Worksheet 03 Group 2

Andrea Staub Emanuel Mauch Guillaume Morlet Holly Vuarnoz
Jan Hohenheim Sophie Haldemann

```
library(tidyverse)
#library(rjags)
library(coda)
library(bayesmeta)
library(pCalibrate)
```

```
## Warning in .recacheSubclasses(def@className, def, env): undefined subclass
## "ndiMatrix" of class "replValueSp"; definition not updated
```

```
library(glue)
library(ggplot2)
```

Exercise 3:

a)

$y_1, \dots, y_n \sim \mathcal{N}(m, k^{-1}) \qquad m \sim \mathcal{N}(\mu, \lambda^{-1})$

Prior: $f(m) = \left(\frac{\lambda}{2\pi}\right)^{\frac{1}{2}} \exp\left(-\frac{\lambda}{2}(m-\mu)^2\right)$

Likelihood: $f(y \mid m) = \left(\frac{k}{2\pi}\right)^{\frac{1}{2}} \exp\left(-\frac{k}{2}(y-m)^2\right)$

$$f(y) = \overbrace{\int_{\mathbb{R}} \underbrace{f(y \mid m)}_{\text{likelihood}} \underbrace{f(m)}_{\text{prior}} \, dm}^{\text{marginal likelihood}}$$

$$= \int_{\mathbb{R}} \left(\frac{k}{2\pi}\right)^{\frac{1}{2}} \exp\left(-\frac{k}{2}(y-m)^2\right) \left(\frac{\lambda}{2\pi}\right)^{\frac{1}{2}} \exp\left(-\frac{\lambda}{2}(m-\mu)^2\right) dm$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \int_{\mathbb{R}} \exp\left(-\frac{k}{2}(y-m)^2 - \frac{\lambda}{2}(m-\mu)^2\right) dm$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}\left(k(y-m)^2 + \lambda(m-\mu)^2\right)\right) dm$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}\left(k(y^2+m^2-2ym) + \lambda(m^2+\mu^2-2m\mu)\right)\right) dm$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}\left(ky^2 + km^2 - 2kym + \lambda m^2 + \lambda\mu^2 - 2\lambda m\mu\right)\right) dm$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}\left(ky^2 + \lambda\mu^2 + km^2 + \lambda m^2 - 2m(ky + \lambda\mu)\right)\right) dm$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}\left(ky^2 + \lambda\mu^2 + km^2 + \lambda m^2 - 2m(ky + \lambda\mu)\right)\right) dm$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \int_{\mathbb{R}} \exp\left\{-\frac{1}{2}\left[(\lambda+k)\left(\frac{km^2}{\lambda+k} + \frac{\lambda m^2}{\lambda+k} - 2m\frac{ky+\lambda\mu}{\lambda+k}\right) + ky^2 + \lambda\mu^2\right]\right\} dm$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \int_{\mathbb{R}} \exp\left\{-\frac{1}{2}\left[(\lambda+k)\left(m^2 - 2m\frac{ky+\lambda\mu}{\lambda+k}\right) + ky^2 + \lambda\mu^2\right]\right\} dm$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \int_{\mathbb{R}} \exp\left\{-\frac{1}{2}\left[(\lambda+k)\left(m^2 - 2m\frac{ky+\lambda\mu}{\lambda+k}\right) + ky^2 + \lambda\mu^2\right]\right\} dm$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \int_{\mathbb{R}} \exp\left\{-\frac{1}{2}\left[(\lambda+k)\left(m^2 - 2m\frac{ky+\lambda\mu}{\lambda+k} + \left(\frac{ky+\lambda\mu}{\lambda+k}\right)^2\right) - \frac{(ky+\lambda\mu)^2}{\lambda+k} + ky^2 + \lambda\mu^2\right]\right\} dm$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \int_{\mathbb{R}} \exp\left\{-\frac{1}{2}\left[(\lambda+k)\left(m - \frac{ky+\lambda\mu}{\lambda+k}\right)^2 - \frac{(ky+\lambda\mu)^2}{\lambda+k} + ky^2 + \lambda\mu^2\right]\right\} dm$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \int_{\mathbb{R}} \exp\left\{-\frac{1}{2}\left[(\lambda+k)\left(m - \frac{ky+\lambda\mu}{\lambda+k}\right)^2\right]\right\} \exp\left\{\frac{1}{2}\left[-\frac{(ky+\lambda\mu)^2}{\lambda+k} + ky^2 + \lambda\mu^2\right]\right\} dm$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \exp\left\{-\frac{1}{2}\left[-\frac{(ky+\lambda\mu)^2}{\lambda+k} + ky^2 + \lambda\mu^2\right]\right\} \int_{\mathbb{R}} \exp\left\{-\frac{1}{2}\left[(\lambda+k)\left(m - \frac{ky+\lambda\mu}{\lambda+k}\right)^2\right]\right\} dm$$

precision, mean

$$= \frac{\sqrt{k\lambda}}{2\pi} \exp\left\{-\frac{1}{2}\left[-\frac{(ky+\lambda\mu)^2}{\lambda+k} + ky^2 + \lambda\mu^2\right]\right\} \int_{\mathbb{R}} \exp\left\{-\frac{\lambda+k}{2}\left(m - \frac{ky+\lambda\mu}{\lambda+k}\right)^2\right\} dm$$

in general, for $x \sim \mathcal{N}(\mu, \frac{1}{k})$: $f(x) = \sqrt{\frac{k}{2\pi}} \exp\left\{-\frac{k}{2}(x-\mu)^2\right\}$

$$= \frac{\sqrt{k\lambda}}{2\pi} \exp\left\{-\frac{1}{2}\left[-\frac{(ky+\lambda\mu)^2}{\lambda+k} + ky^2 + \lambda\mu^2\right]\right\} \sqrt{\frac{2\pi}{\lambda+k}} \underbrace{\int_{\mathbb{R}} \sqrt{\frac{\lambda+k}{2\pi}} \exp\left\{-\frac{\lambda+k}{2}\left(m - \frac{ky+\lambda\mu}{\lambda+k}\right)^2\right\} dm}_{=1}$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \exp\left\{-\frac{1}{2}\left[-\frac{(ky+\lambda\mu)^2}{\lambda+k} + ky^2 + \lambda\mu^2\right]\right\} \sqrt{\frac{2\pi}{\lambda+k}}$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \exp\left\{\frac{(ky+\lambda\mu)^2}{2\lambda+k} - \frac{ky^2}{2} - \frac{\lambda\mu^2}{2}\right\} \sqrt{\frac{2\pi}{\lambda+k}}$$

Because this is an integral over a normal distribution's PDF with $m \sim \mathcal{N}\left(\frac{ky+\lambda\mu}{\lambda+k}, \lambda+k\right)$ we know that

$$= \frac{\sqrt{k\lambda}}{2\pi} \exp\left\{ \frac{1}{2(\lambda+k)} \left[2\lambda\mu y k - \lambda k y^2 - k\lambda\mu^2 \right] \right\} \sqrt{\frac{2\pi}{\lambda+k}}$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \exp\left\{ -\frac{1}{2(\lambda+k)} \left[-2\lambda\mu y k + \lambda k y^2 + k\lambda\mu^2 \right] \right\} \sqrt{\frac{2\pi}{\lambda+k}}$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \exp\left\{ -\frac{k\lambda}{2(\lambda+k)} \left[-2\mu y + y^2 + \mu^2 \right] \right\} \sqrt{\frac{2\pi}{\lambda+k}}$$

$$= \frac{\sqrt{k\lambda}}{2\pi} \exp\left\{ -\frac{k\lambda}{2(\lambda+k)} (y-\mu)^2 \right\} \sqrt{\frac{2\pi}{\lambda+k}}$$

$$= \sqrt{\frac{k\lambda}{4\pi^2}} \sqrt{\frac{2\pi}{\lambda+k}} \exp\left\{ -\frac{k\lambda}{2(\lambda+k)} (y-\mu)^2 \right\}$$

$$= \underbrace{\sqrt{\frac{k\lambda}{2\pi(\lambda+k)}} \exp\left\{ -\frac{k\lambda}{2(\lambda+k)} (y-\mu)^2 \right\}}_{\text{This is the form of a normal distribution's PDF!}}$$

precision: $\frac{k\lambda}{k+\lambda} = \frac{1}{k} + \frac{1}{\lambda} = k^{-1} + \lambda^{-1}$

$\Rightarrow$ var $= \frac{k+\lambda}{k\lambda}$

mean: μ

$$\Rightarrow y \sim \mathcal{N}\left(\mu, \frac{k+\lambda}{k\lambda}\right)$$

b)

$$f(y_{n+1} \mid y_{1:n}) = \int_{\mathbb{R}} f(y_{n+1}, m \mid y_{1:n})\, dm$$

$$= \int_{\mathbb{R}} f(y_{n+1} \mid m, y_{1:n})\, f(m \mid y_{1:n})\, dm = \int_{\mathbb{R}} f(y_{n+1} \mid m)\, f(m \mid y_{1:n})\, dm$$

conditionally independent

Likelihood: $y_{n+1} \mid m \sim \mathcal{N}(m, \kappa^{-1}) \Rightarrow f(y_{n+1} \mid m) = \sqrt{\frac{\kappa}{2\pi}} \exp\left\{ \frac{\kappa}{2} (y_{n+1} - m)^2 \right\}$

Posterior: $m \mid y_{1:n} \sim \mathcal{N}\left(\frac{\kappa n \bar{y} + \lambda\mu}{n\kappa + \lambda}, (n\kappa + \lambda)^{-1} \right)$ per last exercise sheet

$$\mu' := \frac{\kappa n \bar{y} + \lambda \mu}{n\kappa + \lambda} \quad \text{read "}\mu\text{ posterior"}$$

$$\kappa' := n\kappa + \lambda \quad \text{read "}\kappa\text{ posterior"}$$

$$\Rightarrow f(m \mid y_{1:n}) = \sqrt{\frac{\kappa'}{2\pi}} \exp\left\{ -\frac{\kappa'}{2} (m - \mu')^2 \right\}$$

hence

$$f(y_{n+1} \mid y_{1:n}) = \int_{\mathbb{R}} \sqrt{\frac{\kappa}{2\pi}} \exp\left\{ -\frac{\kappa}{2} (y_{n+1} - m)^2 \right\} \sqrt{\frac{\kappa'}{2\pi}} \exp\left\{ -\frac{\kappa'}{2} (m - \mu')^2 \right\} dm$$

$$= \frac{\sqrt{\kappa\kappa'}}{2\pi} \int_{\mathbb{R}} \exp\left\{ -\frac{1}{2} \left[\kappa (y_{n+1} - m)^2 + \kappa' (m - \mu')^2 \right] \right\} dm$$

$$= \frac{\sqrt{\kappa\kappa'}}{2\pi} \int_{\mathbb{R}} \exp\left\{ -\frac{1}{2} \left[\kappa (y_{n+1} - m)^2 + \kappa' (m - \mu')^2 \right] \right\} dm$$

$$= \frac{\sqrt{\kappa\kappa'}}{2\pi} \int_{\mathbb{R}} \exp\left\{ -\frac{1}{2} \left[\kappa y_{n+1}^2 - 2\kappa m y_{n+1} + \kappa m^2 + \kappa' m^2 - 2\kappa' m \mu' + \kappa' \mu'^2 \right] \right\} dm$$

$$= \frac{\sqrt{\kappa\kappa'}}{2\pi} \int_{\mathbb{R}} \exp\left\{ -\frac{1}{2} \left[(\kappa + \kappa') \left(m^2 - 2m \frac{\kappa y_{n+1} + \kappa'\mu'}{\kappa + \kappa'} \right) + \kappa y_{n+1}^2 + \kappa' \mu'^2 \right] \right\} dm$$

$$= \frac{\sqrt{\kappa\kappa'}}{2\pi} \int_{\mathbb{R}} \exp\left\{ -\frac{1}{2} \left[(\kappa + \kappa') \left(m^2 - 2m \frac{\kappa y_{n+1} + \kappa'\mu'}{\kappa + \kappa'} + \left(\frac{\kappa y_{n+1} + \kappa'\mu'}{\kappa + \kappa'} \right)^2 \right) - \frac{(\kappa y_{n+1} + \kappa'\mu')^2}{\kappa + \kappa'} + \kappa y_{n+1}^2 + \kappa' \mu'^2 \right] \right\} dm$$

completing the square

$$= \frac{\sqrt{\kappa\kappa'}}{2\pi} \int_{\mathbb{R}} \exp\left\{ -\frac{1}{2} \left[(\kappa + \kappa') \left(m^2 - \frac{\kappa y_{n+1} + \kappa'\mu'}{\kappa + \kappa'} \right)^2 - \frac{(\kappa y_{n+1} + \kappa'\mu')^2}{\kappa + \kappa'} + \kappa y_{n+1}^2 + \kappa' \mu'^2 \right] \right\} dm$$

$$= \frac{\sqrt{\kappa\kappa'}}{2\pi} \int_{\mathbb{R}} \exp\left\{ -\frac{1}{2} \left[(\kappa + \kappa') \left(m^2 - \frac{\kappa y_{n+1} + \kappa'\mu'}{\kappa + \kappa'} \right)^2 \right] \right\} \exp\left\{ -\frac{1}{2} \left[- \frac{(\kappa y_{n+1} + \kappa'\mu')^2}{\kappa + \kappa'} + \kappa y_{n+1}^2 + \kappa' \mu'^2 \right] \right\} dm$$

$$= \frac{\sqrt{\kappa\kappa'}}{2\pi} \exp\left\{ -\frac{1}{2} \left[- \frac{(\kappa y_{n+1} + \kappa'\mu')^2}{\kappa + \kappa'} + \kappa y_{n+1}^2 + \kappa' \mu'^2 \right] \right\} \int_{\mathbb{R}} \exp\left\{ -\frac{1}{2} \left[(\kappa + \kappa') \left(m^2 - \frac{\kappa y_{n+1} + \kappa'\mu'}{\kappa + \kappa'} \right)^2 \right] \right\} dm$$

$$= \frac{\sqrt{\kappa\kappa'}}{2\pi} \exp\left\{ -\frac{1}{2} \left[- \frac{(\kappa y_{n+1} + \kappa'\mu')^2}{\kappa + \kappa'} + \kappa y_{n+1}^2 + \kappa' \mu'^2 \right] \right\} \int_{\mathbb{R}} \exp\left\{ -\underbrace{\frac{\kappa + \kappa'}{2}}_{\text{precision}} \left(m^2 - \underbrace{\frac{\kappa y_{n+1} + \kappa'\mu'}{\kappa + \kappa'}}_{\text{mean}} \right)^2 \right\} dm$$

this looks like a normal distribution's PDF, so let's add a normalizing constant

$$= \frac{\sqrt{\kappa\kappa'}}{2\pi} \exp\left\{ -\frac{1}{2} \left[- \frac{(\kappa y_{n+1} + \kappa'\mu')^2}{\kappa + \kappa'} + \kappa y_{n+1}^2 + \kappa' \mu'^2 \right] \right\} \sqrt{\frac{2\pi}{\kappa + \kappa'}} \underbrace{\int_{\mathbb{R}} \sqrt{\frac{\kappa + \kappa'}{2\pi}} \exp\left\{ -\frac{\kappa + \kappa'}{2} \left(m^2 - \frac{\kappa y_{n+1} + \kappa'\mu'}{\kappa + \kappa'} \right)^2 \right\} dm}_{=1}$$

$$= \frac{\sqrt{\kappa\kappa'}}{2\pi}\sqrt{\frac{2\pi}{\kappa+\kappa'}}\exp\left\{-\frac{1}{2}\left[-\frac{(\kappa y_{n+1}+\kappa'\mu')^2}{\kappa+\kappa'}+\kappa y_{n+1}^2+\kappa'\mu'^2\right]\right\}$$

$$= \sqrt{\frac{\kappa\kappa'}{2\pi(\kappa+\kappa')}}\exp\left\{-\frac{1}{2}\left[-\frac{(\kappa y_{n+1}+\kappa'\mu')^2}{\kappa+\kappa'}+\kappa y_{n+1}^2+\kappa'\mu'^2\right]\right\}$$

$$= \sqrt{\frac{\kappa\kappa'}{2\pi(\kappa+\kappa')}}\exp\left\{\frac{1}{2}\left[\frac{(\kappa y_{n+1}+\kappa'\mu')^2}{\kappa+\kappa'}-\kappa y_{n+1}^2-\kappa'\mu'^2\right]\right\}$$

$$= \sqrt{\frac{\kappa\kappa'}{2\pi(\kappa+\kappa')}}\exp\left\{\frac{1}{2}\left[\frac{(\kappa y_{n+1}+\kappa'\mu')^2}{\kappa+\kappa'}-\kappa y_{n+1}^2-\kappa'\mu'^2\right]\right\}$$

$$= \sqrt{\frac{\kappa\kappa'}{2\pi(\kappa+\kappa')}}\exp\left\{\frac{1}{2(\kappa+\kappa')}\left[\kappa^2 y_{n+1}^2+2\kappa y_{n+1}\kappa'\mu'+\kappa'^2\mu'^2-\kappa^2 y_{n+1}^2-\kappa\kappa' y_{n+1}^2-\kappa\kappa'\mu'^2-\kappa'^2\mu'^2\right]\right\}$$

$$= \sqrt{\frac{\kappa\kappa'}{2\pi(\kappa+\kappa')}}\exp\left\{\frac{1}{2(\kappa+\kappa')}\left[2\kappa y_{n+1}\kappa'\mu'-\kappa\kappa' y_{n+1}^2-\kappa\kappa'\mu'^2\right]\right\}$$

$$= \sqrt{\frac{\kappa\kappa'}{2\pi(\kappa+\kappa')}}\exp\left\{-\frac{1}{2(\kappa+\kappa')}\left[-2\kappa y_{n+1}\kappa'\mu'+\kappa\kappa' y_{n+1}^2+\kappa\kappa'\mu'^2\right]\right\}$$

$$= \sqrt{\frac{\kappa\kappa'}{2\pi(\kappa+\kappa')}}\exp\left\{-\frac{\kappa\kappa'}{2(\kappa+\kappa')}\left[-2y_{n+1}\mu'+y_{n+1}^2+\mu'^2\right]\right\}$$

$$= \sqrt{\frac{\kappa\kappa'}{2\pi(\kappa+\kappa')}}\exp\left\{-\underbrace{\frac{\kappa\kappa'}{2(\kappa+\kappa')}}_{\text{precision}}(y_{n+1}-\underbrace{\mu'}_{\text{mean}})^2\right\}$$

$$\text{Variance} = \left(\frac{\kappa\kappa'}{\kappa+\kappa'}\right)^{-1} = \frac{\kappa+\kappa'}{\kappa\kappa'} = \kappa^{-1}+\kappa'^{-1}$$

$$\Rightarrow y_{n+1} \mid y_{1:n} \sim \mathcal{N}(\mu', \kappa^{-1}+\kappa'^{-1})$$

where

$$\boxed{\begin{aligned}\mu' &:= \frac{\kappa n\bar{y}+\lambda\mu}{n\kappa+\lambda}\\ \kappa' &:= n\kappa+\lambda\end{aligned}}$$

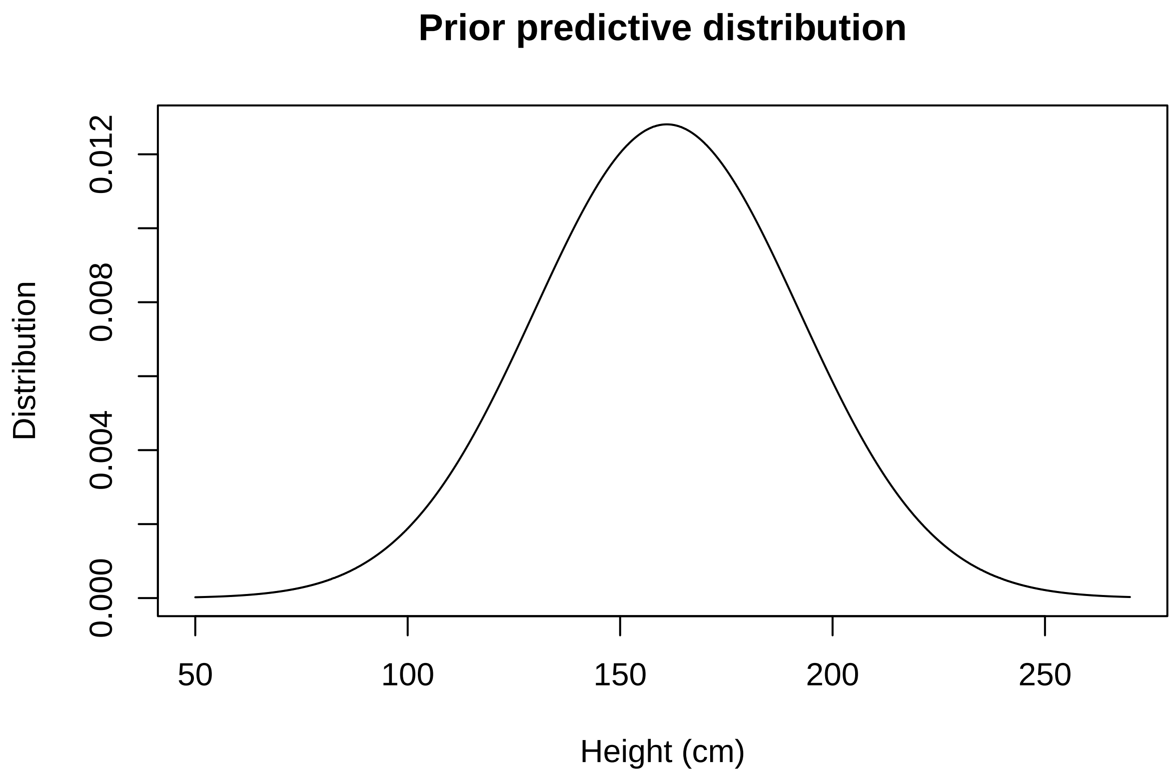
Exercise 4:

Apply analytical formulas derived in Exercise 3 above to the vector of height (cm) measurements 166, 168, 168, 177, 160, 170, 172, 159, 175, 164, 175, 167, 164 of 13 Swiss females. Assume that $y_1, ..., y_n$ are observations generated by N(m, κ^{-1}) distribution with $\kappa = 1/900$. Moreover, assume a N(μ, λ^{-1}) prior for m with $\mu = 161$ and $\lambda = 1/70$.

a) Plot the prior predictive distribution for one observation y and compute its expectation and standard deviation. Estimate P[y > 200] for one future observation of Height.

The prior predictive distribution of one future observation y is N(μ, λ^{-1}+κ^{-1}).

```
my_mu <- 161
my_lambda <- 1/70
my_kappa <- 1/900
my_seq <- seq(50, 270, by = 0.01)
plot(my_seq, dnorm(my_seq, mean = my_mu, sd = sqrt(1/my_lambda+1/my_kappa)), type = "l", main = "Prior
```



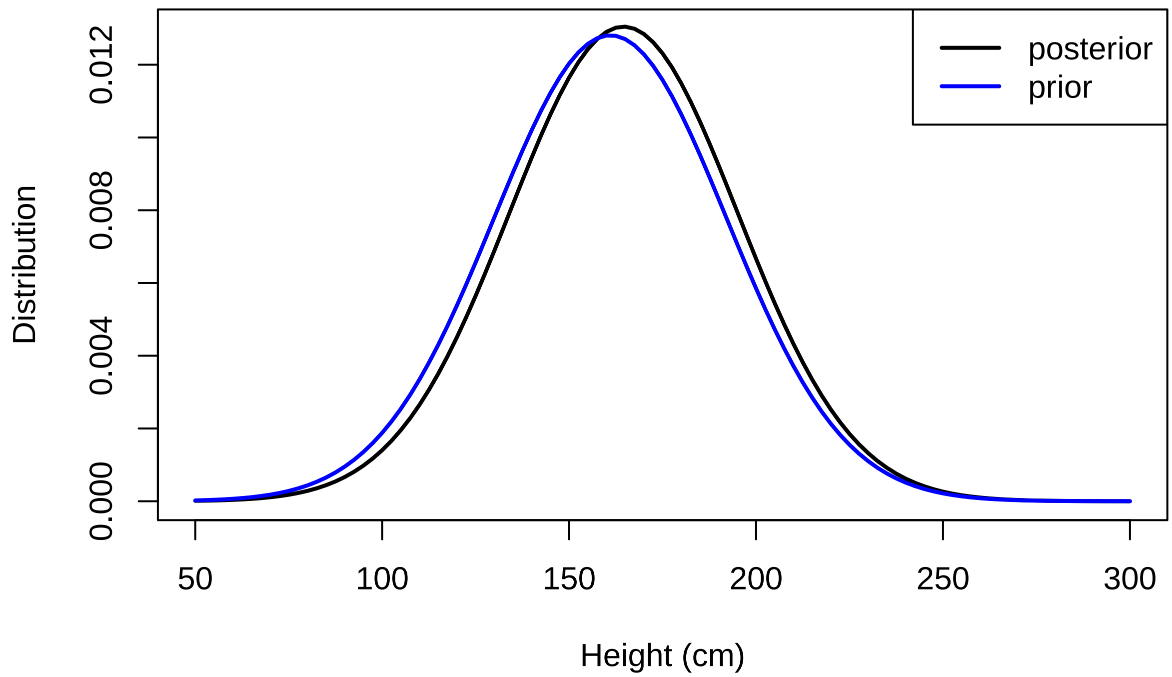

```
expect <- my_mu
stand_dev <- sqrt(1/my_lambda + 1/my_kappa)
prob <- 1-pnorm(200, mean = my_mu, sd = sqrt(1/my_lambda + 1/my_kappa))
```

The expectation is 161 and the standard deviation is 31.145.

P[y >200] for one future observation of Height is 0.105.

b) Plot the posterior predictive distribution for one future observation y_{n+1} given that $y_1, ..., y_n$ have been observed and compute its expectation and standard deviation. Estimate P$[y_{n+1} > 200 y_1, ..., y_n]$ for one future observation y_{n+1} of Height.

```
obs <- c(166, 168, 168, 177, 160, 170, 172, 159, 175, 164, 175, 167, 164)
n <- length(obs)
mean_obs <- mean(obs)

mu_post <- (my_kappa*n*mean_obs+my_lambda*my_mu)/(n*my_kappa+my_lambda)
lambda_post <- n*my_kappa+my_lambda

curve(dnorm(x, mean = mu_post, sd = sqrt(1/lambda_post+1/my_kappa)), xlab = "Height (cm)", ylab = "Dist
curve(dnorm(x, mean = my_mu, sd = sqrt(1/my_lambda+1/my_kappa)), from = 50, to = 300, lwd = 2, add = TRU
legend("topright", legend = c("posterior", "prior"), col = c("black", "blue"), lwd = 2)
```




```r
prob_post <- 1-pnorm(200, mean = mu_post, sd = sqrt(1/lambda_post+1/my_kappa))
```

The expectation is 164.558 and the standard deviation is 30.575.

P[y$_{n+1}$ >200] for one future observation of Height is 0.123.

c) Compare the results obtained for predictive distribution with those obtained for the posterior in Exercise 4 of Worksheet 2. Discuss how much posterior, prior predictive, and posterior predictive distributions differ.

In Exercise 4 of Worksheet 2 the posterior distribution of m$|y_1, ..., y_n$ was derived. Whereas in this worksheet's exercise 4 the posterior distribution of $y_{n+1}|y_1, ..., y_n$ was derived. The values for the mean of the posterior predictive distribution are the same (164.558). However, the standard deviations differ with much larger values for the posterior distribution of y_{n+1}.

The mean of the prior predictive distribution is with 161 lower than the mean of the posterior predictive distribution (164.558). This shift is due to a sample mean larger than 161. The variance of the posterior predictive distribution is with 934.807 slightly smaller than the variance of the prior predictive distribution (970).

Exercise 5:

Change-of-variables formula. Derivation of Inverse Gamma (IG) and Square Root Inverse Gamma (SIG) with parameters $a = 1.6$ and $b = 0.4$.

Change of variables formula

r.v. $X \sim G(a,b)$

density $f(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}$

→ density of $Y = \frac{1}{X}$ (IG prior for σ^2)

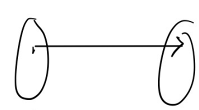
→ density of $Z = \sqrt{Y} = \sqrt{\frac{1}{X}}$ (SIG prior for σ)

Change of variable formula

$$\boxed{f_Y(y) = \underbrace{f_X(g^{-1}(y))}_{\text{inverse}} \cdot \underbrace{\left| \frac{d^{-1}(y)}{dy} \right|}_{\text{derivative of inverse}}}$$

· applied when transformation $g(\cdot)$ is one-to-one & differentiable

↳ bijective



$X \xrightarrow{g(x)} y$

$X \xleftarrow{g^{-1}(y)} y$

Figure 1: Change of variable formula

1) derive density of $Y = \frac{1}{X}$

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \left| \frac{d^{-1}(y)}{dy} \right|$$

find inverse transformation

$$g(x) = y = \frac{1}{X}$$

$$x = \frac{1}{y} \rightarrow g^{-1}(y) = \frac{1}{y}$$

derive

$$\frac{d\,g^{-1}(y)}{dy} = y^{-1}$$

$$= -y^{-2} = -\frac{1}{y^2}$$

plug in

$$f(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}$$

$$f(y) = \frac{b^a}{\Gamma(a)} \left(\frac{1}{y}\right)^{(a-1)} e^{-b/y} \left| -\frac{1}{y^2} \right|$$

$$= \frac{b^a}{\Gamma(a)} \left(y^{-1}\right)^{(a-1)} e^{-b/y} \; y^{-2}$$

$$= \frac{b^a}{\Gamma(a)} y^{-a+1-2} e^{-b/y}$$

$$= \frac{b^a}{\Gamma(a)} y^{-(a+1)} e^{-b/y}$$

→ density function of inverse gamma (IG distribution)

Figure 2: Inverse Gamma

2) derive density of $Z = \sqrt{Y} = \sqrt{\frac{1}{X}}$

inverse function:

$$g(x) = z = \sqrt{\frac{1}{x}}$$

$$z^2 = \frac{1}{x}$$

$$x = \frac{1}{z^2} \rightarrow g^{-1}(z) = \frac{1}{z^2}$$

$$\frac{d\, g^{-1}(z)}{dz} = -2\, z^{-3}$$

$$= \frac{-2}{z^3}$$

plug in

$$f(z) = \frac{b^a}{\Gamma(a)} \left(\frac{1}{z^2}\right)^{(a-1)} e^{-b/z^2} \left| \frac{-2}{z^3} \right|$$

$$= \frac{b^a}{\Gamma(a)} \left(z^{-2}\right)^{(a-1)} \cdot 2\, z^{-3} \cdot e^{-b/z^2}$$

$$= \frac{2b^a}{\Gamma(a)} z^{-2a+2-3} e^{-b/z^2}$$

$$= \frac{2b^a}{\Gamma(a)} z^{-2a-1} e^{-b/z^2}$$

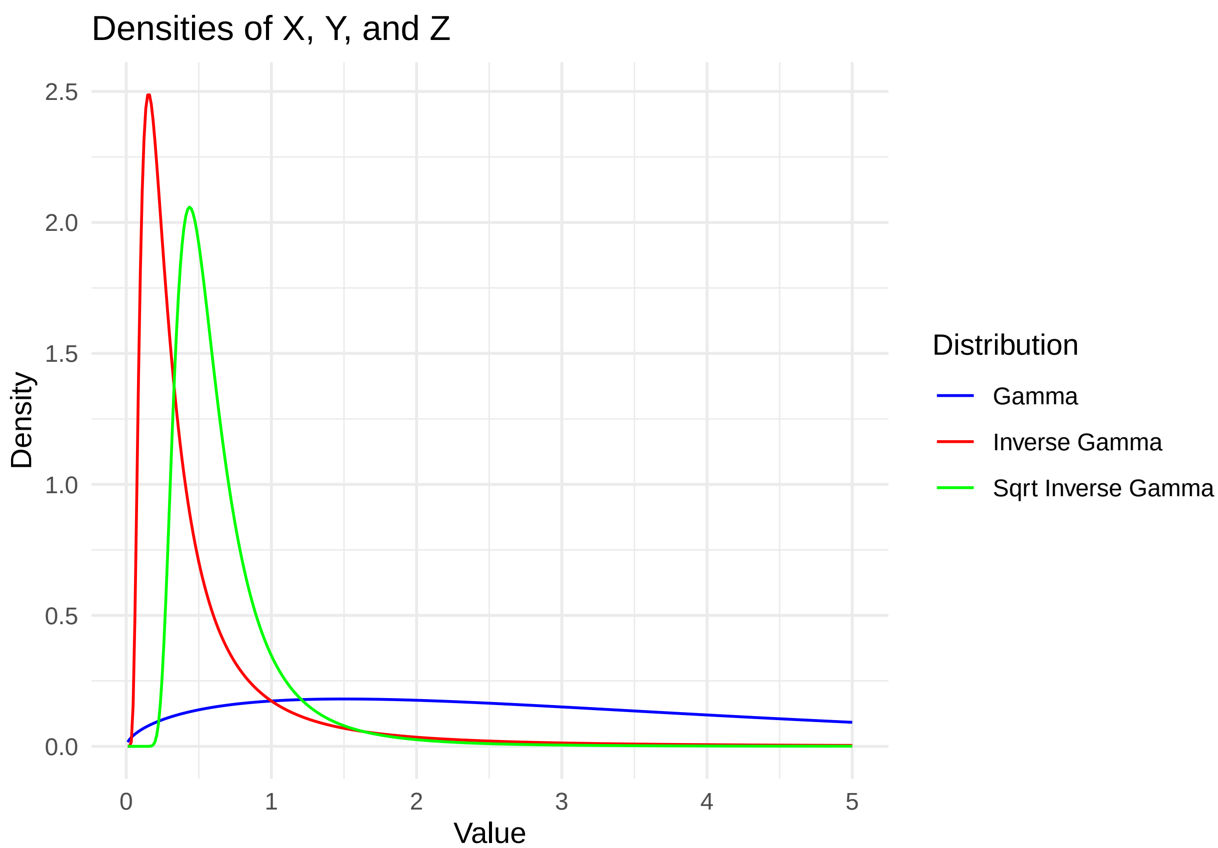
$$= \frac{2b^a}{\Gamma(a)} z^{-(2a+1)} \cdot e^{-b/z^2}$$

density function of Square root Inverse G. (SIG)

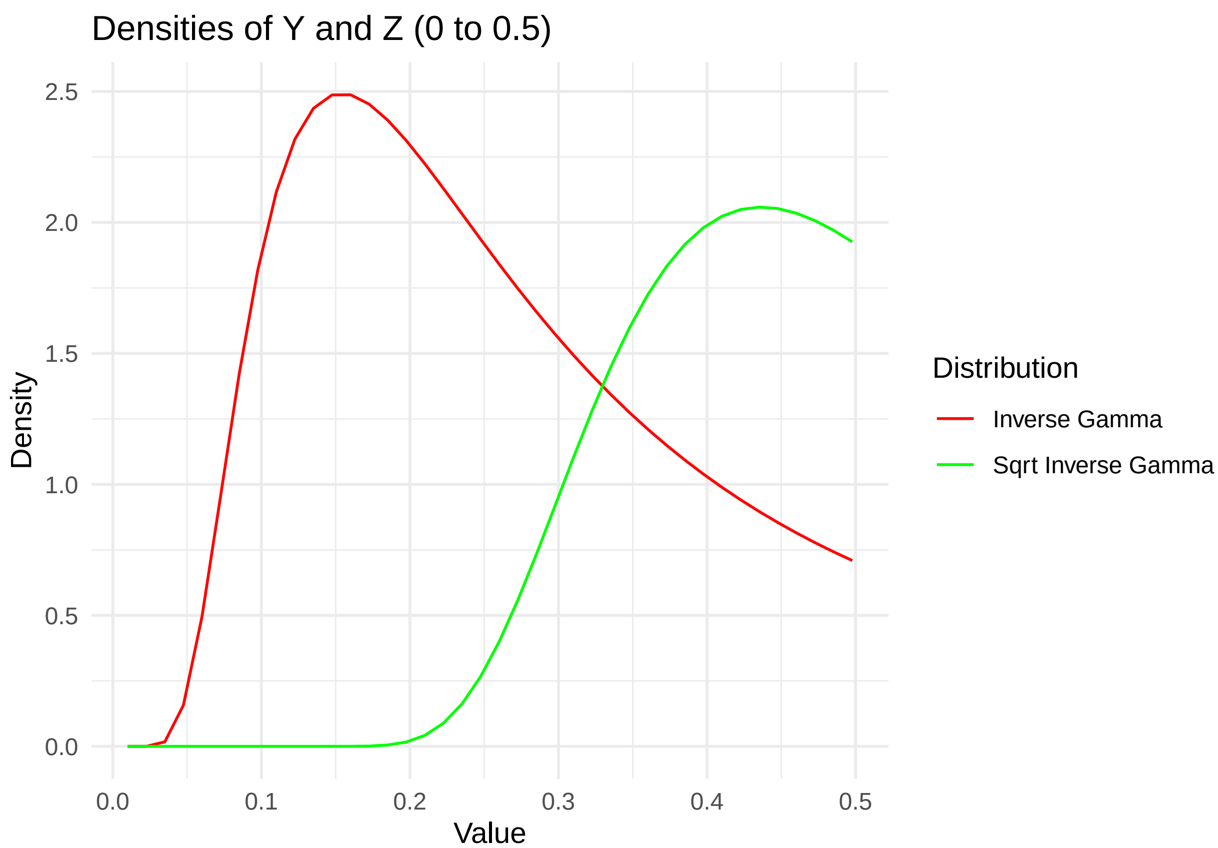
Figure 3: Square root Inverse Gamma

Plots:

```r
# shape parameters
a <- 1.6
b <- 0.4

# gamma
gamma_pdf <- function(x, a, b) {b^a / gamma(a) * x^(a-1) * exp(-b*x)
}

# inverse gamma
IG_pdf <- function(y, a, b) b^a / gamma(a) * y^(-(a + 1)) * exp(-b / y)


# square root gamma
SIG_pdf <- function(z, a, b) 2 * b^a / gamma(a) * z^(-(2 * a + 1)) * exp(-b / (z^2))

# generate values
x_values <- seq(0.01, 5, length = 400)

# data frames for plotting
df_gamma <- data.frame(x = x_values, y = sapply(x_values, gamma_pdf, a, b),
                       Distribution = 'Gamma')
df_IG <- data.frame(x = x_values, y = sapply(x_values, IG_pdf, a, b),
                    Distribution = 'Inverse Gamma')
df_SIG <- data.frame(x = x_values, y = sapply(x_values, SIG_pdf, a, b),
                     Distribution = 'Sqrt Inverse Gamma')

# combine data frames
df <- rbind(df_gamma, df_IG, df_SIG)

# plotting all
ggplot(df, aes(x = x, y = y, color = Distribution)) +
  geom_line() +
  theme_minimal() +
  labs(title = 'Densities of X, Y, and Z', x = 'Value', y = 'Density') +
  scale_color_manual(values = c('Gamma' = 'blue',
                                'Inverse Gamma' = 'red',
                                'Sqrt Inverse Gamma' = 'green'))
```




```
# filter data frame for domain range between 0 and 0.5
df_filtered <- df[df$x <= 0.5,]

# plotting Y and Z for the range 0 to 0.5
ggplot(df_filtered[df_filtered$Distribution != 'Gamma', ],
       aes(x = x, y = y, color = Distribution)) +
  geom_line() +
  theme_minimal() +
  labs(title = 'Densities of Y and Z (0 to 0.5)', x = 'Value', y = 'Density') +
  scale_color_manual(values = c('Inverse Gamma' = 'red',
                                'Sqrt Inverse Gamma' = 'green'))
```



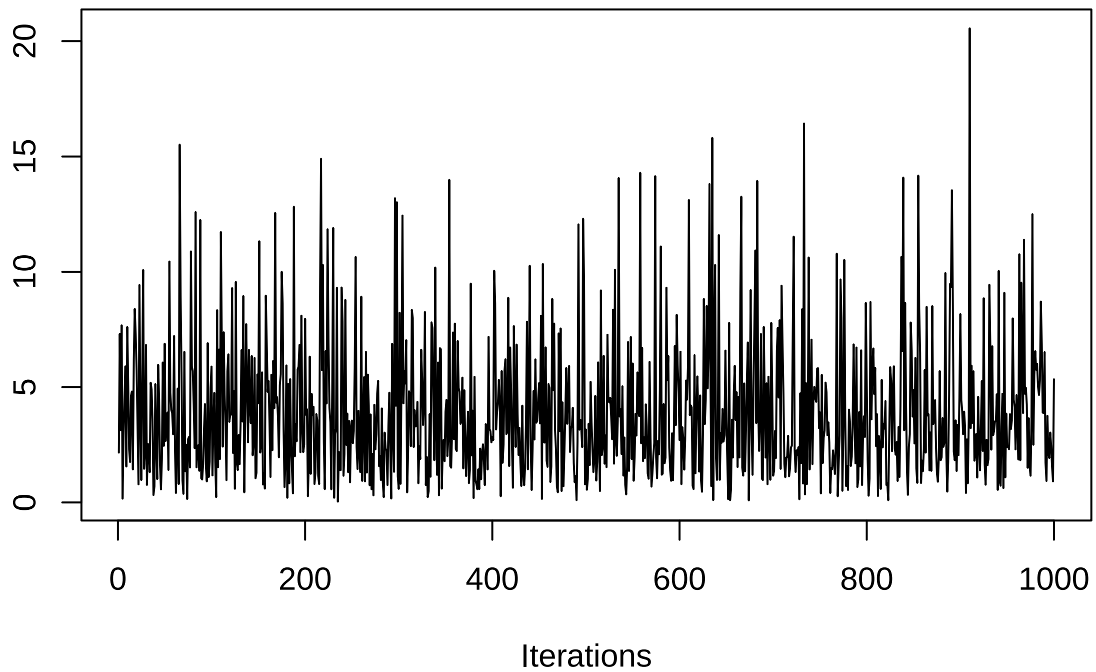
Interpretation:

- The **IG(1.6, 0.4)** distribution spikes quickly in the assignment of probabilities, showing a sharp increase in probability for small values.
- The **SIG(1.6, 0.4)** distribution assigns probability 0 to values (very!) close to 0, indicating a rapid decrease in probability near 0. However, this increase in probability doesn't occur until around 0.2, indicating a slower rise in probability compared to the IG distribution.

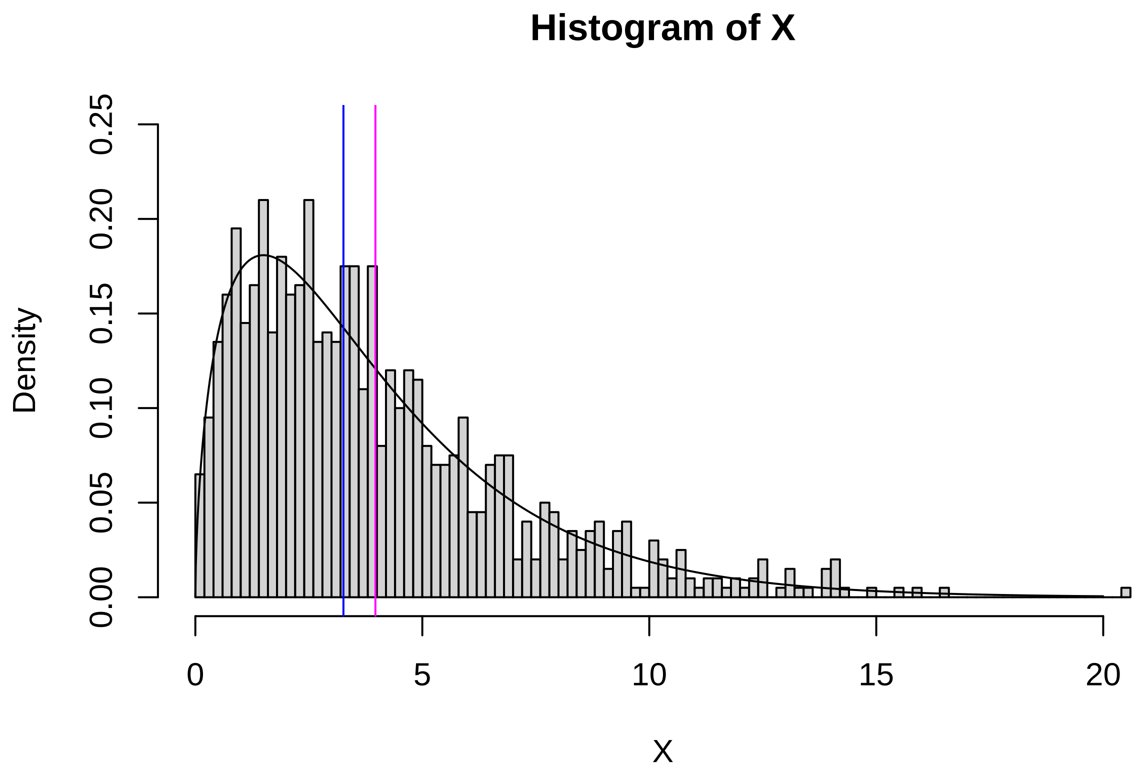
Exercise 6:

a)

```
library(coda)

set.seed(44566)

M <- 1000

X <- rgamma(n = M, shape = 1.6, rate = 0.4)

X <- as.mcmc(X)

# Traceplot
coda::traceplot(X)
```




```r
# Histogram overlayed true density
seq <- seq(0, 20, by = 0.01)

hist(X, probability = TRUE, ylim = c(0, 0.25), breaks = 100)
lines(seq, dgamma(seq, shape = 1.6, rate = 0.4))
abline(v = mean(X), col = "magenta")
abline(v = median(X), col = "blue")
```




```r
# Sample mean and sample median
mean(X)
```

```
## [1] 3.9667
```

```r
median(X)
```

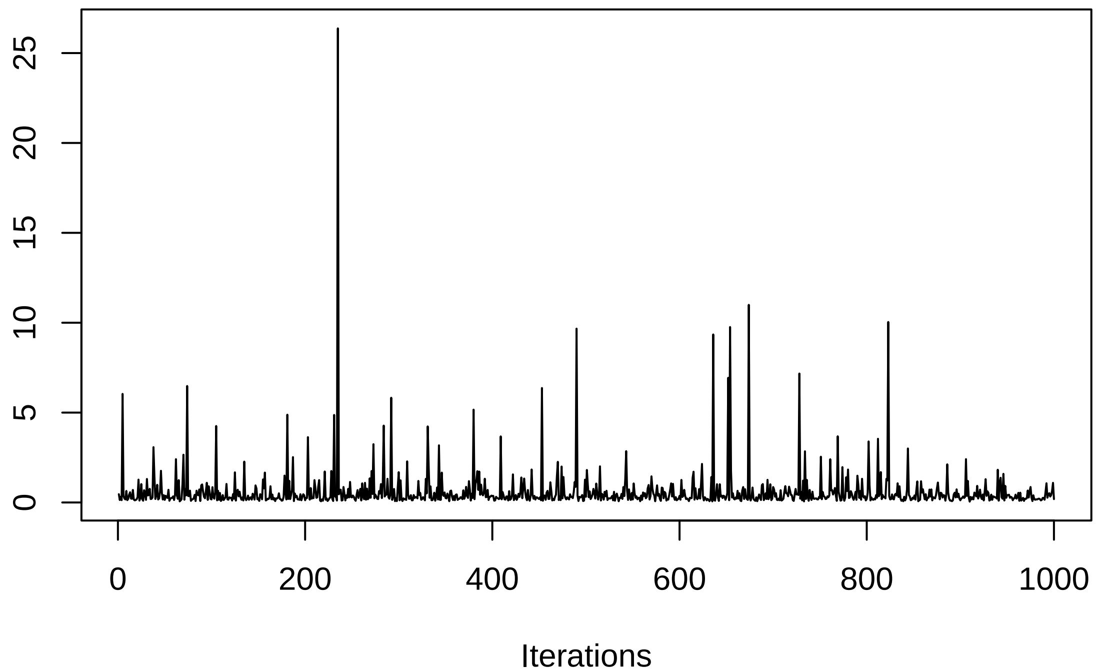
```
## [1] 3.262569
```

b)

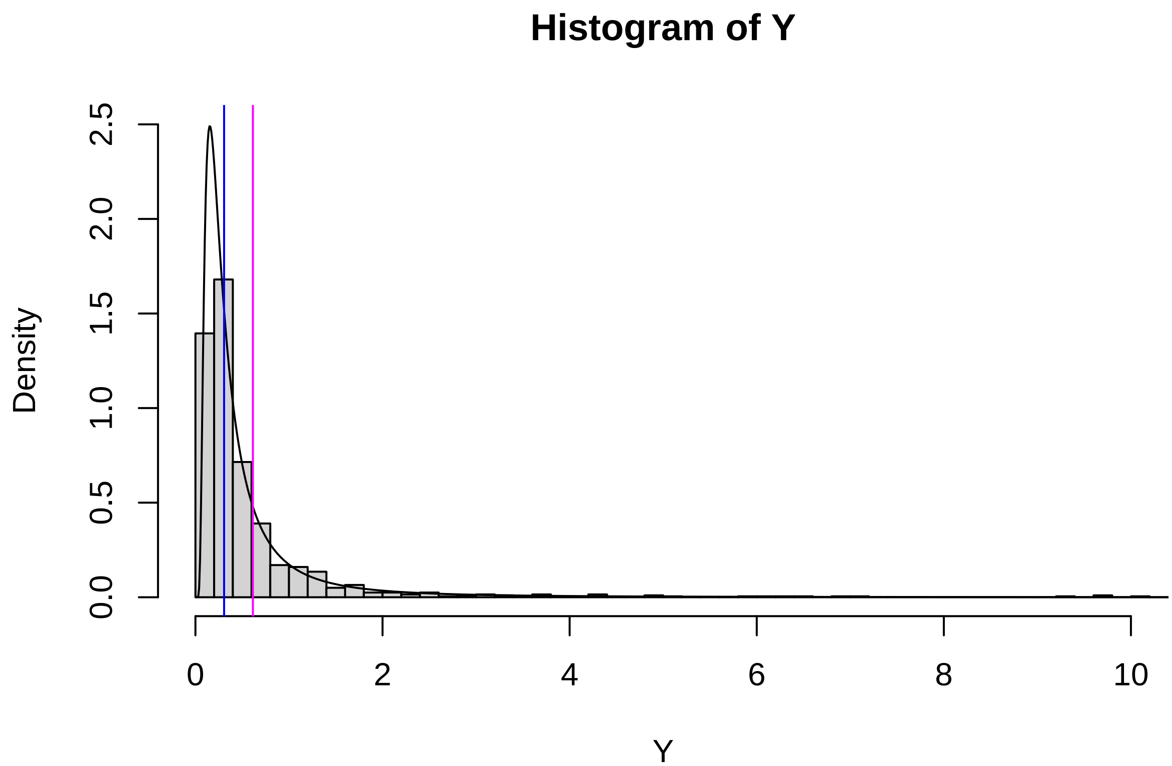
```r
library(invgamma)
```

```
##
## Attaching package: 'invgamma'

## The following objects are masked from 'package:MCMCpack':
##
##     dinvgamma, rinvgamma
```

```r
Y <- 1/X
Y <- as.mcmc(Y)

# Traceplot
coda::traceplot(Y)
```




```r
# Histogram overlayed true density
hist(Y, probability = TRUE, xlim = c(0, 10), ylim = c(0, 2.5), breaks = 100)
lines(seq, dinvgamma(seq, shape = 1.6, rate = 0.4))
abline(v = mean(Y), col = "magenta")
abline(v = median(Y), col = "blue")
```




```
# Sample mean and sample median
mean(Y)
```

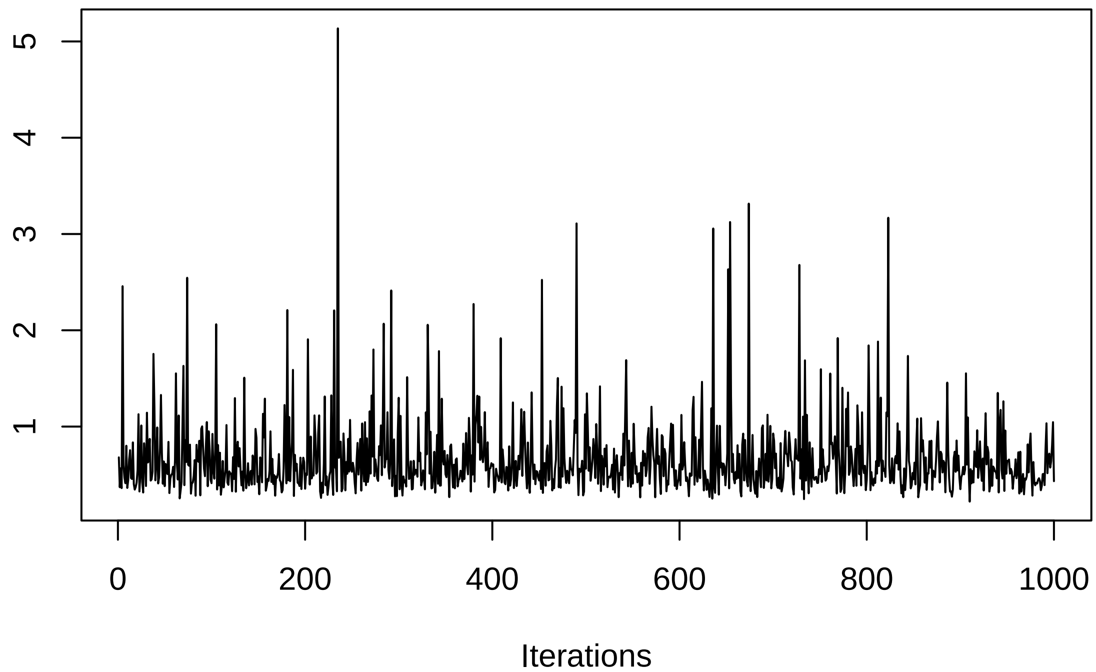
```
## [1] 0.6143637
```

```
median(Y)
```

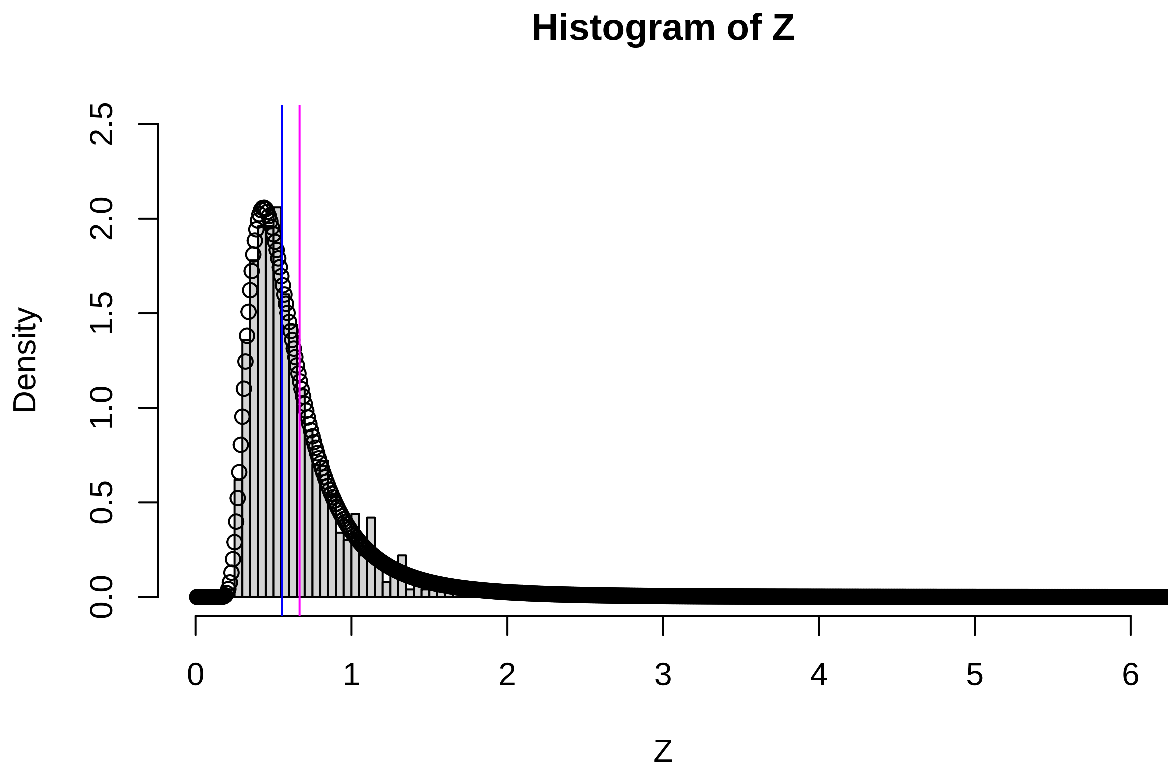
```
## [1] 0.3065072
```

c)

```
Z <- sqrt(1/X)

# Density of the square-root inverse gamma distribution
sqinvgamma <- function(z, alpha, beta){
  beta^alpha/gamma(alpha) * 2/z^(1 + 2*alpha) * exp(-beta*1/z^2)
}


# Traceplot
coda::traceplot(Z)
```




```r
# Histogram overlayed true density
hist(Z, probability = TRUE, xlim = c(0, 6), ylim = c(0, 2.5), breaks = 100)
for (i in seq){
  points(i, sqinvgamma(z = i, alpha = 1.6, beta = 0.4))
}
abline(v = mean(Z), col = "magenta")
abline(v = median(Z), col = "blue")
```




```r
# Sample mean and sample median
mean(Z)
```

```
## [1] 0.6672328
```

```r
median(Z)
```

```
## [1] 0.5536309
```

The Median of Y is 1/Median(X), and the median of Z is sqrt(1/Median(X)) Since the median divides the probability mass of the density into to equally sized areas, the transformations can be applied to the median as well. For the mean, this does not work.