Rafael Cátora Pulgrons

6x1

2.11 Exercises

1. Posterior inference: suppose you have a Beta(4, 4) prior distribution on the probability θ that a coin will yield a 'head' when spun in a specified manner. The coin is independently spun ten times, and 'heads' appear fewer than 3 times. You are not told how many heads were seen, only that the number is less than 3. Calculate your exact posterior density (up to a proportionality constant) for θ and sketch it.

$$\pi(\theta) \propto \theta^{4-1}(1-\theta)^{4-1}$$

$$P(X=x|\theta) \propto \theta^{x}(1-\theta)^{1-x} \quad X_i \sim Ber(p)$$

$$Y=\sum_{i=1}^{10} X_i \sim Bin(10,\theta) \Rightarrow$$

$$P(Y \leq 2) = \binom{10}{0}\theta^{0}(1-\theta)^{10} + \binom{10}{1}\theta^{1}(1-\theta)^{9} + \binom{10}{2}\theta^{2}(1-\theta)^{8}$$

So, $$\theta|x \propto \theta^{4-1}(1-\theta)^{4-1} \cdot \left[(1-\theta)^{10} + 10\,\theta^{1}(1-\theta)^{9} + 45\,\theta^{2}(1-\theta)^{8}\right]$$

$$\propto \theta^{4-1}(1-\theta)^{14-1} + 10\,\theta^{5-1}(1-\theta)^{13-1} + 45\,\theta^{6-1}(1-\theta)^{12-1}$$

Sketch is in the markdown file.

Ex 2:

2. Predictive distributions: consider two coins, C_1 and C_2, with the following characteristics: $\Pr(\text{heads}|C_1) = 0.6$ and $\Pr(\text{heads}|C_2) = 0.4$. Choose one of the coins at random and imagine spinning it repeatedly. Given that the first two spins from the chosen coin are tails, what is the expectation of the number of additional spins until a head shows up?

. If $X \sim \text{Geometric}(p) \Rightarrow \mathbb{P}(X=x) = (1-p)^{k-1} p$, $k = 1, 2, \dots$ $E[X] = 1/p$

Let $Y|p$ denotes the number of heads in two spins

$Y|p \sim \text{Bin}(2, p)$ $\qquad \mathbb{P}(p=0.6) = \mathbb{P}(p=0.4) = 1/2$

$$\mathbb{P}(p=0.6|Y=0) = \frac{\mathbb{P}(Y=0|p=0.6)\cdot\mathbb{P}(p=0.6)}{\mathbb{P}(Y=0|p=0.6)\cdot\mathbb{P}(p=0.6) + \mathbb{P}(Y=0|p=0.4)\cdot\mathbb{P}(p=0.4)}$$

$$= \frac{0.4^2 \cdot 1/2}{0.4^2 \cdot 1/2 + 0.6^2 \cdot 1/2} = \frac{0.4^2}{0.4^2 + 0.6^2} \approx 0.307$$

And, $\mathbb{P}(p=0.4|Y=0) = \frac{0.6^2}{0.6^2 + 0.4^2} \approx 0.692$

So, $p|Y=0 = \begin{cases} 0.6 & \text{with prob. } 0.307 \\ 0.4 & \text{with prob } 0.693 \end{cases}$

Now, let's consider $X|p \sim \text{Geometric}(p)$ where X denotes the numbers of failures until one success.

$E[X|p] = 1/p$, and $E[E[X|p]] = E[X]$

$$1/0.6 \cdot 0.307 + 1/0.4 \cdot 0.693 = 2.24$$

Ex 3

5. Posterior distribution as a compromise between prior information and data: let y be the number of heads in n spins of a coin, whose probability of heads is θ.

(a) If your prior distribution for θ is uniform on the range $[0, 1]$, derive your prior predictive distribution for y,

$$\Pr(y = k) = \int_0^1 \Pr(y = k|\theta)d\theta,$$

for each $k = 0, 1, \ldots, n$.

(b) Suppose you assign a Beta(α, β) prior distribution for θ, and then you observe y heads out of n spins. Show algebraically that your posterior mean of θ always lies between your prior mean, $\frac{\alpha}{\alpha+\beta}$, and the observed relative frequency of heads, $\frac{y}{n}$.

a) $Y \sim Bin(n, \theta) \mid P(Y=k|\theta) = \binom{n}{k} \theta^k (1-\theta)^{n-k}$

We want $P(Y=k) = \int_0^1 \binom{n}{k} \theta^k (1-\theta)^{n-k} d\theta = \binom{n}{k} \cdot \underbrace{\int_0^1 \theta^k (1-\theta)^{n-k} d\theta}_{\text{Kernel Beta}(k+1, n-k+1)}$

$$= \binom{n}{k} \frac{\Gamma(k+1)\cdot\Gamma(n-k+1)}{\Gamma(k+1+n-k+1)} =$$

$$= \frac{n!}{\cancel{k!(n-k)!}} \cdot \frac{\cancel{k!\,(n-k)!}}{\Gamma(n+2)} = \frac{n!}{(n+1)\Gamma(n+1)} = \frac{\cancel{n!}}{(n+1)\cdot\cancel{n!}} = \boxed{\frac{1}{n+1}}$$

---

b) $p(\theta) \propto \theta^{\alpha-1}(1-\theta)^{\beta-1}$

$f_{Y|\theta}(y) \propto \theta^y (1-\theta)^{n-y}$ $\Rightarrow$ $\theta | y \propto \theta^{\alpha-1}(1-\theta)^{\beta-1} \cdot \theta^y (1-\theta)^{n-y}$

$\propto \theta^{\alpha+y-1}(1-\theta)^{\beta+n-y-1}$ $\Rightarrow \theta|y \sim Beta(\alpha+y, \beta+n-y)$

Now, we know that the expected value of a Beta$(a,b) = \frac{a}{a+b}$, so

$$E[\theta|y] = \frac{\alpha+y}{\alpha+y+\beta+n-y} = \frac{\alpha+y}{\alpha+\beta+n} = \lambda\frac{\alpha}{\alpha+\beta} + (1-\lambda)\frac{y}{n}$$, now

if we show that $\lambda \in (0,1)$ it means that the expected value of the posterior is a linear combination of the prior mean and the likelihood mean.

Let's now verify that $\lambda \in (0,1)$:

$$\frac{\alpha+y}{\alpha+\beta+n} = \frac{\lambda\alpha}{\alpha+\beta} + \frac{y}{n} - \lambda\frac{y}{n} \Rightarrow \lambda \cdot \left(\frac{\alpha}{\alpha+\beta} - \frac{y}{n}\right) + \frac{y}{n} = \frac{\alpha+y}{\alpha+\beta+n} \Rightarrow$$

$$\Rightarrow \lambda\left(\frac{\alpha}{\alpha+\beta} - \frac{y}{n}\right) = \frac{\alpha+y}{\alpha+\beta+n} - \frac{y}{n} \Rightarrow$$

$$\Rightarrow \lambda \cdot \left(\frac{n\alpha - \alpha y - \beta y}{(\alpha+\beta)\cdot n}\right) = \frac{n\alpha + \cancel{ny} - \alpha y - \beta y - \cancel{ny}}{(\alpha+\beta+n)n}$$

$$\Rightarrow \lambda \left(\frac{n\alpha - \alpha y - \beta y}{(\alpha+\beta)n}\right) = \frac{n\alpha - \alpha y - \beta y}{(\alpha+\beta+n)\cdot n}$$

$$\Rightarrow \lambda = \frac{\cancel{n\alpha - \alpha y - \beta y}}{(\alpha+\beta+n)\cancel{n}} \cdot \frac{(\alpha+\beta)\cancel{n}}{\cancel{n\alpha - \alpha y - \beta y}} = \frac{\alpha+\beta}{\alpha+\beta+n}$$

Since $\lambda \in (0,1)$ we showed that the expected posterior is a weighted average of the prior mean and the likelihood mean //

(c) Show that, if the prior distribution on θ is uniform, the posterior variance of θ is always less than the prior variance.

(d) Give an example of a Beta(α, β) prior distribution and data y, n, in which the posterior variance of θ is higher than the prior variance.

$\theta \sim Unif(0,1)$ and $P(Y=y|\theta) = \binom{n}{y} \theta^y (1-\theta)^{n-y}$

The posterior distribution is given by:

$\theta|y \propto \theta^y (1-\theta)^{n-y} \Rightarrow \theta|y \sim Beta(y+1, n-y+1)$ //

The prior variance is: $\frac{1}{12}$

And the posterior variance is: $\frac{(y+1)\cdot(n-y+1)}{(y+1+n-y+1)^2 \cdot (y+1+n-y+1+1)} =$

$$= \frac{(y+1)\cdot(n-y+1)}{(n+2)^2 \cdot (n+3)} = \frac{1}{(n+3)} \cdot \underbrace{\frac{(y+1)}{(n+2)}}_{(1)} \cdot \underbrace{\frac{(n-y+1)}{(n+2)}}_{(2)} ,$$

note that $y+1+n-y+1 =$ so $0 < \frac{n+1}{n+2} < 1$ and $0 < \frac{n-y+1}{n+2} < 1$

this way $(1)\cdot(2)$ is always smaller than $1/4$

Since $\frac{1}{n+3}$ is always $\leq 1/4$, because $n \geq 1$, than Var

$1/4 \cdot 1/4 < \frac{1}{12}$, so posterior variance is smaller than the prior variance

---

d) As we did in part b, we know that if $\theta \sim Beta(\alpha, \beta)$ and having observed y heads in n draws, $\theta|y \sim Beta(\alpha+y, \beta+n-y)$

$$Var[\theta|y] = \frac{(\alpha+y)\cdot(\beta+n-y)}{(\alpha+y+\beta+n-y)^2 \cdot (\alpha+y+\beta+n-y+1)} = \frac{(\alpha+y)\cdot(\beta+n-y)}{(\alpha+\beta+n)^2 (\alpha+\beta+n+1)}$$

this way, now we want an example of α, β such that

$$Var[\theta] > Var[\theta|y] \Rightarrow \frac{\alpha \cdot \beta}{(\alpha+\beta)^2(\alpha+\beta+1)} > \frac{(\alpha+y)\cdot(\beta+n-y)}{(\alpha+\beta+n)^2(\alpha+\beta+n+1)}$$

now note that if n is large, $Var[\theta|y] \to 0$, so it's harder to verify the statement under this condition. So let's set $n=1$ and $y=0$.

$$Var[\theta|y] = \frac{(\alpha+0)\cdot(\beta+1-0)}{(\alpha+\beta+1)^2(\alpha+\beta+2)} = \frac{\alpha\beta+\alpha}{(\alpha+\beta+1)^2(\alpha+\beta+2)},$$ Now, setting $\alpha=4$, $\beta=1$

$$Var[\theta] = \frac{4}{4^2\cdot(5)} = \frac{1}{20} \quad \text{and} \quad Var[\theta|y] = \frac{8}{5^2\cdot 6} = \frac{8}{25\cdot 6} = \frac{8}{150} = 4/75$$

Note that $1/20 < 4/75$, so for $n=1, y=0, \alpha=4$, and $\beta=1$ the posterior variance is greater than the prior variance

Ex 4:

7. Noninformative prior densities:

(a) For the binomial likelihood, $y \sim \text{Bin}(n, \theta)$, show that $p(\theta) \propto \theta^{-1}(1-\theta)^{-1}$ is the uniform prior distribution for the natural parameter of the exponential family.

(b) Show that if $y = 0$ or n, the resulting posterior distribution is improper.

a) $p(y|n,\theta) = \binom{n}{y} \theta^y (1-\theta)^{n-y}$, let's rewrite it in the exponential family form

$$p(y_i|\theta) = f(y_i)g(\theta)e^{\phi(\theta)^T u(y_i)}$$

$$p(y|n,\theta) = \exp\{y \ln(\theta) + n \ln(1-\theta) - y \ln(1-\theta)\}\binom{n}{y}$$

$$= \binom{n}{y} \exp\{y \ln\left(\frac{\theta}{1-\theta}\right) + n\ln(1-\theta)\}$$

this way, $f(y) = \binom{n}{y}$, $u(y) = y$, $\phi(\theta) = \ln\left(\frac{\theta}{1-\theta}\right)$ and $g(\theta) = \exp\{n\ln(1-\theta)\}$

thus, if $\phi(\theta) = \ln\left(\frac{\theta}{1-\theta}\right) \Rightarrow e^{\phi} = \frac{\theta}{1-\theta} \Rightarrow e^{\phi}(1-\theta) = \theta \Rightarrow e^{\phi} - \theta e^{\phi} = \theta$

$\Rightarrow e^{\phi} = \theta + \theta e^{\phi} \Rightarrow \theta = \frac{e^{\phi}}{1+e^{\phi}}$ //

We want a distribution proportional to a constant on the parametric space of the natural parameter

this way in order for $p(\phi) \propto 1$, with $\phi = g(\theta)$ we have that

$$p(\theta) \propto 1 \left|\frac{dg(\theta)}{d\theta}\right| \Rightarrow p(\theta) \propto \left|\ln\left(\frac{\theta}{1-\theta}\right)\right| \Rightarrow p(\theta) \propto \left|\ln(\theta) - \ln(1-\theta)\right|$$

$$\Rightarrow p(\theta) \propto \left|\frac{1}{\theta} + \frac{1}{(1-\theta)}\right| \Rightarrow p(\theta) \propto \left|\frac{(1-\theta)+\theta}{\theta(1-\theta)}\right| \Rightarrow p(\theta) \propto \frac{1}{\theta(1-\theta)} \Rightarrow$$

$\Rightarrow p(\theta) \propto \theta^{-1}(1-\theta)^{-1}$ //

Ex 5

11. Computing with a nonconjugate single-parameter model: suppose $y_1, \ldots, y_5$ are independent samples from a Cauchy distribution with unknown center θ and known scale 1: $p(y_i|\theta) \propto 1/(1 + (y_i - \theta)^2)$. Assume, for simplicity, that the prior distribution for θ is uniform on $[0, 100]$. Given the observations $(y_1, \ldots, y_5) = (43, 44, 45, 46.5, 47.5)$:

   (a) Compute the unnormalized posterior density function, $p(\theta)p(y|\theta)$, on a grid of points $\theta = 0, \frac{1}{m}, \frac{2}{m}, \ldots, 100$, for some large integer m. Using the grid approximation, compute and plot the normalized posterior density function, $p(\theta|y)$, as a function of θ.

   (b) Sample 1000 draws of θ from the posterior density and plot a histogram of the draws.

   (c) Use the 1000 samples of θ to obtain 1000 samples from the predictive distribution of a future observation, y_6, and plot a histogram of the predictive draws.

MarKdown file

6x6

6. Consider a distribution with p.d.f

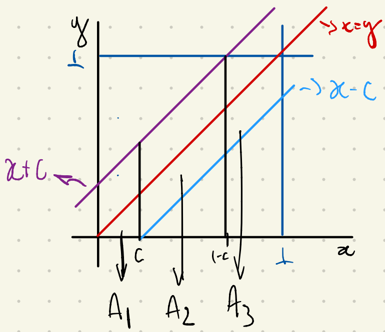
$$p(x, y) \propto \mathbf{1}(|x-y| < c)\mathbf{1}(x, y \in (0,1))$$

(a) Derive a Gibbs sampler to sample from this distribution.

(b) Implement the Gibbs sampler for 1000 iterations for each of the following: $c = 0.3, c = 0.05$ and $c = 0.01$.

(c) For each value of the c make a traceplot of x and a scatter plot of (x, y).

(d) What do you see as c gets smaller?



if $x < y \Rightarrow |x-y| = y - x$ (1)

$x \geq y \Rightarrow |x-y| = x - y$ (2)

for (1): $y - x < c \Leftrightarrow y < x + c$

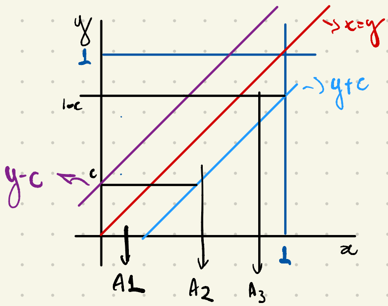
for (2) $x - y < c \Leftrightarrow x < y + c \Rightarrow y > x - c$

$$p(x,y) = \mathbb{1}(|x-y|<c)\,\mathbb{1}(x,y \in (0,1))$$

first, let's find $p(x|y)$ and $p(y|x)$. Let's work on the marginal

$$p(x) \propto \int p(x,y)\,dy \Rightarrow \begin{cases} \int_0^{x+c} dy = x + c, \text{ if } x < c \\ \int_{x-c}^{x+c} dy = 2c, \text{ if } c < x < 1-c \\ \int_{x-c}^{1} dy = 1 - x + c, \text{ if } x > 1-c \end{cases}$$

$$\text{So, } p(y|x) \propto \frac{\mathbb{1}(|x-y|<c)\,\mathbb{1}(x,y \in (0,1))}{(x+c)\cdot\mathbb{1}_{(0,c)}(x) + 2c\,\mathbb{1}_{(c,1-c)}(x) + (1-x+c)\cdot\mathbb{1}_{(1-c,1)}(x)}$$



Now let's find $p(x|y) \propto \dfrac{p(x,y)}{p(y)}$ and

$$\boxed{\begin{array}{lll} x<y \Rightarrow |x-y| = y-x \Rightarrow y-x<c \Leftrightarrow x>y-c & \Rightarrow & x=y-c \Leftrightarrow y=x+c \\ x \geq y \Rightarrow |x-y| = x-y \Rightarrow x-y<c \Leftrightarrow x<y+c & \Rightarrow & x=y+c \Leftrightarrow y=x-c \end{array}}$$

$$p(y) = \int p(x,y)\,dx = \begin{cases} \text{if } y<c = \int_0^{y+c} dx = y+c \\ \text{if } c<y<1-c = \int_{y-c}^{y+c} dx = y+c-y+c = 2c \\ \text{if } 1-c<y\leq 1 \Rightarrow \int_{y-c}^{1} dx = 1-y-c = 1-y+c \end{cases}$$

So, $p(x|y) \propto \dfrac{\mathbb{1}(|x-y|<c)\,\mathbb{1}(x,y\in(0,1))}{(y+c)\cdot\mathbb{1}_{(0,c)}(y) + 2c\,\mathbb{1}_{(c,1-c)}(y) + (1-y+c)\cdot\mathbb{1}_{(1-c,1)}(y)}$

Now, with the conditional distribution on hands, we can derive our Gibb's sampler.

If they are proportional to a constant in an interval, after normalized we can find the exact conditional distribution, which is uniform.

$y|X=x \sim \mathrm{Unif}(a,b)$ with $a=\max(0,x-c)$ and $b=\min(x+c,1)$

$x|Y=y \sim \mathrm{Unif}(a^*,b^*)$ with $a^*=\max(0,y-c)$ and $b=\min(y+c,1)$

Ex 7

3.10 Exercises

1. Binomial and multinomial models: suppose data $(y_1, \dots, y_J)$ follow a multinomial distribution with parameters $(\theta_1, \dots, \theta_J)$. Also suppose that $\theta = (\theta_1, \dots, \theta_J)$ has a Dirichlet prior distribution. Let $\alpha = \frac{\theta_1}{\theta_1+\theta_2}$.

   (a) Write the marginal posterior distribution for α.

   (b) Show that this distribution is identical to the posterior distribution for α obtained by treating y_1 as an observation from the binomial distribution with probability α and sample size $y_1 + y_2$, ignoring the data $y_3, \dots, y_J$.

a) First, let's work out on the posterior distribution:

$$p(\underset{\sim}{\theta} \mid \underset{\sim}{y}) \propto p(\underset{\sim}{y} \mid \underset{\sim}{\theta}) \cdot p(\theta) \propto \prod_{i=1}^{J} \theta_i^{y_i} \prod_{i=1}^{J} \theta_i^{a_i - 1}$$

$$\propto \prod_{i=1}^{J} \theta_i^{a_i + y_i - 1} \Rightarrow \underset{\sim}{\theta} \mid \underset{\sim}{y} \sim \text{Dir}(a_1 + y_1, a_2 + y_2, \dots, a_J + y_J)$$

We know that the Dirichlet distribution has a property that

$$p(\theta_1, \theta_2 \mid y) \propto \theta_1^{y_1 + a_1} \theta_2^{y_2 + a_2} (1 - \theta_1 - \theta_2)^{\left(\sum_{i=3}^{J} y_j + a_i\right) - 1},$$

Now, let's use a transformation of variables. Let $\alpha = \frac{\theta_1}{\theta_1 + \theta_2}$ and $\beta = \theta_1 + \theta_2$

↳ $\alpha = \frac{\theta_1}{\beta}$; $\theta_1 = \alpha\beta$; $\beta = \alpha\beta + \theta_2$ and $\theta_2 = \beta(1-\alpha)$ So,

$$|J| = \begin{vmatrix} \frac{d\theta_1}{d\alpha} & \frac{d\theta_2}{d\alpha} \\ \frac{d\theta_1}{d\beta} & \frac{d\theta_2}{d\beta} \end{vmatrix} = \begin{vmatrix} \beta & -\beta \\ \alpha & 1-\alpha \end{vmatrix} = \beta - \cancel{\beta\alpha} + \cancel{\beta\alpha} = \beta$$

$$p(\alpha, \beta \mid y) \propto \beta (\alpha\beta)^{y_1 - a_1 - 1} (\beta(1-\alpha))^{y_2 + a_2 - 1} (1-\beta)^{\left(\sum_{i=2}^{J} y_i + a_i\right) - 1}$$

$$\propto \underbrace{\alpha^{y_1 + a_1 - 1} (1-\alpha)^{y_2 + a_2 - 1}}_{\alpha \mid y \sim \text{Beta}(y_1 + a_1,\ y_2 + a_2)} \ \underbrace{\beta^{y_1 + y_2 + a_1 + a_2 - 1} (1-\beta)^{\left(\sum_{i=1}^{J} y_i + a_i\right) - 1}}_{\beta \mid y \sim \text{Beta}\left(y_1 + y_2 + a_1 + a_2,\ \sum_{i=1}^{J} y_i + a_i\right)}$$

Since we were able to factor the joint distribution we have that $\alpha \perp \beta$ and that the marginal posterior for $\alpha \mid y \sim \text{Beta}(y_1 + a_1, y_2 + a_2)$ //

part b)

let $y_1|\alpha \sim \text{Bin}(y_1+y_2, \alpha)$ now let $\alpha \sim \text{Beta}(a_1, a_2)$

$$p(\alpha|y_1) \propto \alpha^{y_1}(1-\alpha)^{y_2}\,\alpha^{a_1-1}(1-\alpha)^{a_2-1}$$

$$\propto \alpha^{y_1+a_1-1}(1-\alpha)^{y_2+a_2-1}$$

so $\alpha|y_i \sim \text{Beta}(y_1+a_1, y_2+a_2)$

Ex 8

4. Inference for a 2×2 table: an experiment was performed to estimate the effect of beta-blockers on mortality of cardiac patients. A group of patients were randomly assigned to treatment and control groups: out of 674 patients receiving the control, 39 died, and out of 680 receiving the treatment, 22 died. Assume that the outcomes are independent and binomially distributed, with probabilities of death of p_0 and p_1 under the control and treatment, respectively. We return to this example in Section 5.6.
   (a) Set up a noninformative prior distribution on (p_0, p_1) and obtain posterior simulations.
   (b) Summarize the posterior distribution for the *odds ratio*, $(p_1/(1-p_1))/(p_0/(1-p_0))$.
   (c) Discuss the sensitivity of your inference to your choice of noninformative prior density.

8. BDA3 Problem 3.4. Also compute the following two quantities for the problem: (1) Posterior probabilities of $p_0 > p_0^{(MLE)}$ and $p_1 > p_1^{(MLE)}$ respectively; (2) Posterior probability of the odds ratio being smaller than 1. Also repeat the problem with the modified sample size so that 34 patients were assigned to each group with 2 died in the control and 1 died in the treatment group.

control: $n_c = 674$ $\quad y_c = 39$ $\quad \bar{y}_c = \frac{39}{674}$

treatment: $n_t = 680$ $\quad y_t = 22$ $\quad y_t = \frac{22}{680}$

$$p_0, p_1 \propto \mathbb{1}_{(0,1)}(p_0)\ \mathbb{1}_{(0,1)}(p_1) \Rightarrow \begin{array}{l} p_0 \sim \text{Beta}(1,1) \\ p_1 \sim \text{Beta}(1,1) \end{array}$$

Now, let's find the posterior distribution

$$p_0 | y_c \propto p_0^{1-1}(1-p_0)^{1-1} \cdot p_0^{39}(1-p_0)^{674-39} \Rightarrow p_0 | y \sim \text{Beta}(40, 636)$$

$$p_1 | y_t \propto p_1^{1-1}(1-p_1)^{1-1} \cdot p_1^{22}(1-p_1)^{680-22} \Rightarrow p_1 | y \sim \text{Beta}(23, 659)$$

Now, we can draw from the posterior.

Exercise continues on the markdown file.

Ex 9

5. Rounded data: it is a common problem for measurements to be observed in rounded form (for a review, see Heitjan, 1989). For a simple example, suppose we weigh an object five times and measure weights, rounded to the nearest pound, of 10, 10, 12, 11, 9. Assume the unrounded measurements are normally distributed with a noninformative prior distribution on the mean μ and variance σ^2.

(a) Give the posterior distribution for (μ, σ^2) obtained by pretending that the observations are exact unrounded measurements.

(b) Give the correct posterior distribution for (μ, σ^2) treating the measurements as rounded.

(c) How do the incorrect and correct posterior distributions differ? Compare means, variances, and contour plots.

(d) Let $z = (z_1, \ldots, z_5)$ be the original, unrounded measurements corresponding to the five observations above. Draw simulations from the posterior distribution of z. Compute the posterior mean of $(z_1 - z_2)^2$.

a) using a non-informative prior for μ, σ^{-2} assumming independence between μ and σ we have that $p(\mu, \sigma^2) \propto \frac{1}{\sigma^2}$ which is the Jeffreys prior.

Also, $p(y \mid \mu, \sigma) \sim \text{Normal}(\mu, \sigma^2)$

This way
$$p(\mu, \sigma^2 \mid y) \propto \frac{1}{\sigma^2} \cdot (\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^2\right\}$$
$$\propto \frac{1}{\sigma^2} (\sigma)^{-n} \exp\left\{\frac{-1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^2\right\}$$
$$\propto \sigma^{-n-2} \exp\left\{\frac{-1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^2\right\}$$
$$\propto \underbrace{\sigma^{-n-2} \exp\left\{\frac{-1}{2\sigma^2} [(n-1)s^2 + n(\bar{y} - \mu)^2]\right\}}$$

$p(\mu, \sigma^2 \mid y) = p(\mu \mid \sigma^2, y) \cdot p(\sigma^2 \mid y)$

$\sigma^2 \mid y \sim \text{Inv-}\chi^2$

$\mu \mid \sigma^2, y \sim N(\bar{y}, \sigma^2/n)$

Ex 10

10. Let $Y \sim N(\theta, 1)$ and $\theta \sim \text{Cauchy}(0, 1)$. Derive the posterior of $\theta|y$ up to a proportional constant. Describe and implement a Metropolis-Hastings sampler to generate posterior draws of $\theta|y$ for $y = 1.5$.

$f_{Y|\theta}(y) \propto \exp\left\{-\frac{1}{2}(y-\theta)^2\right\}$ | $p(\theta) \propto \frac{1}{(1+\theta^2)}$

$$\pi(\theta|y) \propto \frac{\exp\left\{-\frac{1}{2}(y-\theta)^2\right\}}{(1+\theta^2)}$$

The rest is on the markdown file.

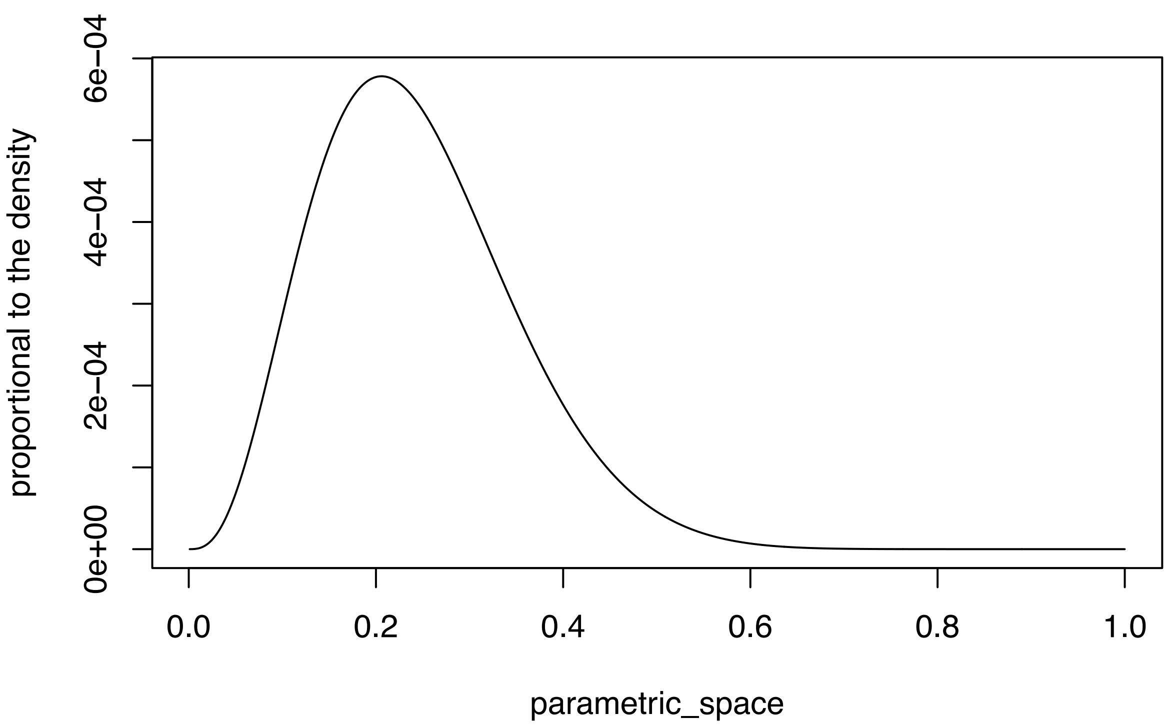
HW1_207

RafaelCatoiaPulgrossi

2023-04-06

Ex 1

```
posterior_prop <- function(theta){
  return(
    (theta^3)*(1-theta)^13+
      (theta^4)*(1-theta)^12+
      (theta^5)*(1-theta)^11
  )
}
parametric_space <- seq(0.001,1,0.001)

plot(parametric_space,posterior_prop(theta=parametric_space),
     type='l',ylab='proportional to the density')
```

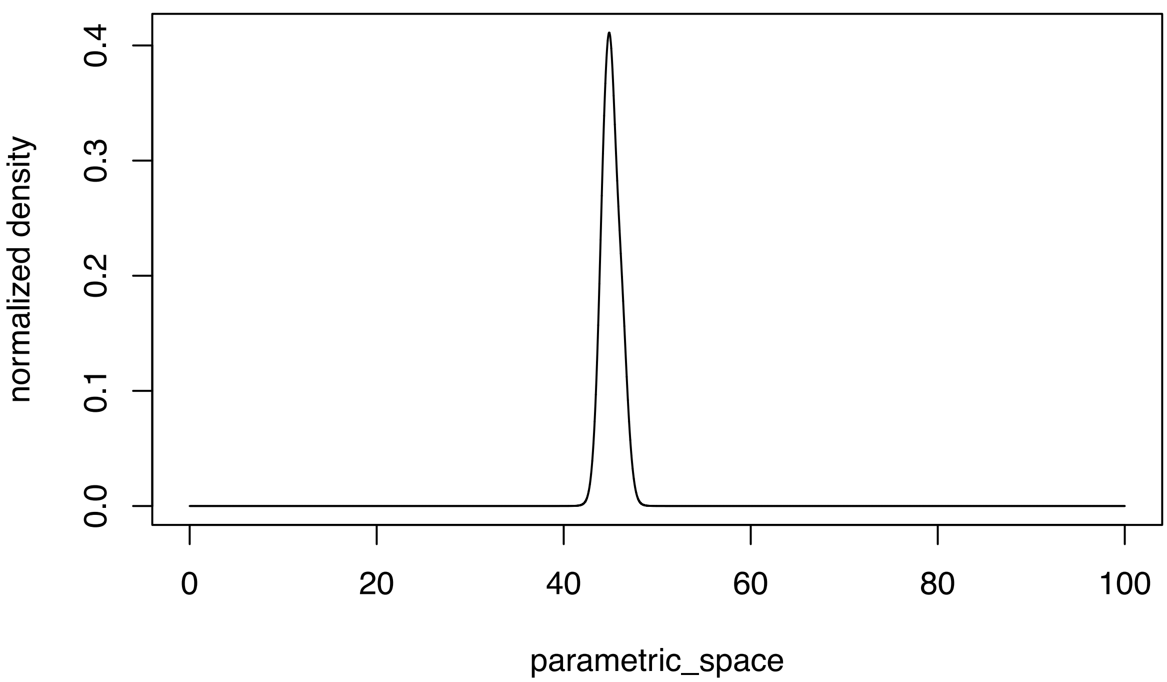


Ex 5

a

Here is the normalized density.

```
y <- c(43,44,45,46.5,47.5)
M=1000

parametric_space <- seq(0,100,1/M)

q_theta_given_y = function(x,theta){
  result <- 1
  for ( i in 1:length(x)){
    result = result * (1/(1+(x[i]-theta)^2))
  }
  return(result/100)
}


## unormalized density
unormalized_density <- q_theta_given_y(y,parametric_space)

## summing the heights of the bins
normalizing_constant <-  sum(unormalized_density/M)

#plotting the normalized density
plot(parametric_space,
     unormalized_density/normalizing_constant,
     type='l',
     ylab = 'normalized density')
```

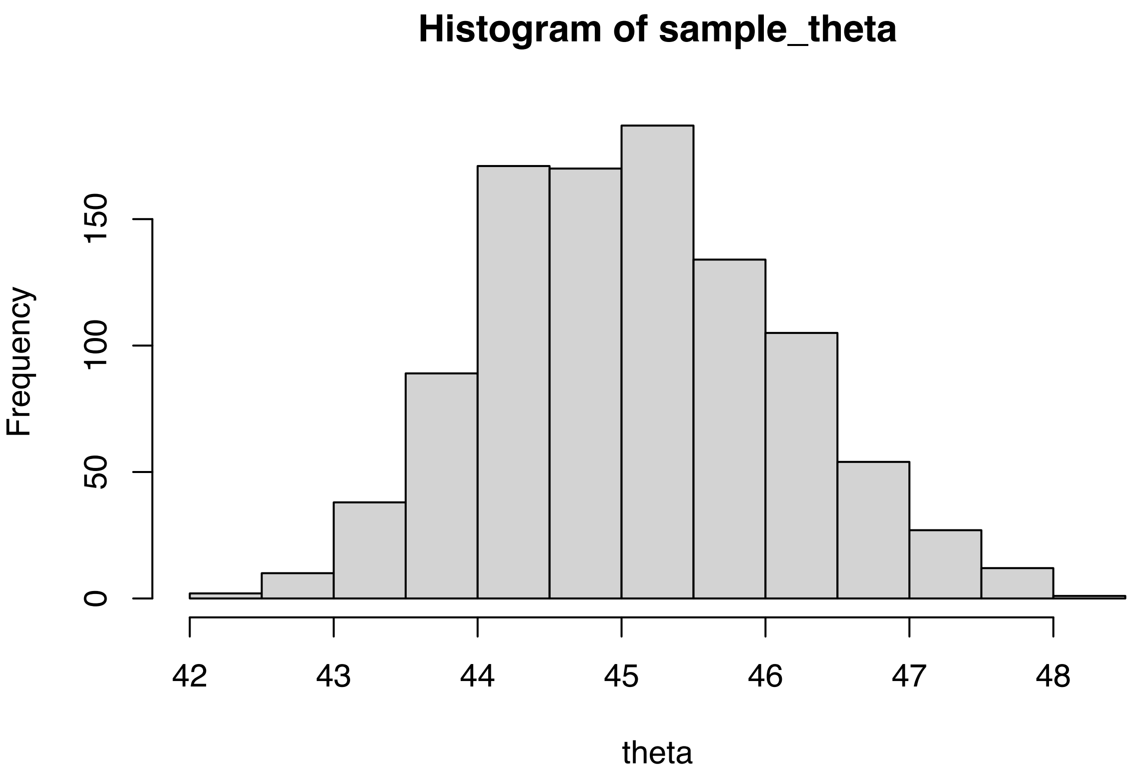



```
normalized_density <-  unormalized_density/normalizing_constant
```

b

Sampling from the normalized density

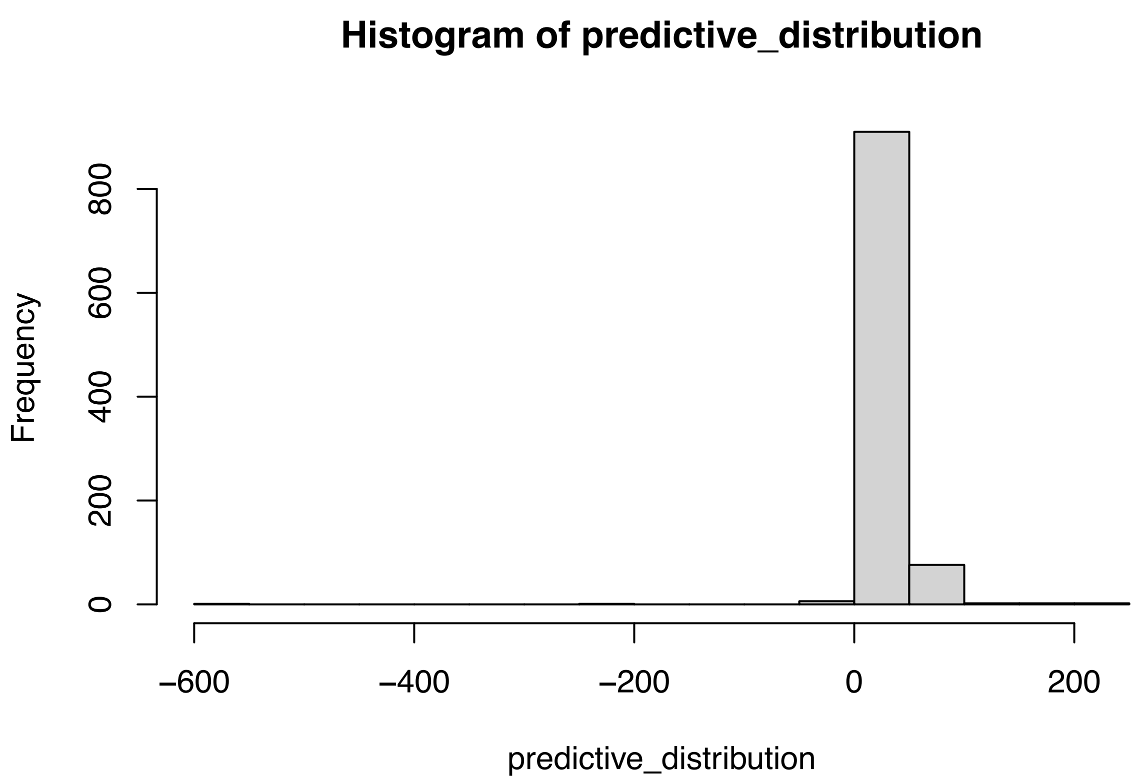
```
sample_theta <- sample(parametric_space,
                       prob = normalized_density,size = 1000)
hist(sample_theta,xlab='theta')
```



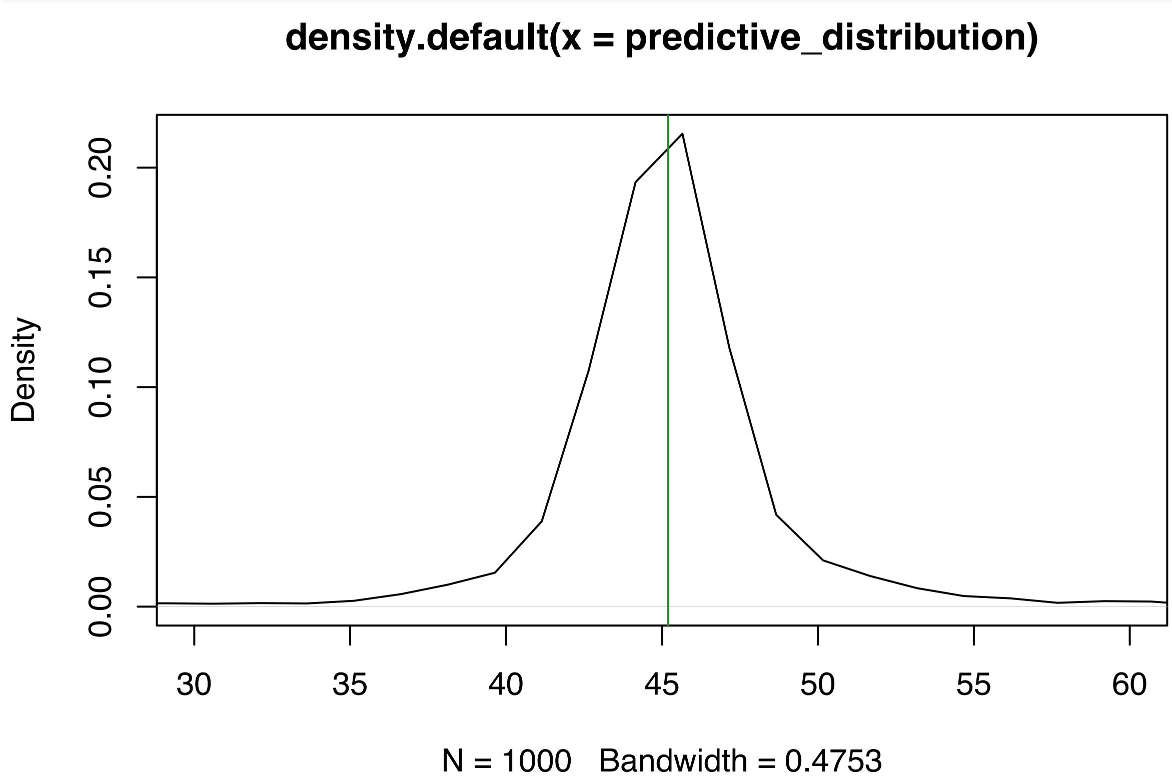
c

Histogram for the predictive distribution.

```
set.seed(57)
predictive_distribution <-
  rcauchy(1000,scale = 1,location = sample_theta)
hist(predictive_distribution)
```




```r
#zoom
plot(density(predictive_distribution),xlim=c(30,60))
abline(v=mean(y),col='forestgreen')
```



```
mean(predictive_distribution) ; sd(predictive_distribution)
```

```
## [1] 44.7186
```

```
## [1] 24.22165
```

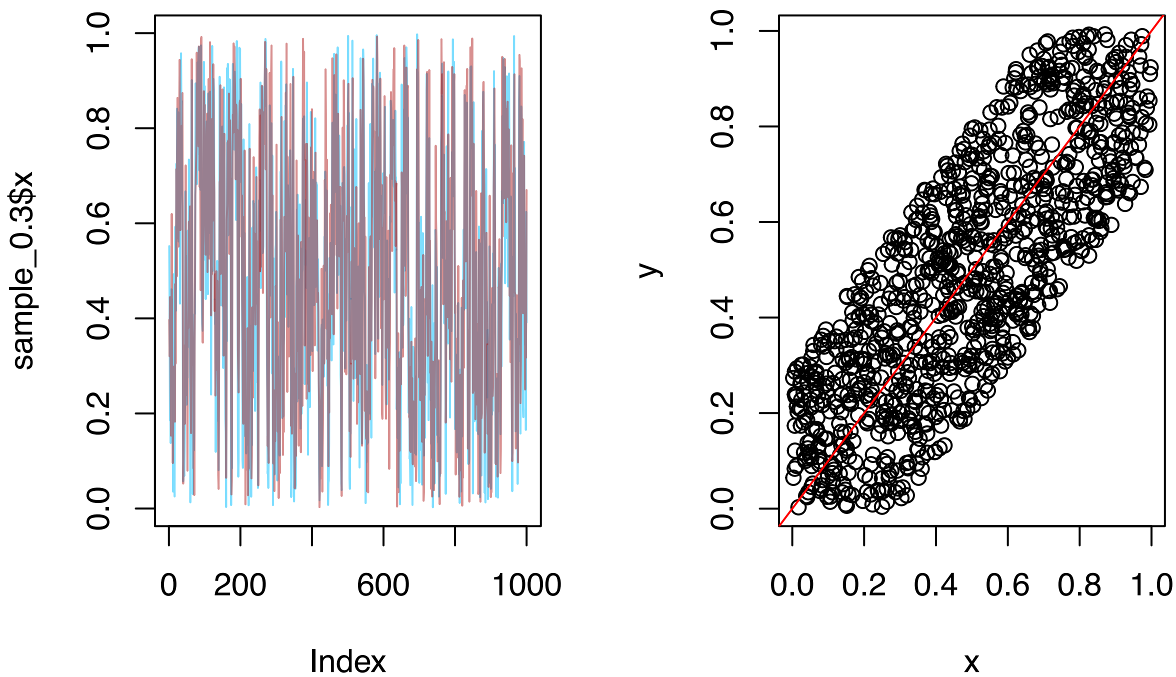
Ex 6

Here we are implementing the Gibbs Sampler.

```
x_given_y <- function(constant,y){
  a <- max(0,y-constant)
  b <- min(y+constant,1)
  return(runif(1,a,b))
}

y_given_x <- function(constant,x){
  a <- max(0,x-constant)
  b <- min(x+constant,1)
  return(runif(1,a,b))
}


gibbs_ex6 <- function(B=1000,constant){
  set.seed(1000)
  samples <- data.frame(x=NA,y=NA)
  new_x <- 0.5 #initial value for x
  for(i in 1:(B)){
    #calculating p(y|x=new_x)
    new_y <- y_given_x(x = new_x,constant = constant)
    #calculating p(x|y=new_y)
    new_x <- x_given_y(y = new_y,constant = constant)
    samples <-rbind(samples, c(new_x,new_y))
  }
  return(samples[-1,])
}

sample_0.3 <- gibbs_ex6(B = 1000,0.3)
sample_0.05 <- gibbs_ex6(B = 1000,0.05)
sample_0.01 <- gibbs_ex6(B = 1000,0.01)
```

For C=0.3

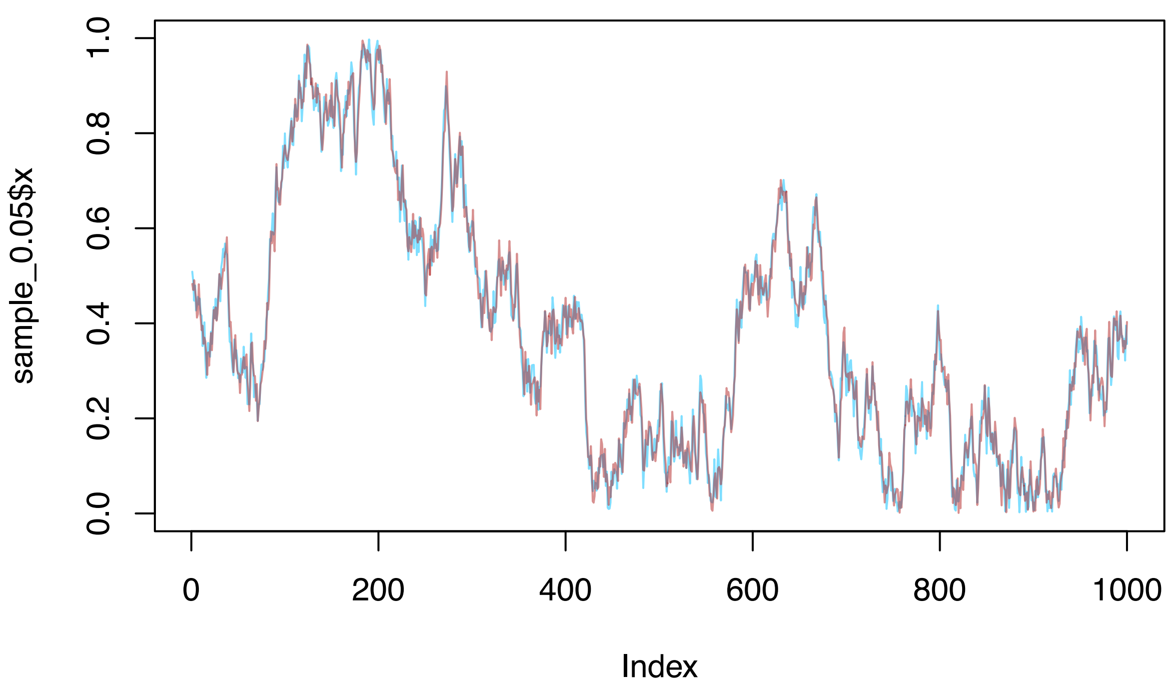
```
par(mfrow=c(1,2))
constant<-0.3
plot(sample_0.3$x,type='l',col=scales::alpha('deepskyblue',0.5))
lines(sample_0.3$y,type='l',col=scales::alpha('firebrick',0.5))
plot(sample_0.3)
abline(0,1,col='red')
```



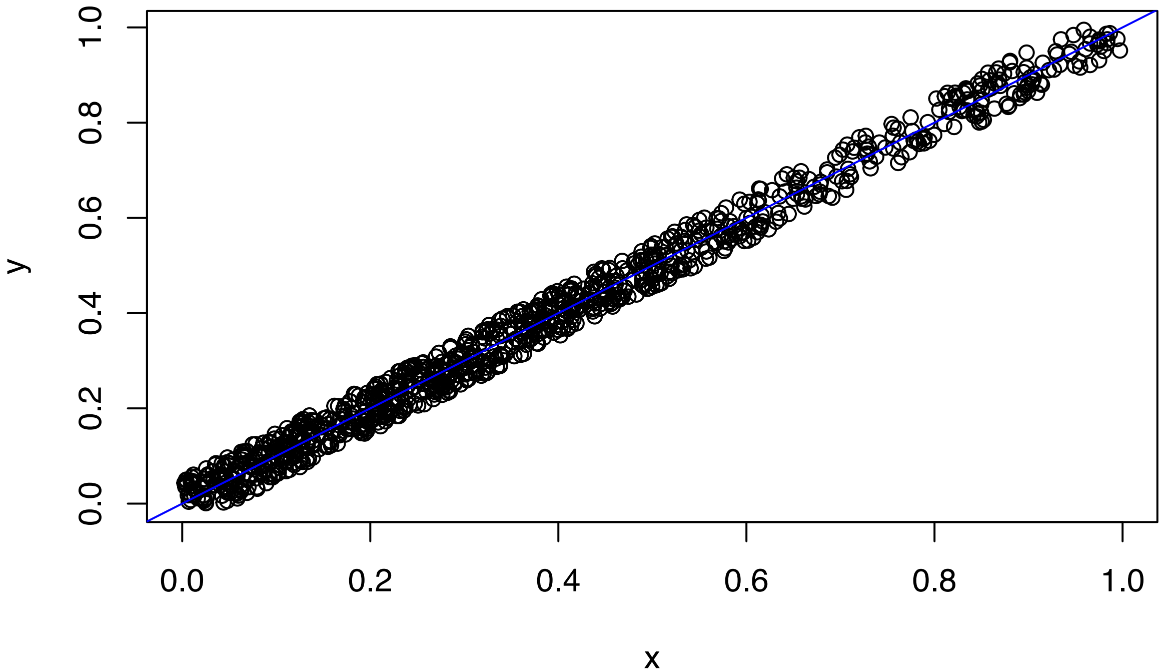

```
#acf(sample_0.3$x)
#acf(sample_0.3$y)
```

For C=0.05

```
constant<-0.05
plot(sample_0.05$x,type='l',col=scales::alpha('deepskyblue',0.5))
lines(sample_0.05$y,type='l',col=scales::alpha('firebrick',0.5))
```



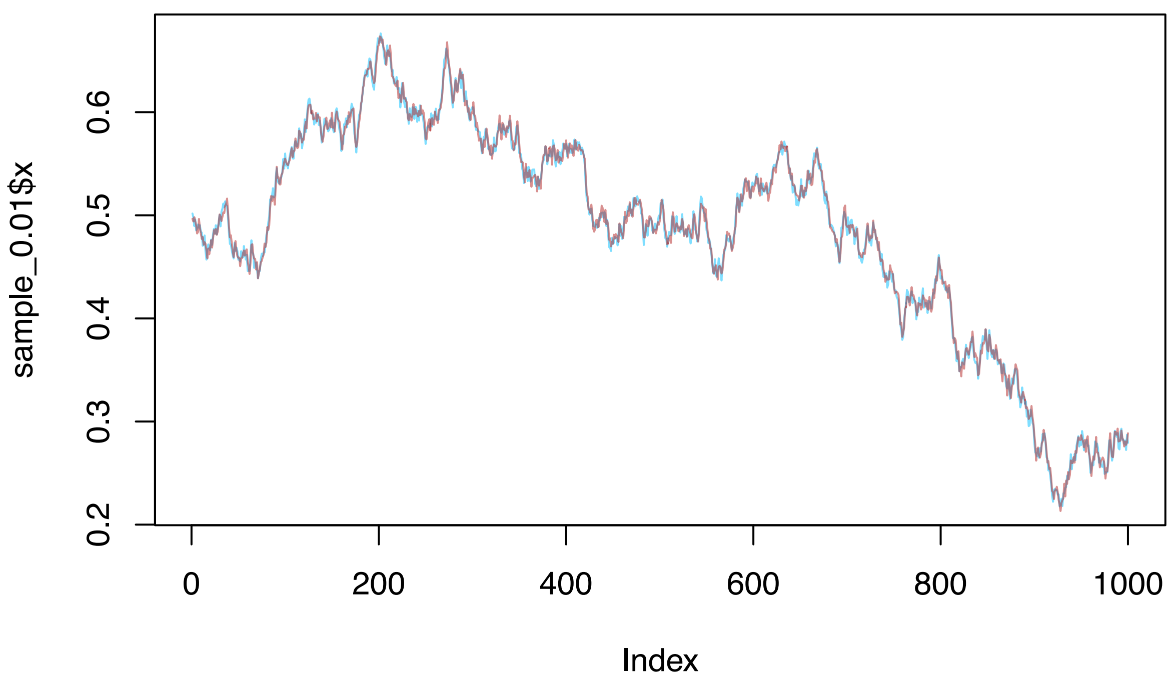

```
plot(sample_0.05)
abline(0,1,col='blue')
```



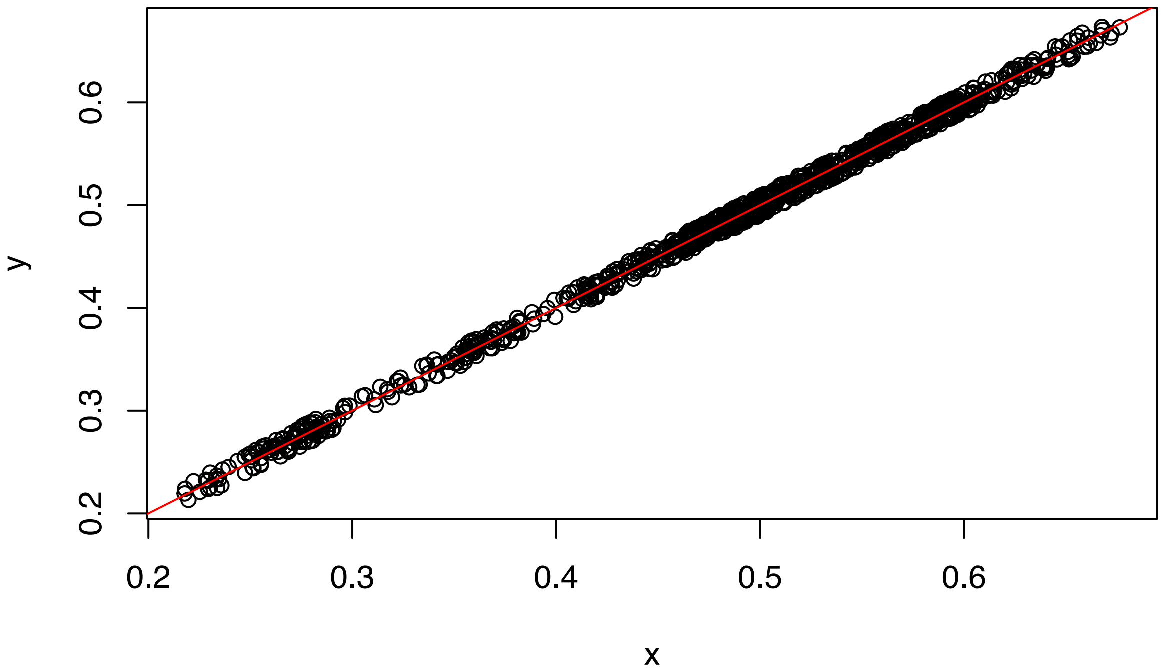

```
#acf(sample_0.05$x)
#acf(sample_0.05$y)
```

For C=0.01

```
constant<-0.01
plot(sample_0.01$x,type='l',col=scales::alpha('deepskyblue',0.5))
lines(sample_0.01$y,type='l',col=scales::alpha('firebrick',0.5))
```




```
plot(sample_0.01)
abline(0,1,col='red')
```




```
#acf(sample_0.01$x)
#acf(sample_0.01$y)
par(mfrow=c(1,1))
```

When C is small, the chain is very dependent on the past observations and as we can see on the last plot, it hasn't visited the entire parametric space yet.

Ex 8 - 3.4

```
nsim <- 10000
posterior_simulation <- data.frame(
  p0 = rbeta(nsim,40,636),
  p1 = rbeta(nsim,23,569)
)
```

Summary for the posterior sample of p_0 and p_1

```
posterior_simulation =
  posterior_simulation %>%
  mutate(odds_p0=p0/(1-p0),
         odds_p1=p1/(1-p1),
         odds_ratio = odds_p1/odds_p0)
```

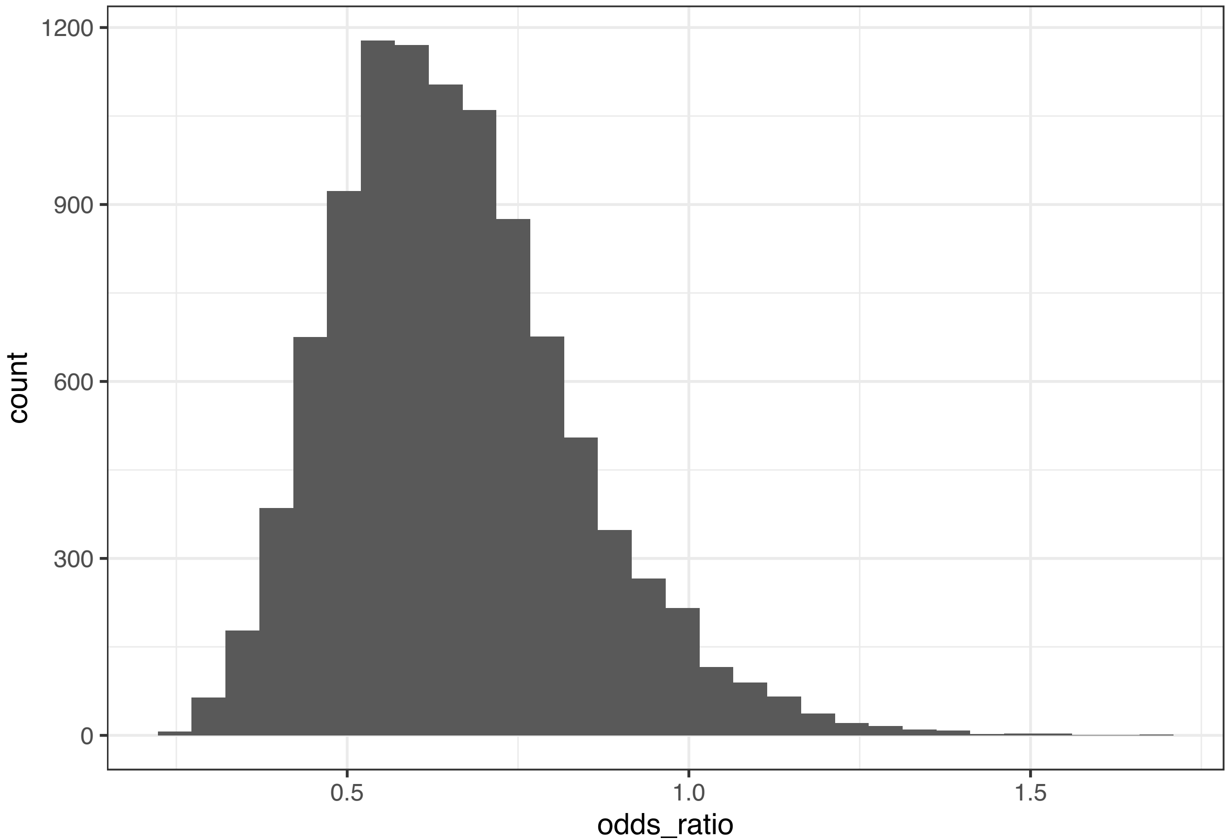
Summary for the posterior sample of odds ratio

```
posterior_simulation %>%
  summarise(
    Min = min(odds_ratio),
    Mean = mean(odds_ratio),
    Median = median(odds_ratio),
    SD = sd(odds_ratio),
    Max = max(odds_ratio)) %>%
  knitr::kable() %>% kableExtra::kable_styling()
```

Min	Mean	Median	SD	Max
0.252106	0.6583732	0.6367122	0.1774619	1.688666

```
posterior_simulation %>%
  ggplot(aes(x=odds_ratio))+
  geom_histogram() +
  theme_bw()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Calculating some posterior probabilities

```
y_bar_control = 39/674
y_bar_treatment = 22/680

#Posterior probability of p0 > p_control_MLE
sum(posterior_simulation$p0 > y_bar_control)/nsim
```

```
## [1] 0.5367
```

```
#Posterior probability of p0 > p_treat_MLE
sum(posterior_simulation$p1 > y_bar_treatment)/nsim
```

```
## [1] 0.7863
```

```
#Posterior probability of the odds ratio being smaller than 1
sum(posterior_simulation$odds_ratio < 1)/nsim
```

```
## [1] 0.9583
```

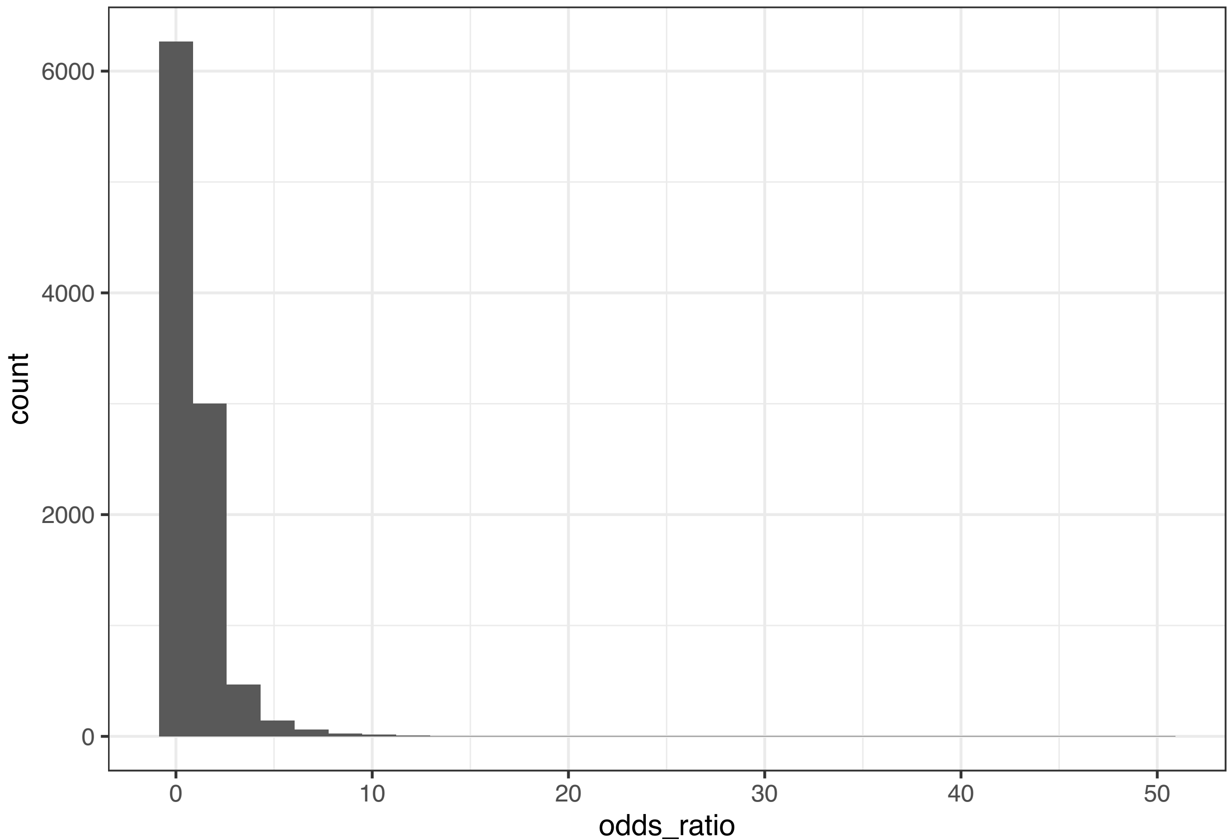
Min	Mean	Median	SD	Max
0.0015135	1.022259	0.6217633	1.451283	50.0575

Now changing the sample size

```
nsim <- 10000
posterior_simulation <- data.frame(
  #34 patients in each group
  p0 = rbeta(nsim,3,35),
  p1 = rbeta(nsim,2,35)
)


posterior_simulation =
  posterior_simulation %>%
  mutate(odds_p0=p0/(1-p0),
         odds_p1=p1/(1-p1),
         odds_ratio = odds_p1/odds_p0)

posterior_simulation %>%
  summarise(
    Min = min(odds_ratio),
    Mean = mean(odds_ratio),
    Median = median(odds_ratio),
    SD = sd(odds_ratio),
    Max = max(odds_ratio)) %>%
  knitr::kable() %>% kableExtra::kable_styling()
```

```
posterior_simulation %>%
  ggplot(aes(x=odds_ratio))+
  geom_histogram() +
  theme_bw()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
y_bar_control = 2/34
y_bar_treatment = 1/34

#Posterior probability
sum(posterior_simulation$p0 > y_bar_control)/nsim
```

```
## [1] 0.6288
```

```
#Posterior probability of p0 > p_treat_MLE
sum(posterior_simulation$p1 > y_bar_treatment)/nsim
```

```
## [1] 0.7147
```

```
#Posterior probability of the odds ratio being smaller than 1
sum(posterior_simulation$odds_ratio < 1)/nsim
```

```
## [1] 0.6801
```

Ex 9

a)

Lets use as non-informative prior $p(\mu, \sigma) \propto \frac{1}{\sigma^2}$, our data-sampling distribution is given by $p(y|\mu, \sigma^2) \sim N(\mu, \sigma^2)$.

Thus, our posterior distribution is going to be:

$$p(\mu, \sigma^2) = p(\mu|\sigma^2, y)p(\sigma^2|y)$$

being $\mu|\sigma^2, y \sim N(\bar{y}, \frac{\sigma^2}{n})$ and $\sigma^2|y \sim Inv\chi^2(n-1, s^2)$

```
y = c(10,10,12,11,9)
ybar = mean(y)
n = length(y)
S2 = var(y)


set.seed(1234)
generate_posterior_untruncated <- function(y){
  ybar = mean(y)
  n = length(y)
  S2 = var(y)
  sigma2 <- (n-1)*S2/rchisq(n = 1,df = (n-1))
  mu <- rnorm(1,mean=ybar,sd=sqrt(sigma2)/sqrt(n))
  return(data.frame(mu,sigma2))
}
```

```
#Calculating the posterior density for the untruncated likelihood

dchisqncp <- function(x, df, ncp){
  (df/2)^(df/2)/gamma(df/2) * sqrt(ncp)^df * x^(-df/2 - 1) * exp(-df*ncp/(2*x))
}

density_posterior_untruncated <- function(mu,sigma2,y){
  ybar = mean(y)
  S2 = var(y)
  n=length(y)
  return(
    dnorm(mu,mean = ybar,sd=sqrt(sigma2)/sqrt(n))*
      dchisqncp(sigma2,df = n-1,ncp = S2)
  )
}
```

b)

Now our likelihood is different, so does he posterior which is given by: $p(\mu, \sigma^2|y) \propto \frac{1}{\sigma^2} \prod_{i=1}^{n} \left[\Phi(\frac{y_i-\mu+0.5}{\sigma}) - \Phi(\frac{y_i+\mu-0.5}{\sigma})\right]$

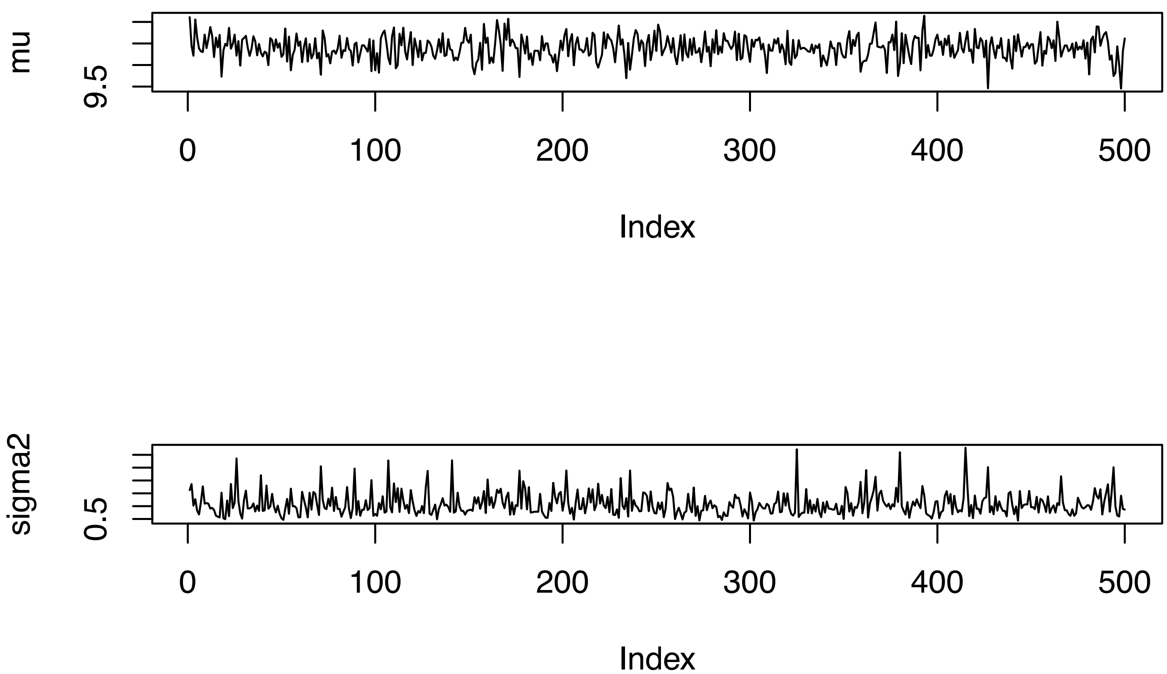
```
#proportional posterior density
prop_density_post_truncated <- function(y,sigma2,mu){
  prop_density <- 0
  for(i in 1:length(y)){
    prop_density =  prop_density +
      log(
        (pnorm((y[i]+0.5-mu)/sqrt(sigma2),mean = 0,sd = 1,)  -
         pnorm((y[i]-0.5-mu)/sqrt(sigma2),mean = 0,sd = 1))
         )
  }
  return(exp(prop_density))
}




#### MCMC ----------------------------------------------------
## Metropolis Hastings ------------------------------------------
```

```r
# candidate will be the posterior without truncation

n_it <- 500 #number of samples to be generated
step <- 10
burn_in = 500
total_samples = burn_in + n_it*step
init_value <- generate_posterior_untruncated(y=y) # initial value of theta
accepted=0
post_samp_truncated <-
  data.frame(mu=rep(NA,n_it),
             sigma2=rep(NA,n_it))

## Let's use as candidate the posterior
## pretending unrounded measurements
post_samp_truncated[1,]<-init_value
for (i in 2:total_samples){
  candidate <- generate_posterior_untruncated(y=y)

  ## Calculating r
  #numerators
  num1 = prop_density_post_truncated(
    y = y,
    sigma2 = candidate$sigma2,
    mu = candidate$mu)

  num2 = density_posterior_untruncated(
    y = y,
    mu=candidate$mu,
    sigma2 = candidate$sigma2
  )

  #denominators
  den1 = prop_density_post_truncated(
    y = y,
    sigma2 = post_samp_truncated[i-1,]$sigma2,
    mu = post_samp_truncated[i-1,]$mu)

  den2 = density_posterior_untruncated(
    y = y,
    mu=post_samp_truncated[i-1,]$mu,
    sigma2 = post_samp_truncated[i-1,]$sigma2
  )

  ## r
  ## using the log since the densities can be very small
  r=min(1,exp(log(num1)+log(num2)-(log(den1)+log(den2))))
  #r=min(1,exp(log(num1)-log(den1)))

  if(runif(1)>r){
    # rejected! we will start the new iteration with the same point
    post_samp_truncated[i,] <- post_samp_truncated[i-1,]
  } else {
    # accepted! we will move our chain to the candidate point and start the
```

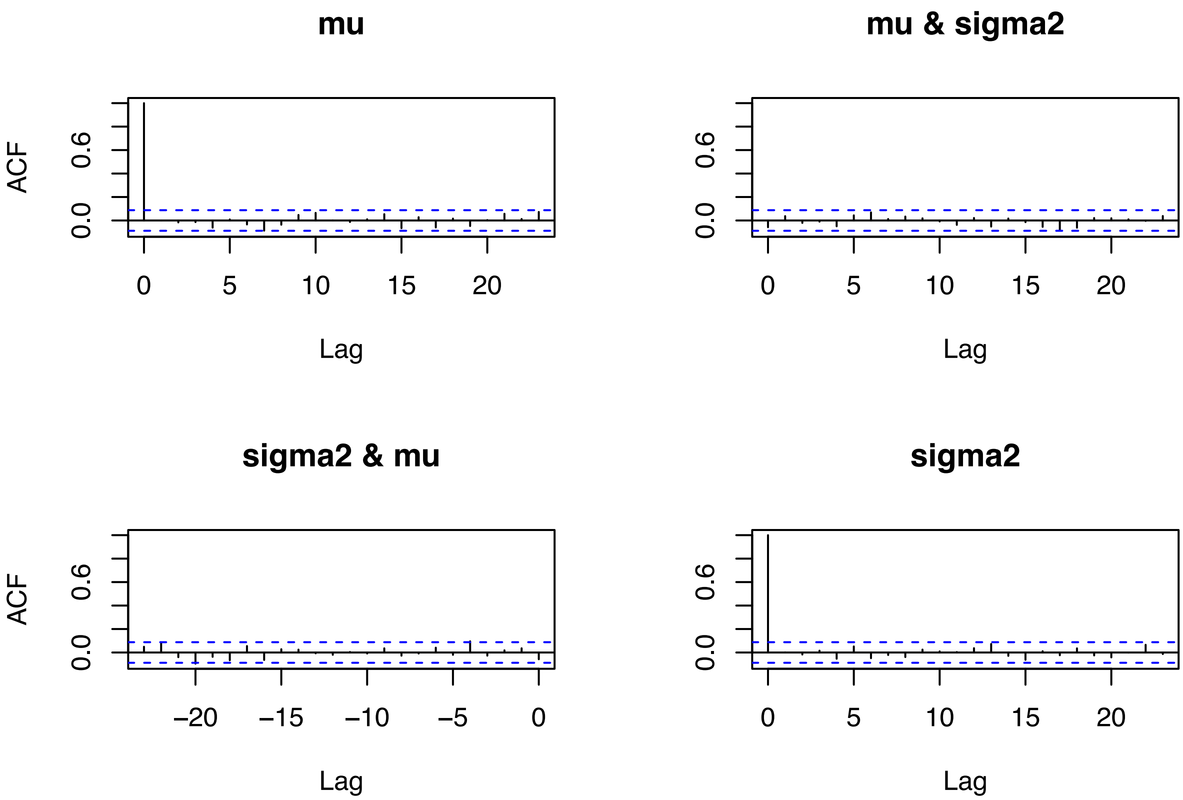
```
    # next iteration from this point
    post_samp_truncated[i,] <- candidate
    accepted<-accepted+1
  }
}

## acceptance rate =
accepted/total_samples
```

```
## [1] 0.3885455
```

```
## brushing the sample
post_sample_clean <- post_samp_truncated[-c(1:burn_in),][seq(1,total_samples-burn_in,by=step),]

par(mfrow=c(2,1))
plot(post_sample_clean$mu,type='l',ylab='mu')
plot(post_sample_clean$sigma2,type='l',ylab='sigma2')
```




```
par(mfrow=c(1,1))
acf(post_sample_clean)
```



c)

```
sample_posterior_untruncated <- data.frame(
  mu=rep(NA,n_it),
  sigma2=rep(NA,n_it)
)

for(i in 1:n_it){
  sample_posterior_untruncated[i,] <- generate_posterior_untruncated(y=y)
}

posterior_samples <- data.frame(
  mu_untruncated = sample_posterior_untruncated$mu,
  mu_truncated = post_sample_clean$mu,
  sigma2_untruncated = sample_posterior_untruncated$sigma2,
  sigma2_truncated = post_sample_clean$sigma2
)

data.frame(
  Mean =  posterior_samples %>%
    summarise(Untruncated = mean(mu_untruncated),
              Truncated = mean(mu_truncated)) %>% t(),
  SD =posterior_samples %>%
    summarise(Untruncated = sd(mu_untruncated),
              Truncated = sd(mu_truncated)) %>% t(),
  Percentile_0.5 =  posterior_samples %>%
    summarise(Untruncated = quantile(mu_untruncated,0.05),
              Truncated = quantile(mu_truncated,0.1)) %>% t(),
```

Table 1: For mu

	Mean	SD	Percentile_0.5	Percentile_1	Percentile_25	Percentile_50
Untruncated	10.39151	0.7224377	9.28141	9.598535	9.984416	10.41365
Truncated	10.39554	0.2725610	10.04506	9.939021	10.218478	10.39331

Table 2: For sigma

	Mean	SD	Percentile_0.5	Percentile_1	Percentile_25	Percentile_50
Untruncated	2.549833	3.4040734	0.4983951	0.6439316	0.9929236	1.5914511
Truncated	1.085370	0.4549925	0.6252114	0.5610709	0.7750116	0.9722241

```
  Percentile_1 = posterior_samples %>%
    summarise(Untruncated = quantile(mu_untruncated,0.1),
              Truncated = quantile(mu_truncated,0.05)) %>% t(),
  Percentile_25 =   posterior_samples %>%
    summarise(Untruncated = quantile(mu_untruncated,0.25),
              Truncated = quantile(mu_truncated,0.25)) %>% t(),
  Percentile_50 = posterior_samples %>%
    summarise(Untruncated = quantile(mu_untruncated,0.5),
              Truncated = quantile(mu_truncated,0.5)) %>% t()
) %>% knitr::kable(caption = 'For mu') %>% kableExtra::kable_styling()
```

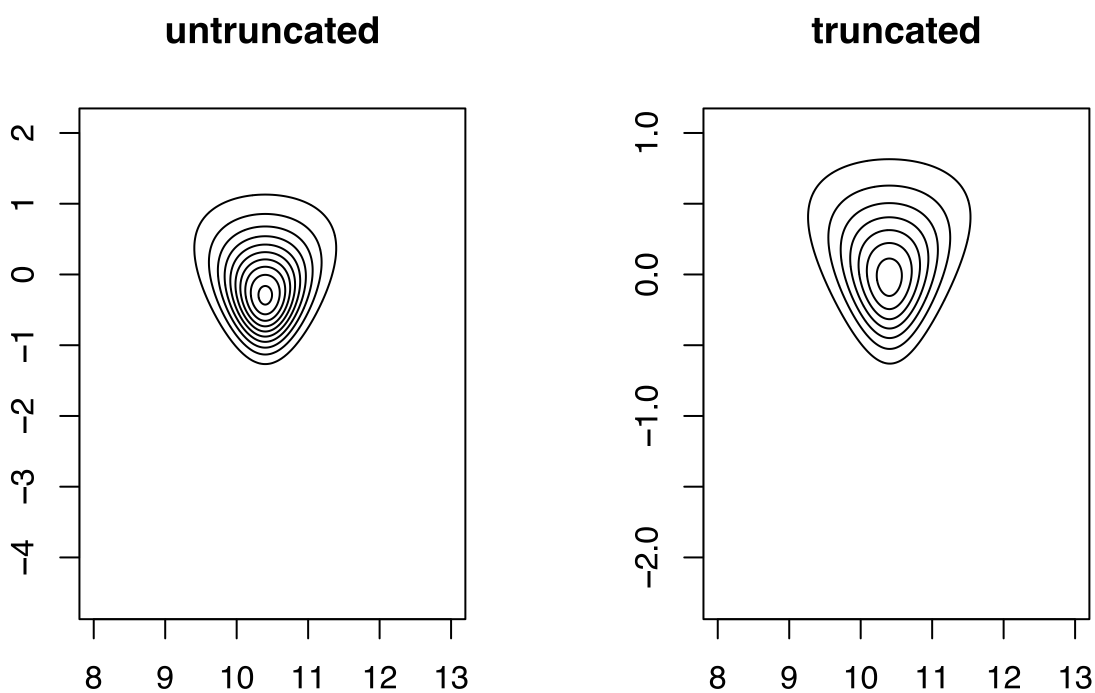
```
data.frame(
  Mean =  posterior_samples %>%
    summarise(Untruncated = mean(sigma2_untruncated),
              Truncated = mean(sigma2_truncated)) %>% t(),
  SD =posterior_samples %>%
    summarise(Untruncated = sd(sigma2_untruncated),
              Truncated = sd(sigma2_truncated)) %>% t(),
  Percentile_0.5 =  posterior_samples %>%
    summarise(Untruncated = quantile(sigma2_untruncated,0.05),
              Truncated = quantile(sigma2_truncated,0.1)) %>% t(),
  Percentile_1 = posterior_samples %>%
    summarise(Untruncated = quantile(sigma2_untruncated,0.1),
              Truncated = quantile(sigma2_truncated,0.05)) %>% t(),
  Percentile_25 =   posterior_samples %>%
    summarise(Untruncated = quantile(sigma2_untruncated,0.25),
              Truncated = quantile(sigma2_truncated,0.25)) %>% t(),
  Percentile_50 = posterior_samples %>%
    summarise(Untruncated = quantile(sigma2_untruncated,0.5),
              Truncated = quantile(sigma2_truncated,0.5)) %>% t()
) %>% knitr::kable(caption = 'For sigma') %>% kableExtra::kable_styling()
```

Now,for the contour plot we are going to use the $log(\sigma^2)$ in order to have a better visualization.

```
#density untruncated
mu_grid <- seq(8,13,0.01)
sigma2_grid <- seq(0.01,8,0.01)
density_values <- matrix(NA,nrow=length(mu_grid),ncol=length(sigma2_grid))

for(i in 1:length(mu_grid)){
  for(j in 1:length(sigma2_grid)){
```

```
  #cat(paste(i,j),'\n')
    density_values[i,j] <- density_posterior_untruncated(
      y = y,
      mu = mu_grid[i],
      sigma2 = sigma2_grid[j]
    )
  }
}
```

```
#density untruncated
density_values_truncated <- matrix(NA,nrow=length(mu_grid),ncol=length(sigma2_grid))

for(i in 1:length(mu_grid)){
  for(j in 1:length(sigma2_grid)){
  #cat(paste(i,j),'\n')
    density_values_truncated[i,j] <- prop_density_post_truncated(
      y = y,
      mu = mu_grid[i],
      sigma2 = sigma2_grid[j]
    )
  }
}

#this one is not normalized! but since we would divide by a constant we can still use the shape to see
par(mfrow=c(1,2))
contour(
  x = mu_grid,
  y = log(sigma2_grid),
  z = density_values,
  main='untruncated',
  drawlabels = F
  )

contour(
  x = mu_grid,
  y = log(sqrt(sigma2_grid)),
  z = density_values_truncated,
  drawlabels = F,main='truncated',
  )
```




```
par(mfrow=c(1,1))
```

d)

For each i, z_i is the the unrounded value of y_i. Therefore, we can generate a sample of z_i conditional on y_i by generating from a truncated normal distribution between $y_i - 5$ and $y_i + 5$ with parameters from our posterior sample.

```
z_samples <- matrix(NA, n_it, 5)
for (i in 1:5) {
  z_samples[, i] <- truncnorm::rtruncnorm(
    n_it, a = y[i] - 0.5, b = y[i] + 0.5,
    mean = sample_posterior_untruncated$mu, sd = sample_posterior_untruncated$sigma2)
}

((z_samples[, 2] - z_samples[, 1])^2) %>% mean()
```

```
## [1] 0.1491312
```

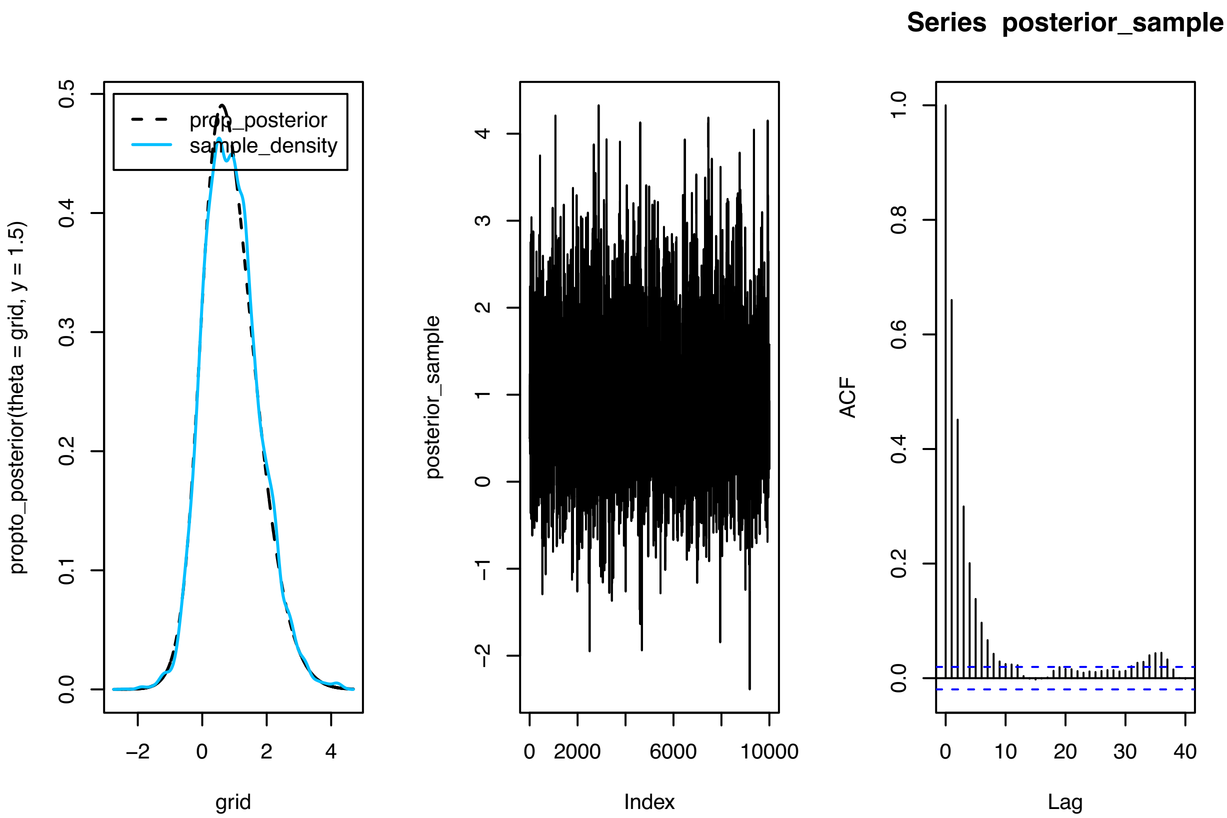
Ex 10

```
# proportional to the posterior distribution that we
# are going to draw samples from
# Witout using log
propto_posterior <- function(y=1.5,theta){
  numerator <- exp(-(1/2)*(y-theta)^2)
  denominator <- (1+theta^2)
  return(numerator/denominator)
}

# Using log
propto_posterior <- function(y=1.5,theta){
```

```r
 numerator <- -(1/2)*(y-theta)^2
 denominator <- -log(1+theta^2)
 return(exp(numerator+denominator))
}

randWalk_MH_diagnostics <- function(posterior_sample){
  lindseys_density  <- density(posterior_sample)
  a <- round(min(lindseys_density$x),3) ; b <-round(max(lindseys_density$x),3)
  grid<- seq(a,b,0.001)
  par(mfrow=c(1,3))
  plot(grid,propto_posterior(theta = grid,y=1.5),type='l',lty=2,lwd=1.5)
  lines(density(posterior_sample),type='l',col='deepskyblue',lwd=1.5)
  legend(a,0.5,legend = c('prop_posterior','sample_density'),
         col=c('black','deepskyblue'),lwd = c(1.5,1.5), lty=c(2,1))
  plot(posterior_sample,type = 'l')
  acf(posterior_sample)
  par(mfrow=c(1,3))
}


n_it <- 10000 #number of samples to be generated
init_value <- 0.5 # initial value of theta
proposal_dist <- function(theta){rnorm(1,theta,sd=2)} # we are going to use the normal distribution as p
post_sample <- rep(NA,n_it) #creating the vector that will store the samples
post_sample[1]=init_value #the first draw of the posterior is the initial value
accepted=0 # counting the accepted values to see the acceptance rate
# initializing the iteration
for (i in 2:n_it){
  candidate <- proposal_dist(post_sample[i-1])

  #calculating acceptance probability of the current point
  r = min(1,propto_posterior(theta = candidate) /
            propto_posterior(theta = post_sample[i-1]))

  #accepting the candidates with probability r
  if(runif(1)>r){
    # if not accepted, we will start the new iteration with the same point
    post_sample[i] <- post_sample[i-1]
  } else {
    # if accepted, we will move our chain to the candidate point and start the
    # next iteration from this point
    post_sample[i] <- candidate
    accepted<-accepted+1
  }
}

accepted/n_it
```

```
## [1] 0.4294
```

```r
randWalk_MH_diagnostics(post_sample)
```




```
summary(post_sample)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -2.3869  0.3080  0.8580  0.9182  1.4414  4.3269
```

```
n_it <- 1000 #number of samples to be generated
init_value <- 0 # initial value of theta
step <- 5
burn_in = 10000
total_samples = burn_in + n_it*step
# we are going to use the normal distribution as proposal
proposal_dist <- function(theta){rnorm(1,theta,sd=2)}
#creating the vector that will store the samples
post_sample <- rep(NA,total_samples)
#the first draw of the posterior is the initial value
post_sample[1]=init_value
accepted=0 # counting the accepted values to see the acceptance rate

# initializing the iteration
for (i in 2:total_samples){
  candidate <- proposal_dist(post_sample[i-1])

  #calculating acceptance probability of the current point
  r = min(1,propto_posterior(theta = candidate) /
            propto_posterior(theta = post_sample[i-1]))

  #accepting the candidates with probability r
  if(runif(1)>r){
    # if not accepted,
```

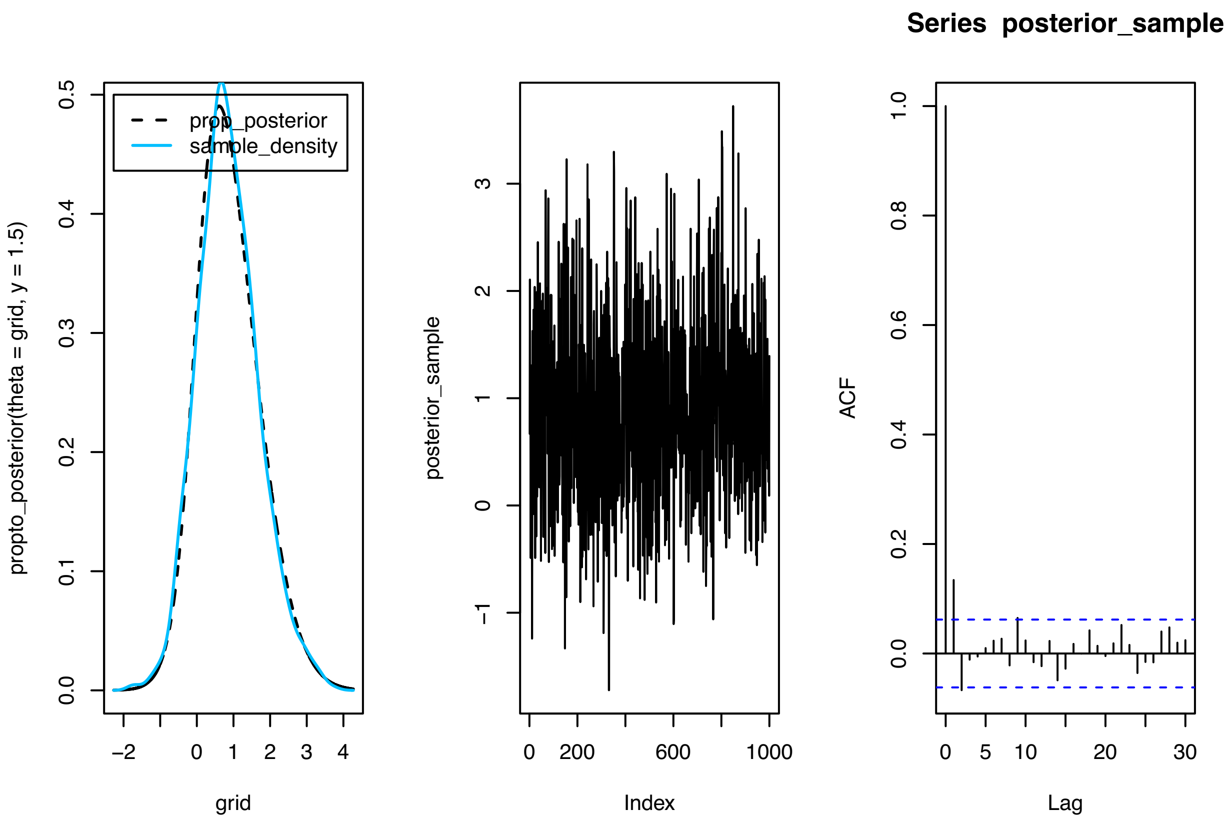
```r
    # we will start the new iteration with the same point
    post_sample[i] <- post_sample[i-1]
  } else {
    # if accepted,
    # we will move our chain to the candidate point and start the
    # next iteration from this point
    post_sample[i] <- candidate
    accepted<-accepted+1
  }
}

#acceptance rate
accepted/total_samples
```

```
## [1] 0.4358
```

```r
post_sample_clean <- post_sample[-c(1:burn_in)][seq(1,length(post_sample)-burn_in,step)]

#verifying if the length of the draw matches with the desired number
post_sample_clean %>% length()
```

```
## [1] 1000
```

```r
randWalk_MH_diagnostics(post_sample_clean)
```




```r
post_sample_clean %>% summary()
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -1.7224  0.3063  0.8010  0.8652  1.3884  3.7239
```

```r
post_sample_clean %>% sd()
```

```
## [1] 0.8282481
```