# Gibbs Sampling for estimate of parameter posterior distributions

Modelling y data as a normal distribution with mean = mu, and variance = 1/tau We read in the data, and we start with an estimate of tau to start the gibbs sampling off. Using the posterior distributions derived by hand, we have come up with the mu and tau sampling distributions in the for loop.

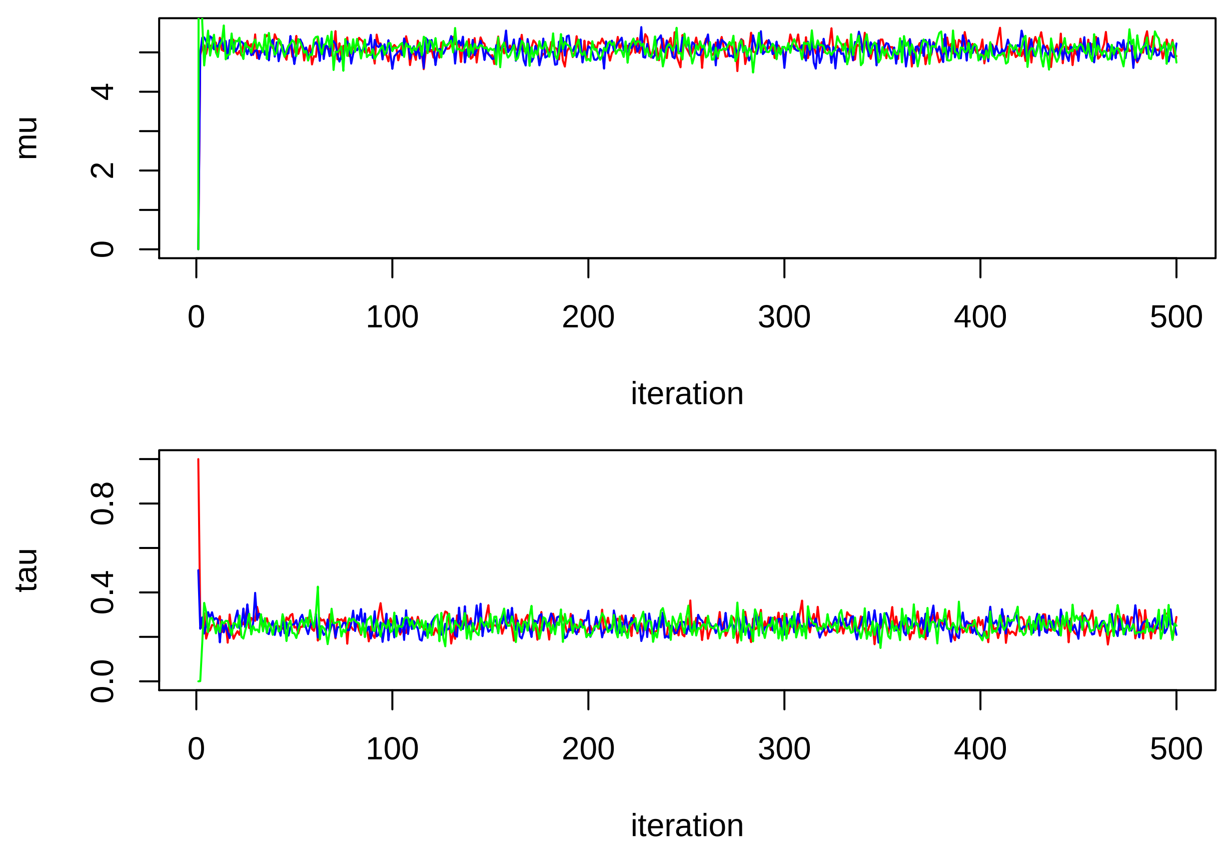
```
gibbs = function(tau1) {

y = scan("y data.txt", what=double())

y_bar = mean(y); s2 = var(y);
n = 100; k = 500;

mu = rep(0, k); tau = rep(0, k);

tau[1] <- tau1

# posterior sampling distibutions
for (i in 2:(k)) {
  mu[i] <- rnorm(1, mean = y_bar, sd = sqrt(1/(n*tau[i-1])))
  k = sum(y^2) - 2*n*y_bar*(mu[i]) + n*((mu[i])^2)
  tau[i] <- rgamma(1, shape = n/2, scale = (2 / k))
}

result <- list(x1 = mu, x2 = tau)
result

}
```

```
means <- function(chain) {
    return(c(mean(chain$x1), mean(chain$x2)))
}
```

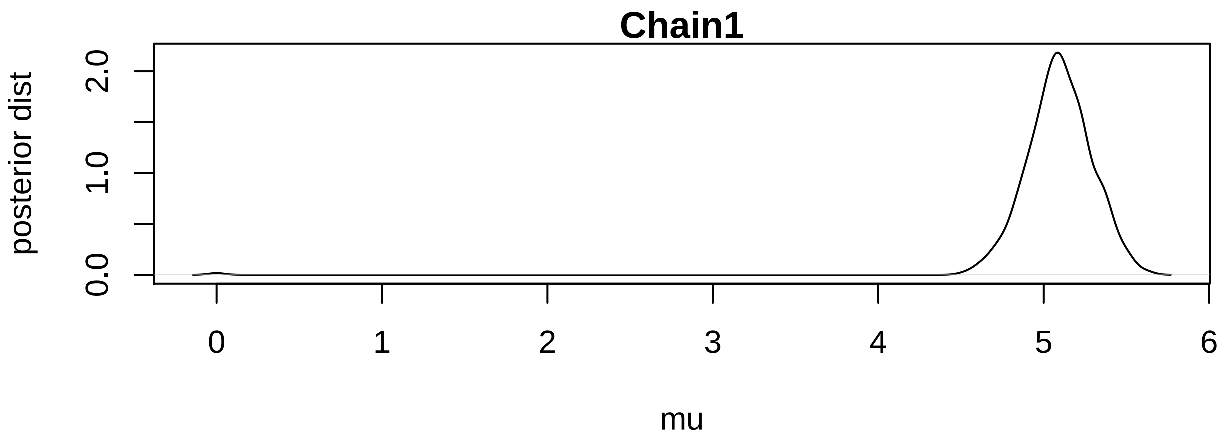
```
cred_int <- function(chain) {
    mu_95 <- quantile(chain$x1, 0.95); mu_05 <- quantile(chain$x1, 0.05)
    tau_95 <- quantile(chain$x2, 0.95); tau_05 <- quantile(chain$x2, 0.05)

    intervals <- matrix(0,2,2)
    intervals[1,1] <- mu_05; intervals[1,2] <- mu_95;
    intervals[2,1] <- tau_05; intervals[2,2] <- tau_95;

    return(intervals)
}
```

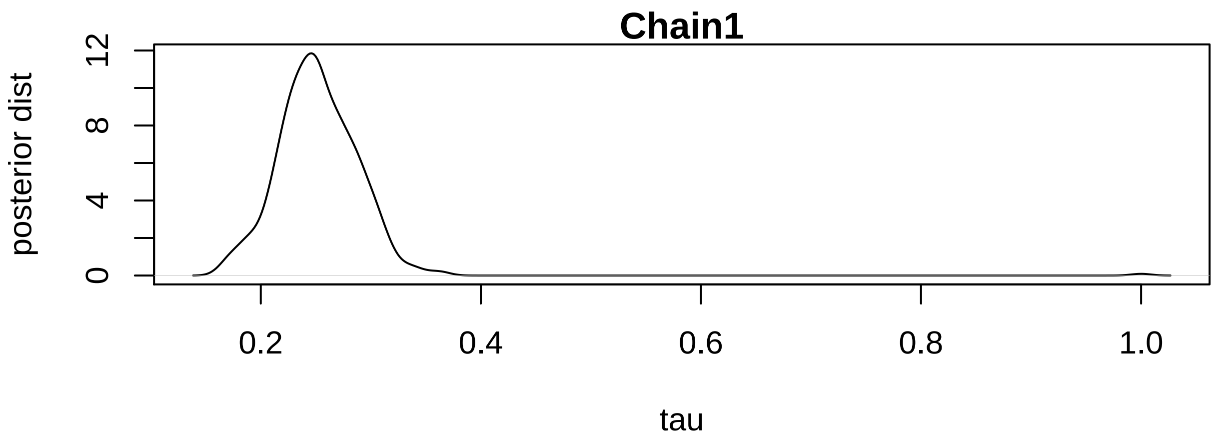
## starting guesses here for the chain

```r
test_result1 = gibbs(tau1 = 1)
test_result2 = gibbs(tau1 = 0.5)

# try a 'bad' starting guess
test_result3 = gibbs(tau1 = 0.000001)

nreps = 500;

par(mfrow=c(2,1), mar=c(4,4,1,1))

plot(1:nreps, test_result1$x1, type="l", col="red", ylim = c(0, max(test_result1$x1, test_result2$x1)),
points(1:nreps, test_result2$x1, type="l", col="blue")
points(1:nreps, test_result3$x1, type="l", col="green")

plot(1:nreps, test_result1$x2, type="l", col="red", ylim = c(0, max(test_result1$x2, test_result2$x2)),
points(1:nreps, test_result2$x2, type="l", col="blue")
points(1:nreps, test_result3$x2, type="l", col="green")
```

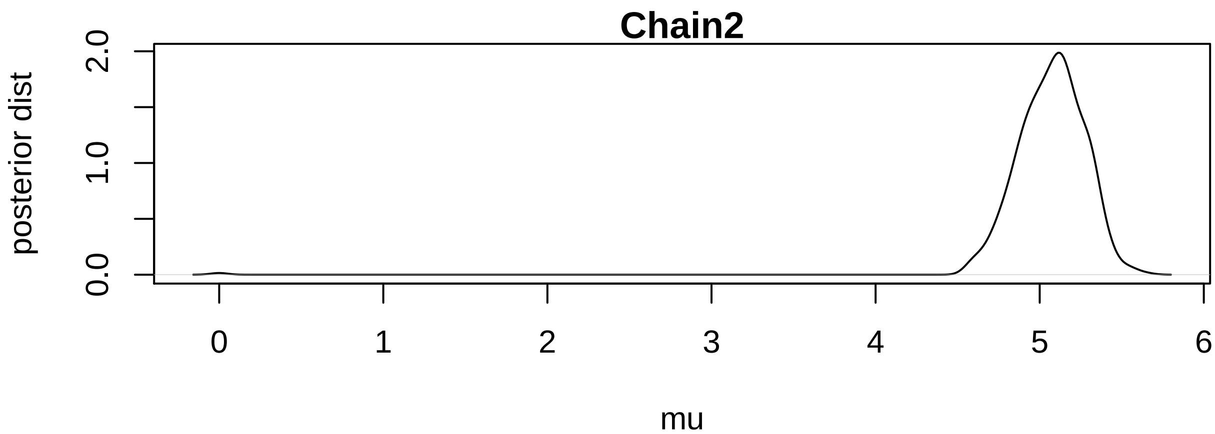


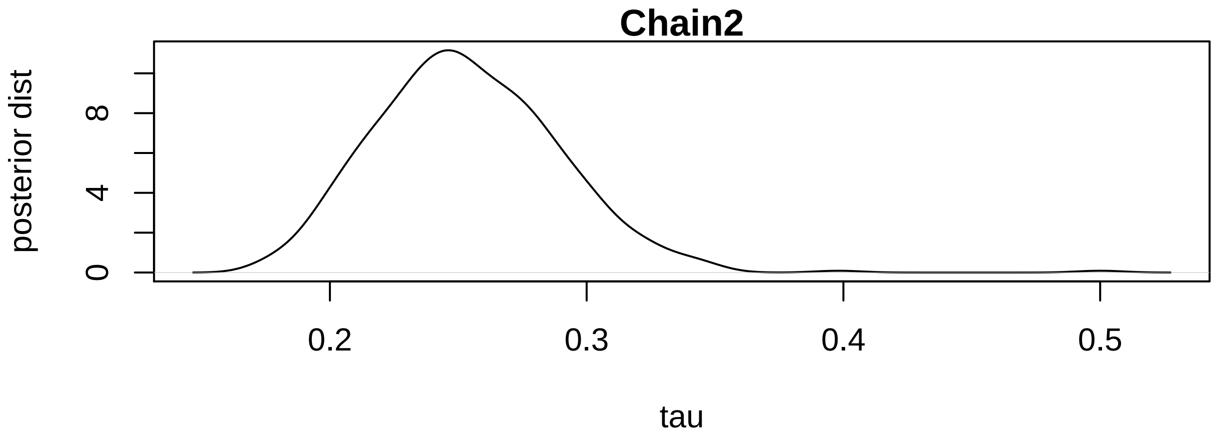

```r
# Densities of 1st chain
plot(density(test_result1$x1), ylab="posterior dist", xlab="mu", main="Chain1")
plot(density(test_result1$x2), ylab="posterior dist", xlab="tau", main="Chain1")
```




```r
# Densities of 2nd chain
plot(density(test_result2$x1), ylab="posterior dist", xlab="mu", main="Chain2")
plot(density(test_result2$x2), ylab="posterior dist", xlab="tau", main="Chain2")
```




```r
(mean_test_result1 <- means(test_result1))
```

```
## [1] 5.086460 0.253068
```

```r
(mean_test_result2 <- means(test_result2))
```

```
## [1] 5.0639462 0.2537444
```

```r
(cred_int_test_result1 <- cred_int(test_result1))
```

```
##           [,1]      [,2]
## [1,] 4.7718470 5.3975394
## [2,] 0.1925715 0.3084869
```

```r
(cred_int_test_result2 <- cred_int(test_result2))
```

```
##           [,1]      [,2]
## [1,] 4.7419524 5.3639195
## [2,] 0.1989986 0.3123465
```

Note: As we can see from the trace plots - all three converge quickly to the correct values of mu = 5, and tau = 0.25. The green line is the worst starting position, but still converges quickly.