Supplement to slide set 5

Ege Rubak and Jesper Møller

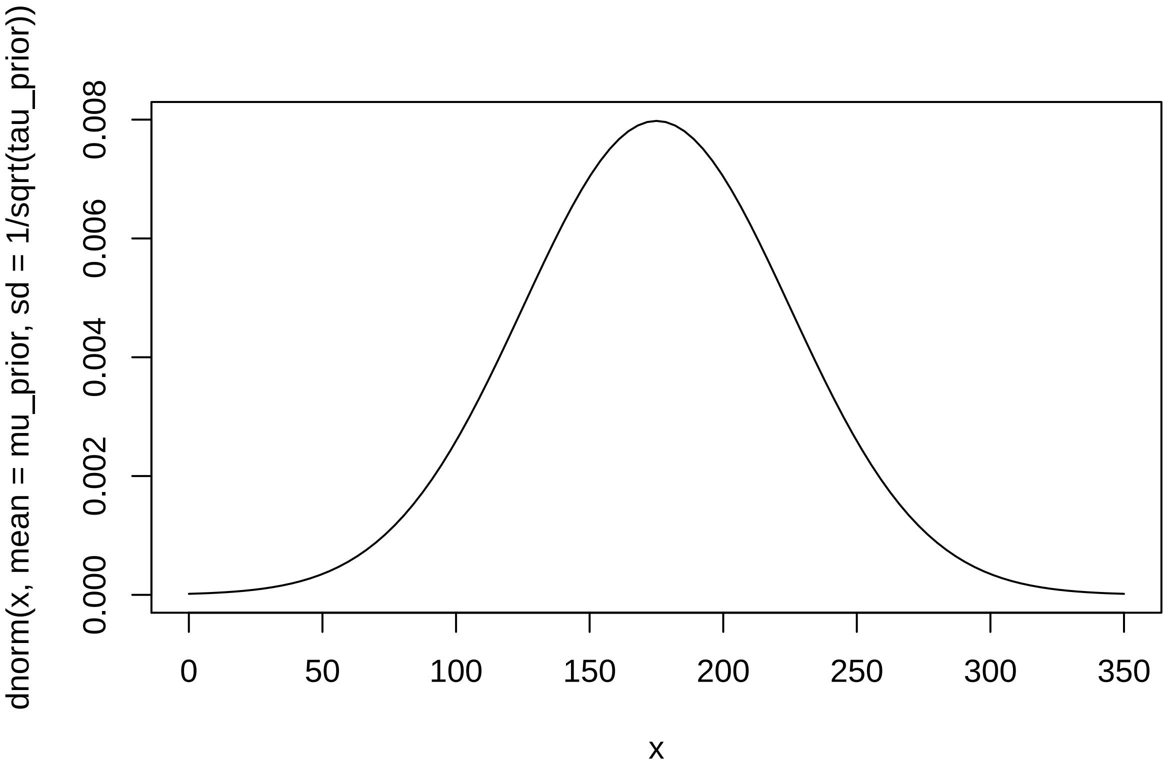
“Scientific problem”

- A researcher (Gulliver, say) travels to a new country and expects the height X of the locals to be normally distributed with unknown mean μ and unknown precision τ.

Prior for the mean

- Based on previous travels he expects the mean height μ could be anything between 50 and 300 cm but most likely something in the middle, which is translated to the prior $\mu \sim N(175, 1/50^2)$.:

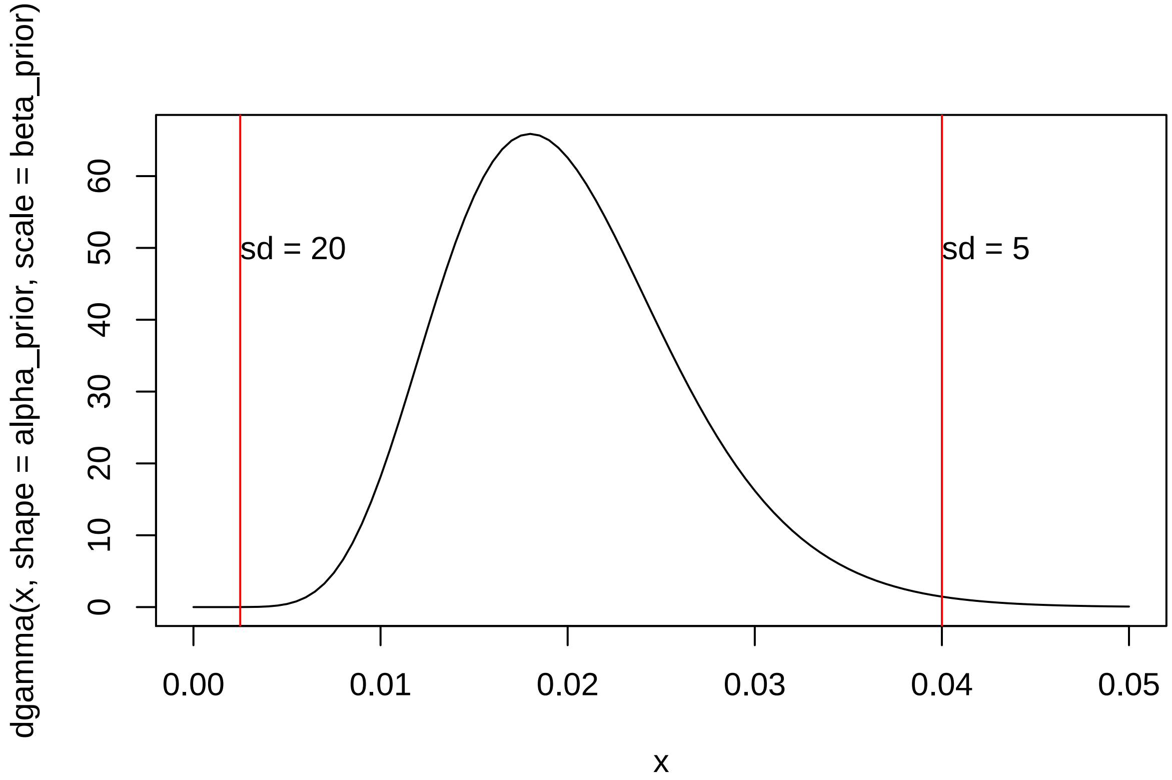
```r
tau_prior <- 1/50^2
mu_prior <- 175
curve(dnorm(x, mean = mu_prior, sd = 1/sqrt(tau_prior)), from = 0, to = 350)
```



Prior for the precision

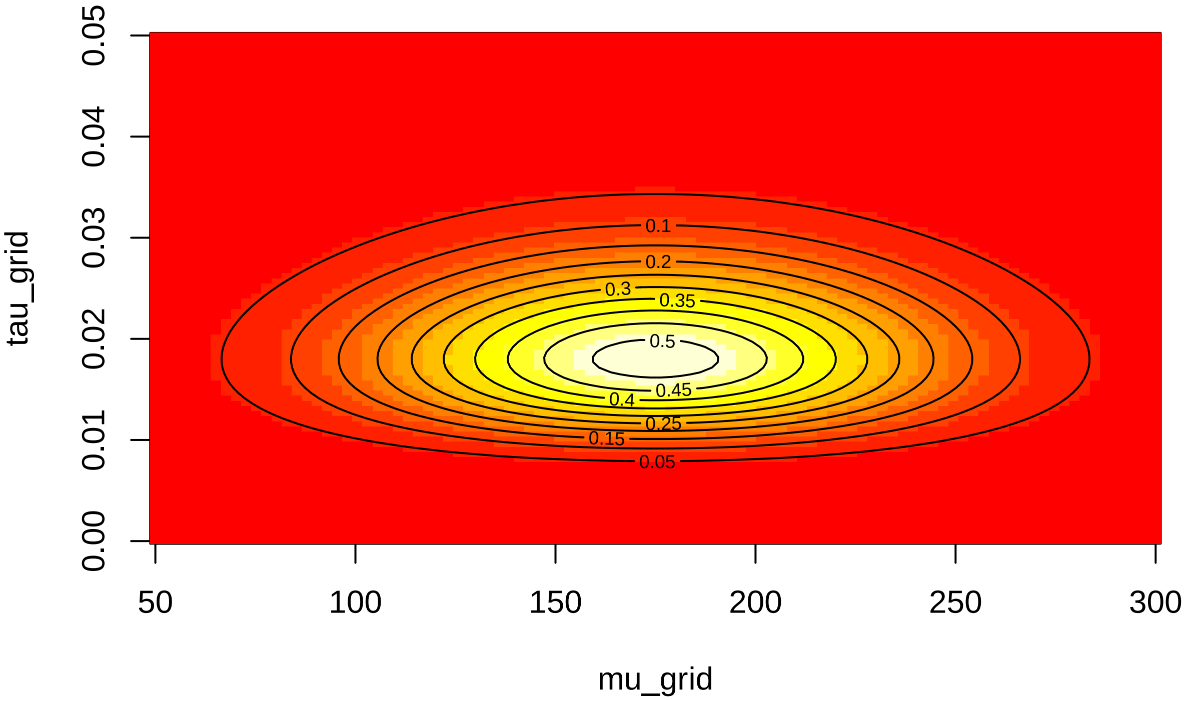
- In the researcher’s experience any country has a pretty homogeneous height distribution so the standard deviation of height should be of the order 5 to 20 cm, i.e. precision between 1/400=0.0025 and 1/25=0.04. After some experiments with the Gamma-distribution the prior Gamma(10, 0.002) is chosen:

```
alpha_prior <- 10
beta_prior <- .002
curve(dgamma(x, shape = alpha_prior, scale = beta_prior), from = 0, to = .05)
abline(v = 1/c(20, 5)^2, col = "red")
text(x=1/c(20, 5)^2, 50, labels = c("sd = 20", "sd = 5"), adj = 0)
```



Joint prior

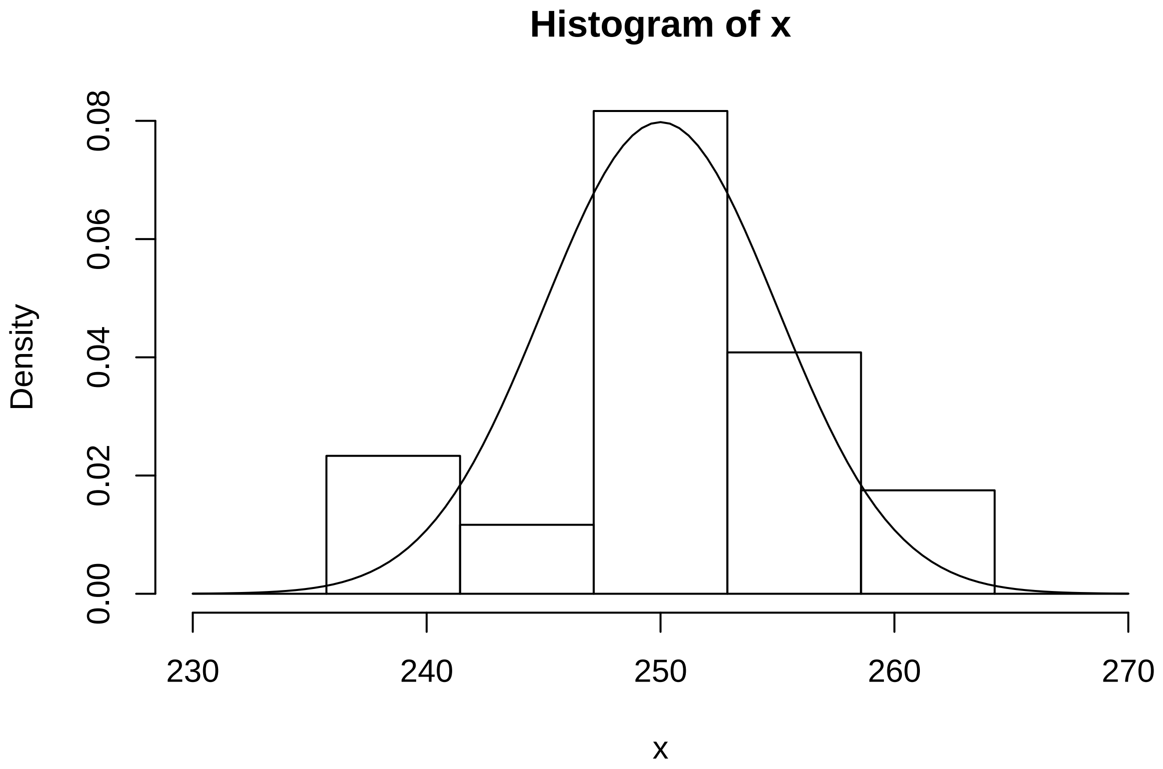
```
mu_grid <- seq(50, 300, length.out = 100)
tau_grid <- seq(0,.05,length.out=100)
prior <- function(mu, tau){
  x <- dnorm(mu, mean = mu_prior, sd = 1/sqrt(tau_prior))
  y <- dgamma(tau, shape = alpha_prior, scale = beta_prior)
  return(x*y)
}
prior_grid <- outer(mu_grid, tau_grid, prior)
image(mu_grid, tau_grid, prior_grid)
contour(mu_grid, tau_grid, prior_grid, add = TRUE)
```



Simulated Gaussian data

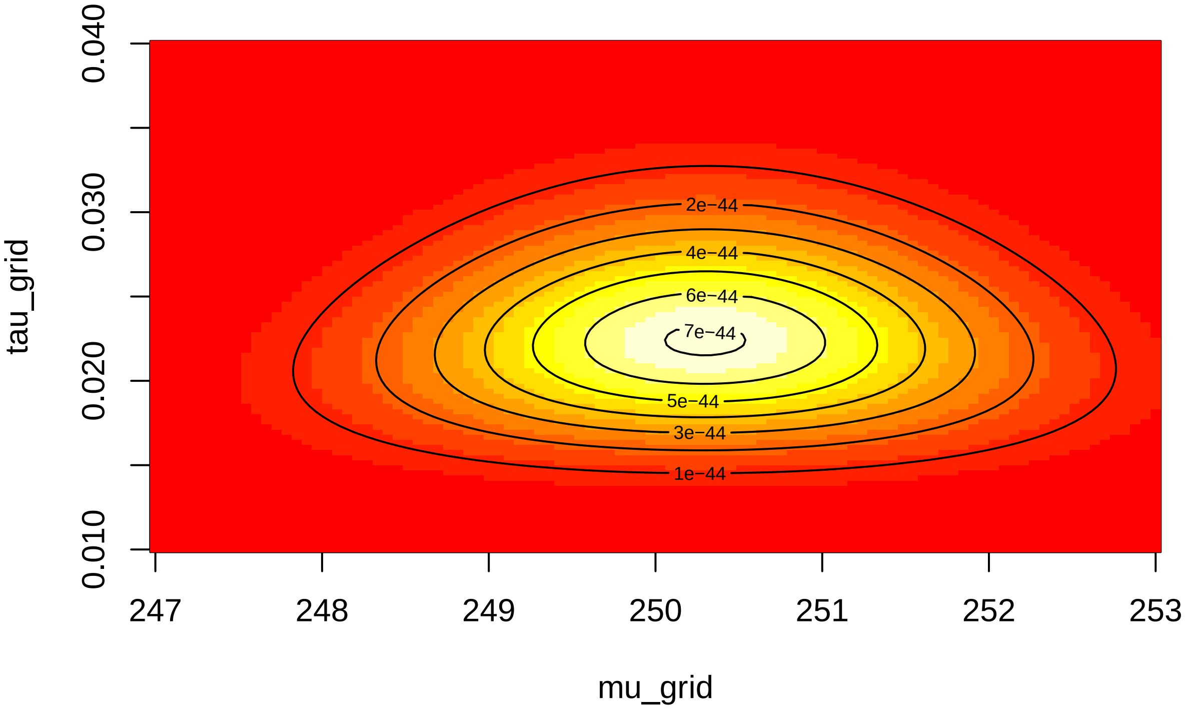
The researcher doesn't know, but in fact the country has a very homogeneous population of tall people, and the available data is:

```
mu_true <- 250
tau_true <- 1/5^2
n <- 30
set.seed(42)
x <- rnorm(n, mean = mu_true, sd = sqrt(1/tau_true))
x_bar <- mean(x)
hist(x, prob = TRUE, breaks = seq(250 - 20, 250 + 20, length.out = 8), ylim = c(0,0.08))
curve(dnorm(x, mean = mu_true, sd = 1/sqrt(tau_true)), add = TRUE)
```



Posterior density

```r
mu_grid <- seq(247, 253, length.out = 100)
tau_grid <- seq(0.01, .04, length.out = 100)
post_grid <- matrix(0, nrow = length(mu_grid), ncol = length(tau_grid))
for(i in seq_along(mu_grid)){
  for(j in seq_along(tau_grid)){
    post_grid[i,j] <- prior(mu_grid[i], tau_grid[j]) *
      prod(dnorm(x, mean = mu_grid[i], sd = 1/sqrt(tau_grid[j])))
  }
}
image(mu_grid, tau_grid, post_grid)
contour(mu_grid, tau_grid, post_grid, add = TRUE)
```



Full conditionals

Functions for sampling from the full conditional distributions:

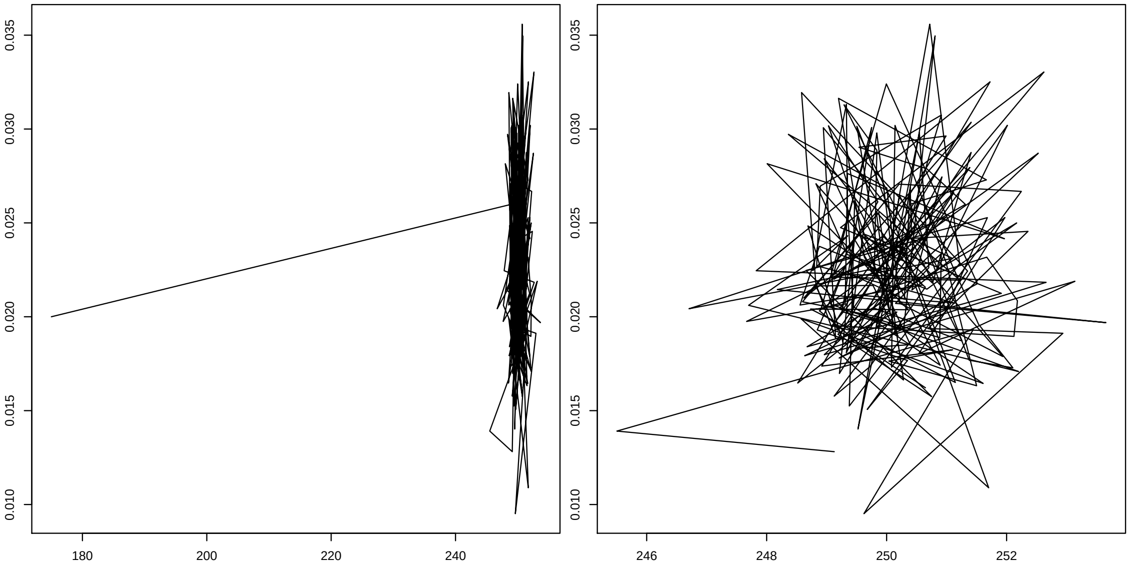
```r
rmu <- function(tau){
  cond_mean <- (n*tau*x_bar + tau_prior*mu_prior)/(n*tau + tau_prior)
  cond_tau <- n*tau + tau_prior
  return(rnorm(1, mean = cond_mean, sd = sqrt(1/cond_tau)))
}
rtau <- function(mu){
  cond_alpha <- n/2 + alpha_prior
  cond_beta <- 1/(sum((x-mu)^2)/2 + 1/beta_prior)
  return(rgamma(1, shape = cond_alpha, scale = cond_beta))
}
```

Running the Gibbs sampler

```r
n_samples <- 200 # Number of samples
samples_tau <- samples_mu <- rep(0, n_samples) # Vectors to write results in
samples_tau[1] <- alpha_prior * beta_prior # Start in prior mean
samples_mu[1] <- mu_prior # Start in prior mean
for(i in 2:n_samples){
  samples_mu[i] <- rmu(samples_tau[i-1])
  samples_tau[i] <- rtau(samples_mu[i])
}
```

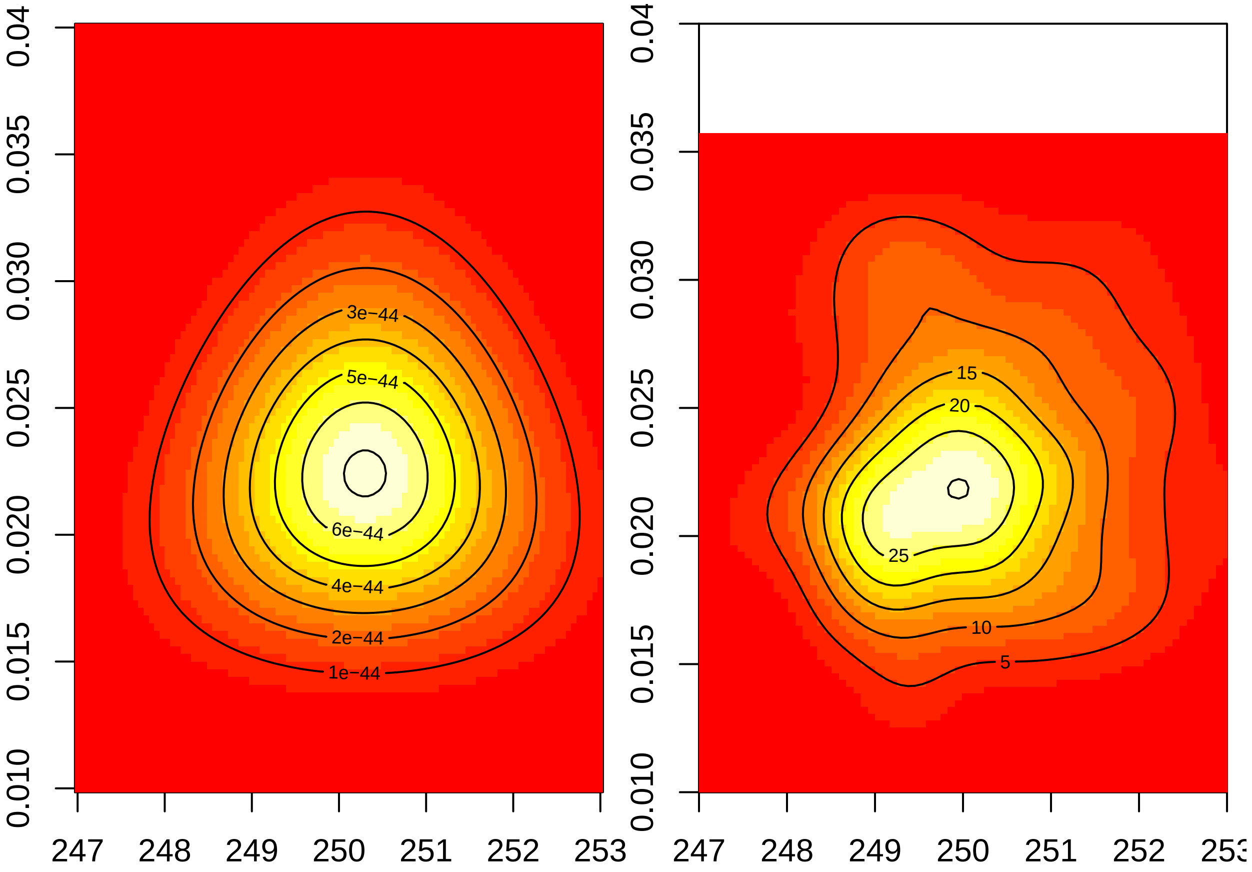
Displaying the results: Bivariate trace

```r
samples <- cbind(samples_mu, samples_tau)
par(mfrow=c(1,3), mar = c(2,2,.5,.5))
plot(samples, type = "l")
burn_in <- 1:4
plot(samples[-burn_in,], type = "l")
chain <- samples[-burn_in,]
```



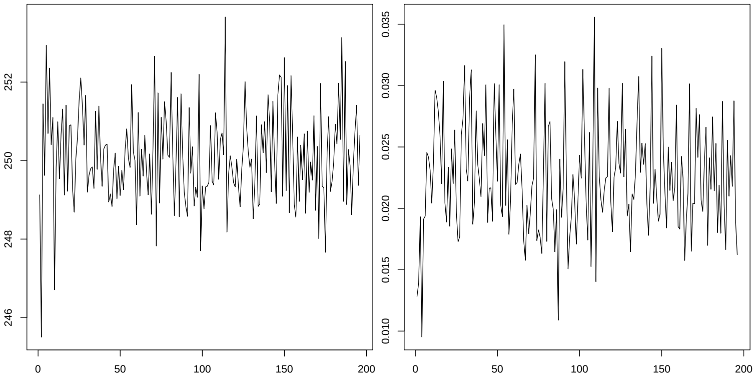
Displaying the results: Comparing densities

```r
par(mfrow=c(1,2), mar = c(2,2,.5,.5))
image(mu_grid, tau_grid, post_grid)
contour(mu_grid, tau_grid, post_grid, add = TRUE)
smooth_chain <- MASS::kde2d(chain[,1], chain[,2], n = 100)
image(smooth_chain, xlim = range(mu_grid), ylim = range(tau_grid))
contour(smooth_chain, add = TRUE)
```



Displaying the results: Uni-variate traces

```r
par(mfrow=c(1,2), mar = c(2,2,.5,.5))
plot(chain[,1], type = "l")
plot(chain[,2], type = "l")
```



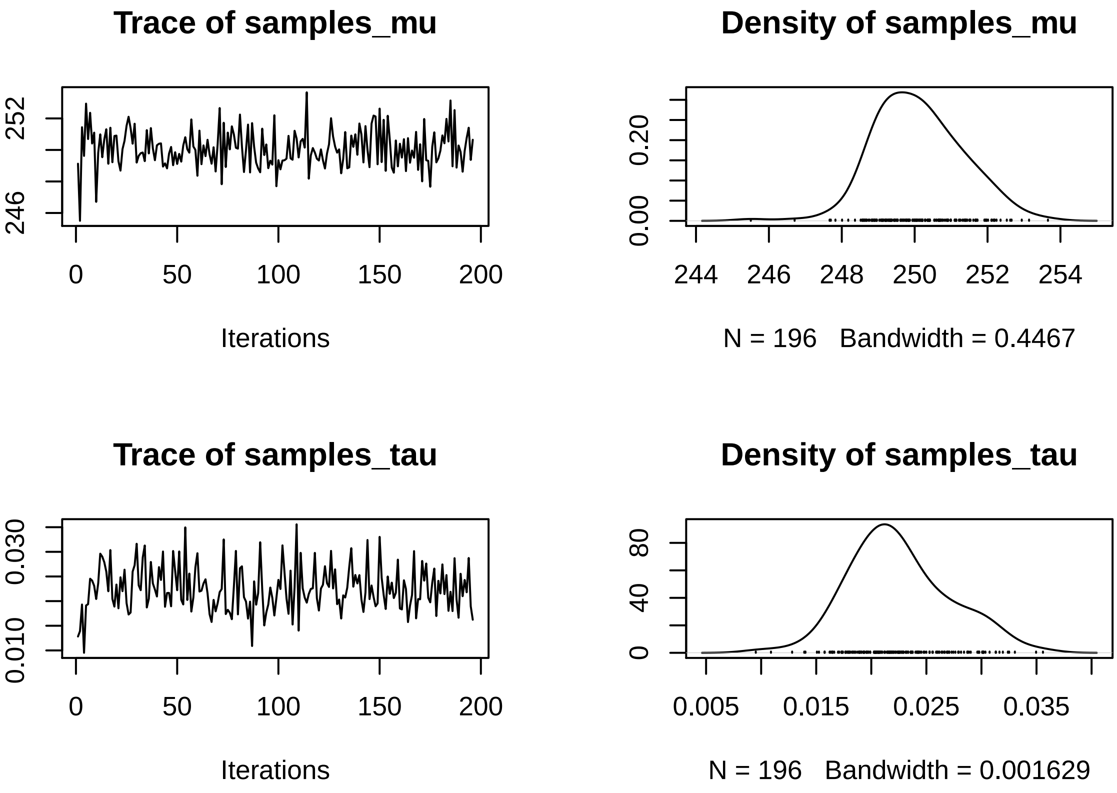
Displaying results with the coda package: Summary

```
library(coda)
chain <- mcmc(chain)
summary(chain)
```

```
##
## Iterations = 1:196
## Thinning interval = 1
## Number of chains = 1
## Sample size per chain = 196
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##                  Mean       SD  Naive SE Time-series SE
## samples_mu  250.10025 1.211037 0.0865027      0.0883548
## samples_tau   0.02253 0.004589 0.0003278      0.0003278
##
## 2. Quantiles for each variable:
##
##                  2.5%       25%       50%       75%   97.5%
## samples_mu  247.98516 249.25124 250.03662 250.89770 252.540
## samples_tau   0.01493   0.01937   0.02192   0.02529   0.032
```

Displaying results with the coda package: Plot

```
plot(chain)
```

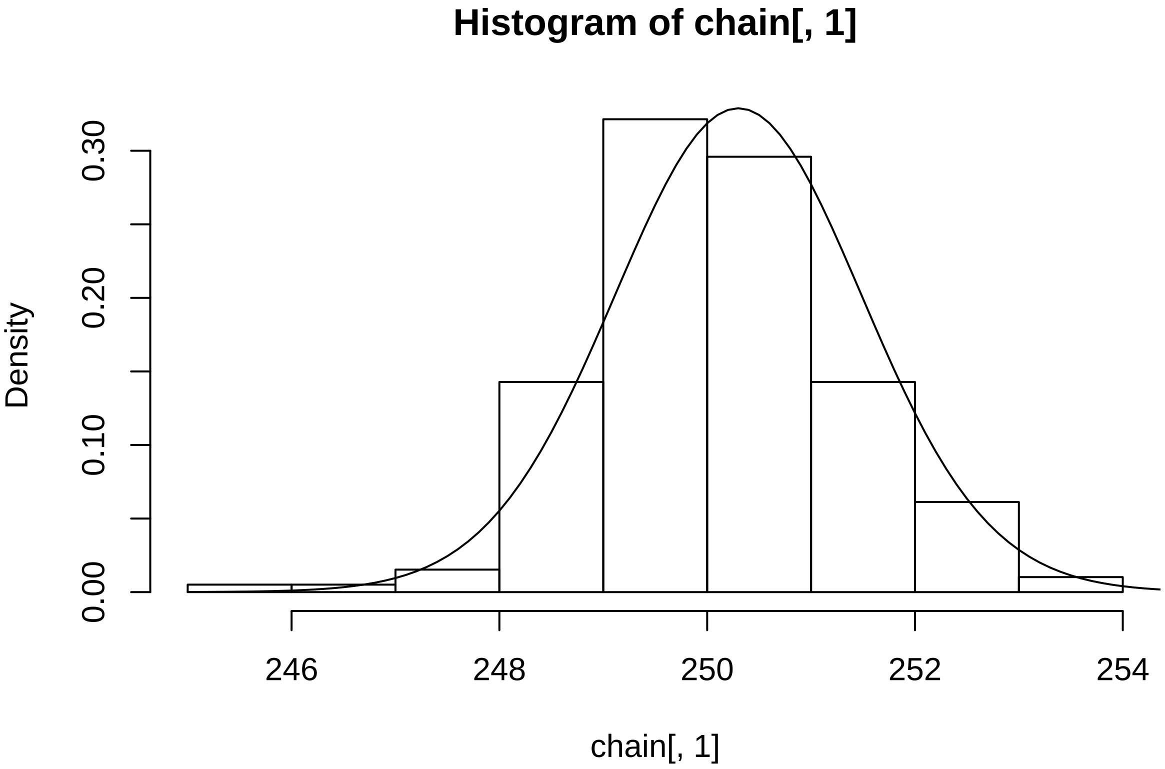


Marginal distribution

Remember the non-starndard marginal density for μ found in the main slides:

$$\pi(\mu|x) \propto \exp\left(-\frac{1}{2}\tau_0(\mu-\mu_0)^2\right)\left(\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2+\frac{1}{\beta}\right)^{-(n/2+\alpha)}$$

```
f <- function(mu){
  exp(-tau_prior*(mu-mu_prior)^2/2)*
    (sum((x-mu)^2)/2+1/beta_prior)^(-(n/2+alpha_prior))
}
f_vec <- Vectorize(f)
int_f <- integrate(f_vec, 230, 270)
hist(chain[,1], prob = TRUE)
curve(f_vec(x)/int_f$value, 245, 255, add = TRUE)
```



What do we use the samples for?

- Based on the plots above it looks like we are really sampling from the correct posterior distribution, but what should we do with these samples?
- Answer: Any statistical inference we like! (Which can be phrased as a mean under the posterior distribution.)
- This could be $P(249 < \mu < 251)$, $\mathrm{E}(\mu)$, ...
- E.g.

```
mean(chain[,1])
```

```
## [1] 250.1003
```