Gibbs Sampling Problem 1

Ganchao Wei

5/26/2021

1. Gibbs Sampling for a Bivariate Normal Distribution

1.1 Notations and Full Conditional Distributions

Let $\boldsymbol{X} = (X_1, X_2)' \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} = (\mu_1, \mu_2)'$ and $\boldsymbol{\Sigma} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$. Therefore, the full conditional distributions of X_1 and X_2 are:

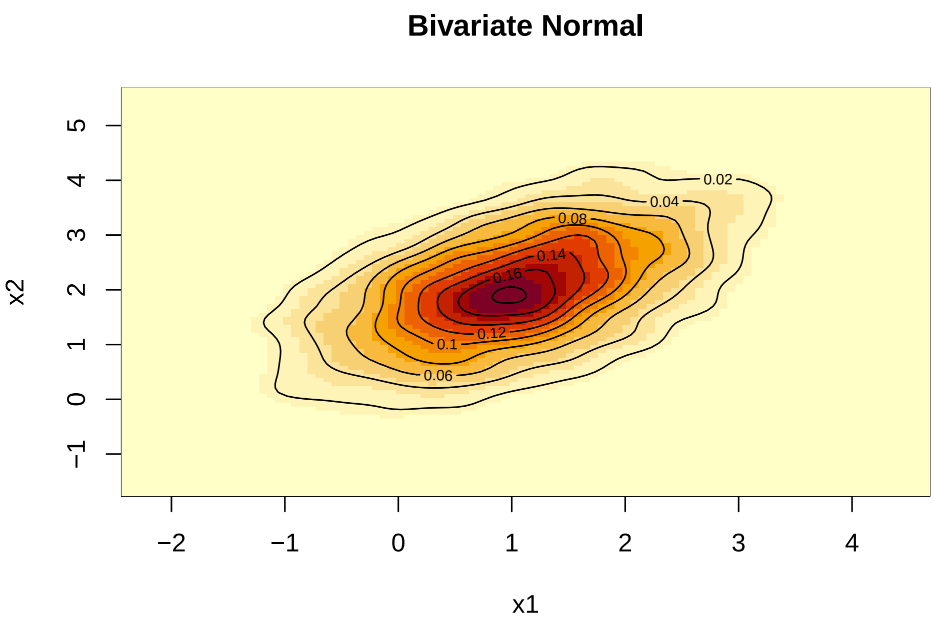
$$X_1|X_2 = x_2 \sim N(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$$

$$X_2|X_1 = x_1 \sim N(\mu_1 + \Sigma_{21}\Sigma_{11}^{-1}(x_1 - \mu_1), \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})$$

1.2 Simulation

Here, I set $\boldsymbol{\mu} = (1, 2)'$ and $\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$. To visualize the distribution, let's draw 5000 samples directly from the distribution, with contours shown (from kernel density estimation).

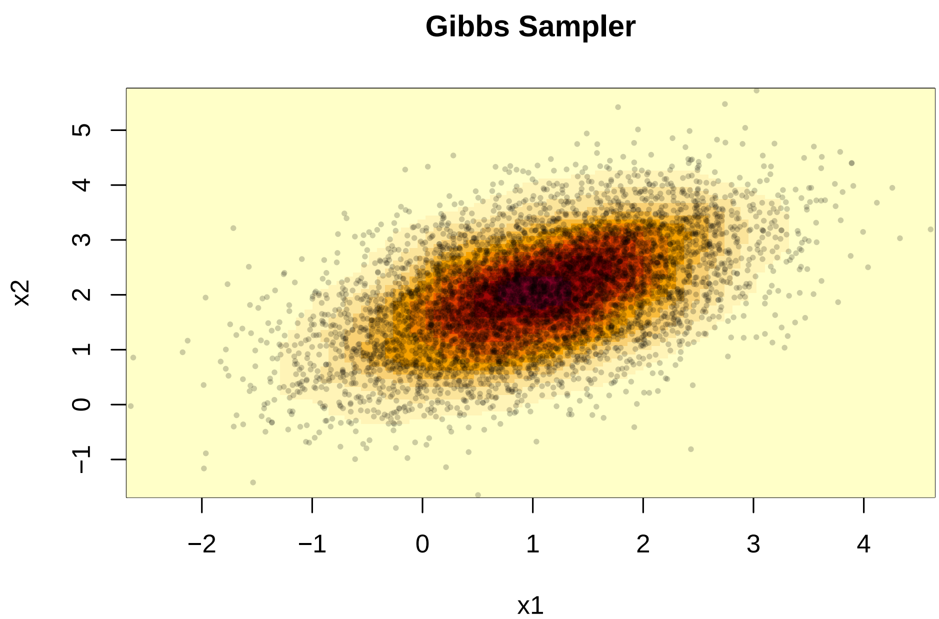
```
set.seed(1)

mu <- c(1, 2)
Sig <- matrix(c(1, .5,
                .5, 1), nrow = 2, byrow = T)
BVNs <- mvrnorm(5000, mu, Sig)

## plot
BVNs.kde <- kde2d(BVNs[, 1], BVNs[, 2], n = 100)
image(BVNs.kde, main = 'Bivariate Normal',
      xlab = 'x1', ylab = 'x2')
contour(BVNs.kde, add = T)
```



Then 5000 Gibbs samples are drawn, with the initial value $\boldsymbol{x}_0 = \boldsymbol{\mu}$. The code for Gibbs sampling are as follows. The contours of kernel density is overlaid by the samples from the Gibbs sampler.

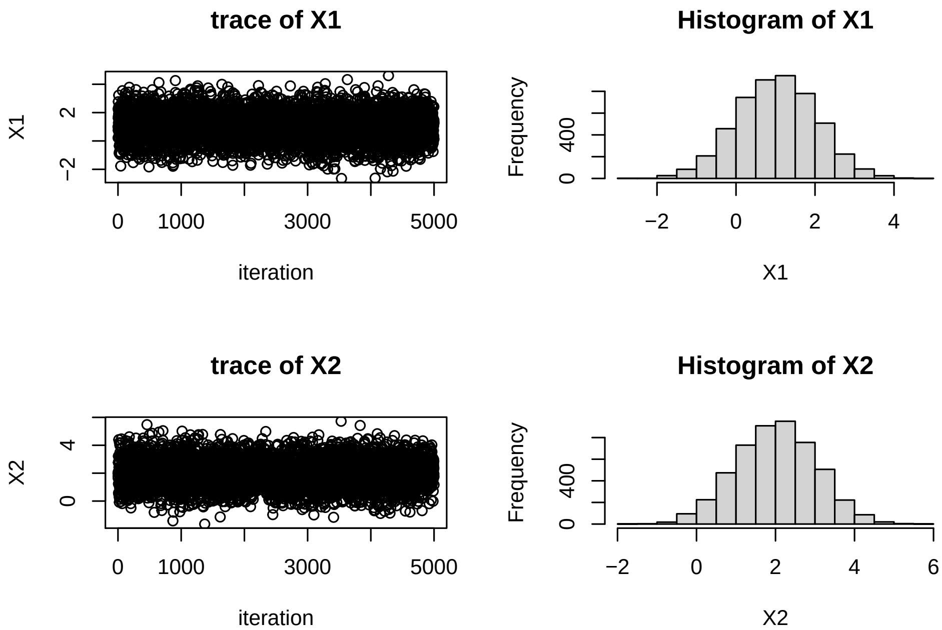
```r
GS_BVN <- function(mu, Sig, nGS, x0){
  X <- matrix(NA, nrow = nGS, ncol = 2)
  X[1, ] <- x0
  for(i in 2:nGS){

    ## update x1
    mu_tmp1 <- mu[1] + Sig[1, 2] %*% solve(Sig[2, 2]) %*% (X[i-1, 2] - mu[2])
    sig_tmp1 <- Sig[1, 1] - Sig[1, 2] %*% solve(Sig[2, 2]) %*% Sig[2, 1]
    X[i, 1] <- rnorm(1, mu_tmp1, sqrt(sig_tmp1))

    ## update x2
    mu_tmp2 <- mu[2] + Sig[2, 1] %*% solve(Sig[1, 1]) %*% (X[i, 1] - mu[1])
    sig_tmp2 <- Sig[2, 2] - Sig[2, 1] %*% solve(Sig[1, 1]) %*% Sig[1, 2]
    X[i, 2] <- rnorm(1, mu_tmp2, sqrt(sig_tmp2))
  }
  return(X)
}

set.seed(2)
x0 <- mu
nGS <- 5000
X <- GS_BVN(mu, Sig, nGS, x0)

## plot
X.kde <- kde2d(X[, 1], X[, 2], n = 100)
image(X.kde, main = 'Gibbs Sampler',
      xlab = 'x1', ylab = 'x2')
lines(X[, 1], X[, 2], type = 'p',
```

```
    col = alpha(1, 0.2), pch = 16, cex = 0.5)
```



We can see the sampling distribution from the Gibbs sampler matches the true distribution well. Further, the 1-D sample traces and histograms are shown below:

```
par(mfrow = c(2, 2))
plot(X[, 1], main = 'trace of X1',
     xlab = 'iteration', ylab = 'X1')
hist(X[, 1], xlab = 'X1', main = 'Histogram of X1')
plot(X[, 2], main = 'trace of X2',
     xlab = 'iteration', ylab = 'X2')
hist(X[, 2], xlab = 'X2', main = 'Histogram of X2')
```




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

The trace plots show that there are no heavy auto-correlation issues, and the 1D histograms show that the marginal distributions are normal.

2. Body Temperature

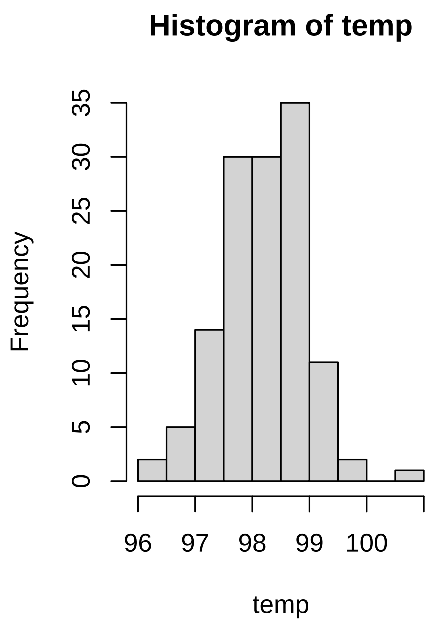
Currently, we focus on temperature only.

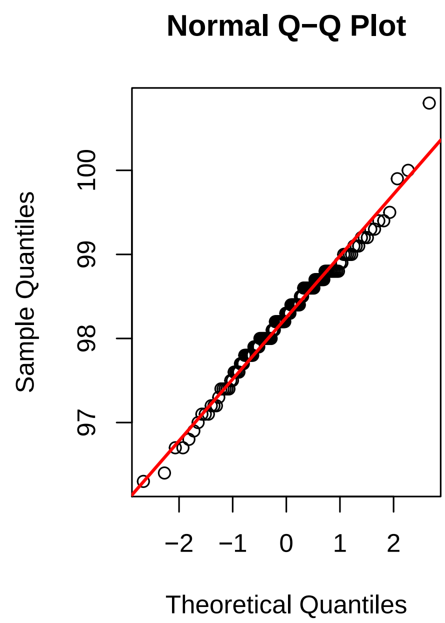
```r
d2 <- read.table('bodytemp.txt')
temp <- d2$temperature
```

2.1 Exploratory Analysis

The histogram and normal Q-Q plot are show below.

```r
par(mfrow = c(1, 2))
hist(temp)
qqnorm(temp)
abline(a = mean(temp), b= sd(temp),
       col = 'red', lwd = 2)
```






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

We can see that the distribution of body temperature is some what normal. The p-value for Shapiro-Wilk test is 0.23, which is larger than 0.05. Therefore, we may assume that the body temperature follows a normal distribution.

```r
shapiro.test(temp)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  temp
## W = 0.98658, p-value = 0.2332
```

The mean and variance:

```r
temp.mean <- mean(temp)
temp.var <- var(temp)
temp.n <- length(temp)

cat('sample mean:', temp.mean, '\n')
```

```
## sample mean: 98.24923
```

```r
cat('sample variance:', temp.var)
```

```
## sample variance: 0.5375575
```

2.2 Gibbs Sampler

Denote $X_i \overset{i.i.d.}{\sim} N(\theta, \sigma^2)$, and $\boldsymbol{X} = (x_1, ..., x_n)'$. Here, we use independent priors for θ and σ^2 as follows:

$$\theta \sim N(\mu_0, \tau_0^2)$$
$$\tilde{\sigma}^2 = 1/\sigma^2 \sim Gamma(\alpha_0, \beta_0)$$

Here, $\tilde{\sigma}^2 = 1/\sigma^2$ is the "precision". Therefore, we can get the full conditional distributions for θ and $\tilde{\sigma}^2$:

$$\begin{aligned}
p(\theta|\tilde{\sigma}^2, \boldsymbol{x}) &\propto p(\boldsymbol{x}|\theta, \tilde{\sigma}^2)p(\theta|\tilde{\sigma}^2)p(\tilde{\sigma}^2) \\
&\propto p(\boldsymbol{x}|\theta, \tilde{\sigma}^2)p(\theta) \\
&\propto N(\frac{\mu_0/\tau_0^2 + n\boldsymbol{x}\tilde{\sigma}^2}{1/\tau_0^2 + n\tilde{\sigma}^2}, (1/\tau_0^2 + n\tilde{\sigma}^2)^{-1}) \\
p(\tilde{\sigma}^2|\theta, \boldsymbol{x}) &\propto p(\boldsymbol{x}|\theta, \tilde{\sigma}^2)p(\tilde{\sigma}^2|\theta)p(\theta) \\
&\propto p(\boldsymbol{x}|\theta, \tilde{\sigma}^2)p(\tilde{\sigma}^2) \\
&\propto Gamma(\alpha_0 + \frac{n}{2}, \beta_0 + \frac{\sum_{i=1}^n (x_i - \theta)^2}{2})
\end{aligned}$$

To help with interpretation, rewrite $\alpha_0 = \frac{\nu_0}{2}$ and $\beta_0 = \frac{\nu_0\sigma_0^2}{2}$. Here, ν_0 can be viewed as the prior sample size, and σ_0^2 can be viewed as the prior sample variance. Further, since $\sum_{i=1}^n (x_i - \theta)^2 = \sum_{i=1}^n (x_i - \bar{\boldsymbol{x}})^2 + \sum_{i=1}^n (\bar{\boldsymbol{x}} - \theta)^2 = (n-1)s^2 + n(\bar{\boldsymbol{x}} - \theta)^2$, where s^2 and $\bar{\boldsymbol{x}}$ are sample variance and mean. So, the full conditional distribution for $\tilde{\sigma}^2$ can be written as:

$$p(\tilde{\sigma}^2|\theta, \boldsymbol{x}) \propto Gamma(\frac{\nu_0 + n}{2}, \frac{\nu_0\sigma_0^2 + (n-1)s^2 + n(\bar{\boldsymbol{x}} - \theta)^2}{2})$$

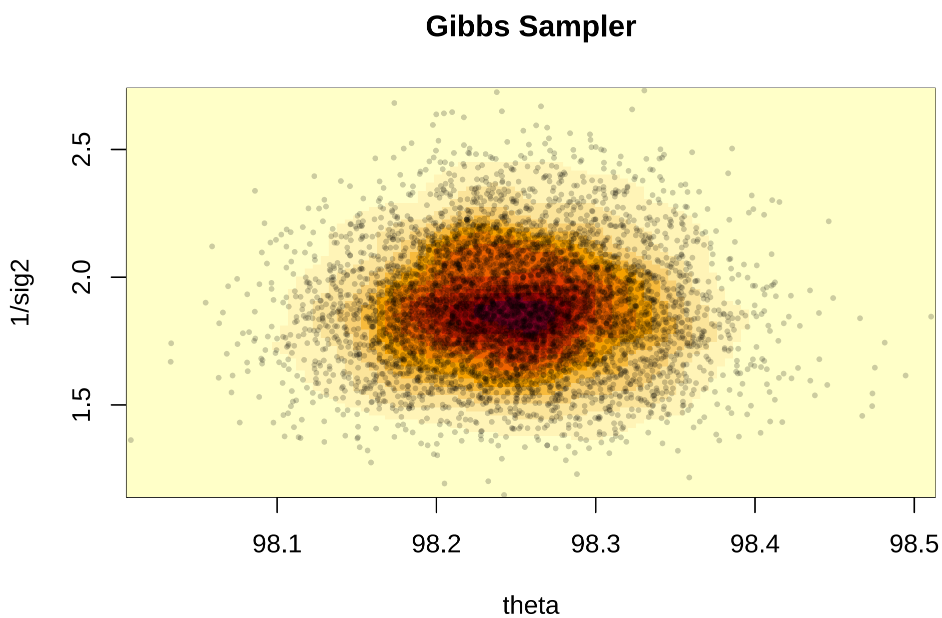
OK, let's begin to do the Gibbs sampler! Set the prior parameters as $\mu_0 = \bar{\boldsymbol{x}}, \tau_0^2 = 1, \nu_0 = 1, \sigma_0^2 = 0.01$. Draw 5000 samples, with initial values be $\theta_0 = \bar{\boldsymbol{x}}$ and $\tilde{\sigma}^2 = 1/s^2$. The code are as follows. In the code, I further denote $\phi = (\theta, \tilde{\sigma}^2)'$

```
nGS <- 5000
PHI <- matrix(NA, nrow = nGS, ncol = 2)

## prior
mu0 <- temp.mean
tau20 <- 1
nu0 <- 1
sig20 <- 0.01

## initialization
PHI[1, ] <- c(temp.mean, 1/temp.var)

## GS
set.seed(3)
for(i in 2:nGS){

  ## update theta
  mun <- (mu0/tau20 + temp.n*temp.mean*PHI[i-1, 2])/
    (1/tau20 + temp.n*PHI[i-1, 2])
  tau2n <- 1/(1/tau20 + temp.n*PHI[i-1, 2])
  PHI[i, 1] <- rnorm(1, mun, sqrt(tau2n))

  ## update sig2
```

```
  alphn <- (nu0 + temp.n)/2
  betan <- (nu0*sig20 + (temp.n-1)*temp.var + temp.n*(temp.mean - PHI[i, 1])^2)/2
  PHI[i, 2] <- rgamma(1, alphn, betan)
}
```

Again, let's first see the samples from the posterior distribution by the Gibbs sampler. The contours of kernel density is overlaid by the samples from the Gibbs sampler.

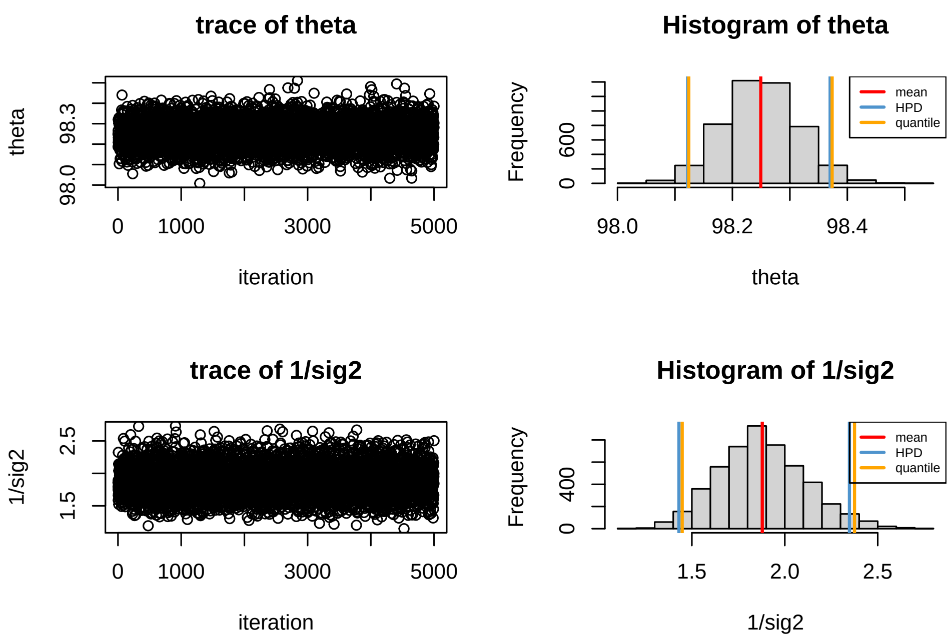
```
PHI.kde <- kde2d(PHI[, 1], PHI[, 2], n = 100)
image(PHI.kde, main = 'Gibbs Sampler',
      xlab = 'theta', ylab = '1/sig2')
lines(PHI[, 1], PHI[, 2], type = 'p',
      col = alpha(1, 0.2), pch = 16, cex = 0.5)
```



The traces and histograms are shown below. Gibbs samples mean, 95% HPD 0.25 quantile and 0.975 quantile are overlaid in the histograms.

```
par(mfrow = c(2, 2))
plot(PHI[, 1], main = 'trace of theta',
     xlab = 'iteration', ylab = 'theta')
hist(PHI[, 1], xlab = 'theta',
     main = 'Histogram of theta')
theta.hdi <- hdi(PHI[, 1], credMass = 0.95)
theta.quantile <- quantile(PHI[, 1], c(0.025, 0.975))

abline(v = mean(PHI[, 1]), lwd = 2, col = 'red')
abline(v = theta.hdi[1], lwd = 2, col = 'steelblue3')
abline(v = theta.hdi[2], lwd = 2, col = 'steelblue3')
abline(v = theta.quantile[1], lwd = 2, col = 'orange')
```

```r
abline(v = theta.quantile[2], lwd = 2, col = 'orange')
legend('topright', legend = c('mean', 'HPD', 'quantile'),
       lwd = 2, col = c('red', 'steelblue3', 'orange'), cex = 0.6)

plot(PHI[, 2], main = 'trace of 1/sig2',
     xlab = 'iteration', ylab = '1/sig2')
hist(PHI[, 2], xlab = '1/sig2',
     main = 'Histogram of 1/sig2')
invSig2.hdi <- hdi(PHI[, 2], credMass = 0.95)
invSig2.quantile <- quantile(PHI[, 2], c(0.025, 0.975))

abline(v = mean(PHI[, 2]), lwd = 2, col = 'red')
abline(v = invSig2.hdi[1], lwd = 2, col = 'steelblue3')
abline(v = invSig2.hdi[2], lwd = 2, col = 'steelblue3')
abline(v = invSig2.quantile[1], lwd = 2, col = 'orange')
abline(v = invSig2.quantile[2], lwd = 2, col = 'orange')
legend('topright', legend = c('mean', 'HPD', 'quantile'),
       lwd = 2, col = c('red', 'steelblue3', 'orange'), cex = 0.6)
```




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

The trace plots show that there are no heavy auto-correlation issues. The histograms show that the posterior distributions are somewhat symmetric for both parameters.

The posterior mean, 95% HPD and 95% symmetric credible interval for θ:

```
## posterior mean: 98.24945
```

```
## 95% HPD: [ 98.12213 , 98.3701 ]
```

```
## 95% symmetric credible interval: [ 98.12409 , 98.37353 ]
```

While the posterior mean, 95% HPD and 95% symmetric credible interval for $\tilde{\sigma}^2$:

```
## posterior mean: 1.878633
```

```
## 95% HPD: [ 1.430734 , 2.34759 ]
```

```
## 95% symmetric credible interval: [ 1.447968 , 2.37505 ]
```

Further, the values for variance σ^2:

```
## posterior mean: 0.540682
```

```
## 95% HPD: [ 0.4120304 , 0.676463 ]
```

```
## 95% symmetric credible interval: [ 0.4210437 , 0.690623 ]
```

The HPD and symmetric credible intervals are close in both parameters.

Also, notice that the observation mean & precision are:

```
## sample mean: 98.24923
```

```
## sample precision: 1.860266
```

2.3 Conclusion

From the above analysis, the 95% HPD for the posterior mean is [98.122, 98.370] and 95% HPD for the posterior variance is [0.412, 0.676]. Since 98.6 is not in 95% HPD of the mean, the normal body temperature is not 98.6.