpha_i, \beta_i \mid \text{others}) \propto \exp\left\{ -\frac{1}{2} \cdot \left[\right. \right.$$

$$\left(\frac{n_i}{\sigma^2} + \frac{1}{\tau^2_\alpha} \right) \alpha_i^2 - \frac{2\alpha_i \beta_i \sum_{j=1}^{n_i} t_{ij}}{\sigma^2} +$$

$$\left(\frac{\sum_{j=1}^{n_i} t_{ij}^2}{\sigma^2} + \frac{1}{\tau^2_\beta} \right) \beta_i^2 +$$

$$\left. \left. \left(\frac{2\sum_{j=1}^{n_i} y_{ij}}{\sigma^2} - \frac{2\alpha}{\tau^2_\alpha} \right) \alpha_i + \left(-\frac{2 \sum_{j=1}^{n_i} y_{ij} t_{ij}}{\sigma^2} - \frac{2\beta}{\tau^2_\beta} \right) \beta_i \right] \right\}$$

It's a quardratic form, so (α_i, β_i) jointly follows a MVN_2 distribution, but it's not independent since we have cross term $\alpha_i\beta_i$.

$$\pi(\beta|\text{others}) \propto \exp\left(-\frac{1}{2}\left(\frac{\sum_{i=1}^{I}(\beta_i-\beta)^2}{\tau_\beta^2} + \frac{\beta^2}{\rho_\beta^2}\right)\right)$$

$$\propto \left(-\frac{1}{2}\left(\beta^2\left(\frac{I}{\tau_\beta^2}+\frac{1}{\rho_\beta^2}\right)-2\beta\left(\frac{\sum_{i=1}^{I}\beta_i}{\tau_\beta^2}\right)+c\right)\right)$$

$$\sim N\left(\frac{\sum_{i=1}^{I}\beta_i / \tau_\beta^2}{I/\tau_\beta^2 + 1/\rho_\beta^2}, \left(I/\tau_\beta^2 + 1/\rho_\beta^2\right)^{-1}\right)$$

c).

Step 1: Initialize a starting value:

$(\alpha^{(1)}, \beta^{(1)}, \tau_\alpha^{(1)}, \tau_\beta^{(1)}, \alpha_1^{(1)}, \alpha_2^{(1)} \cdots \alpha_I^{(1)}, \beta_1^{(1)}, \cdots, \beta_I^{(1)}, \sigma^{2(1)})$

Step 2: for i in 1:10000

$\alpha^{(i+1)}$ = rnorm, $\pi(\alpha|\text{others})$

$\beta^{(i+1)}$ = rnorm, $\pi(\beta|\text{others})$

⋮

$\sigma^{2(i+1)}$ = rnorm, $\pi(\beta|\text{others})$

Step 3: Thinning, Burnin.

d). $a_\sigma = a_\alpha = a_\beta = 1$ $\qquad p_\alpha^2 = 100$

$b_\sigma = b_\alpha = b_\beta = 0.1$ $\qquad p_\beta^2 = 1$

I set a weak prior since if wrong prior is too strong, The result can be very different

(e)

Patient	Alpha_i(mean)	Lower 95% CI of Alpha_i	Upper 95% CI of Alpha_i	Beta_i(mean)	Lower 95% CI of Beta_i	Upper 95% CI of Beta_i
1	174.3102	168.1024	180.8953	6.048707	5.105033	6.914313
2	167.9970	162.2953	173.6362	5.119530	4.252582	5.984225
3	170.5407	163.7297	177.8351	4.690904	3.659325	5.542255
4	173.6425	168.3729	178.8058	5.856221	5.073306	6.631359

Here is a table, the point estimate of α_i and β_i and 95% credible interval estimate.

Model:

$$y_{i,j} \sim \mathrm{N}(\alpha_i + \beta_i t_{i,j}, \sigma^2),$$

for a certain people i, as $t \uparrow 1$ unit, the $E(y_{ij})$ (Average weight) of the patient will increase β_i on average.

PS: It's a panel / longitudinal data interpretation.

	Alpha	Beta	Sigma	Tau_alpha	Tau_beta
Point Estimate	169.2781	5.026451	9.424936	21.808482	1.3289224
95% CI Lower	167.7988	4.697103	8.038228	9.454264	0.8271465
95% CI Upper	170.6956	5.350685	11.078935	37.446560	2.0297705

Here is a posterior for the other parameters.

ps: All the code is included in the attachments.