# HW3 Solutions
STAT431 Spr19

February 2019

## Contents



# 1 Conjugate joint prior density

## 1.1 Values of parameters in Normal-inverse gamma prior

The four parameters in the Normal Inverse-gamma prior are $\mu_0, \kappa, \alpha, \beta$. For the conditional normal part, we want a prior mean of 0 and an equivalent prior sample size of 1, so we set $\mu_0 = 0$ and $\kappa = 1$ for the conditional normal part.

For the inverse-gamma portion, we want $\sigma^2$ to marginally have prior mean $\frac{\beta}{\alpha-1} = 1$ and prior variance $\frac{\beta^2}{(\alpha-1)^2(\alpha-2)} = 1$. We then solve for the values of $\alpha, \beta$.

$$
\begin{aligned}
\frac{\beta}{\alpha - 1} = 1 \Rightarrow \frac{\beta^2}{(\alpha-1)^2} &= 1 \qquad (1)\\
\Rightarrow \frac{1}{(\alpha - 2)} &= 1 \qquad (2)\\
\Rightarrow \alpha &= 3 \qquad (3)\\
\Rightarrow \beta &= 2 \qquad (4)
\end{aligned}
$$

**We set $\alpha = 3$ and $\beta = 2$ for the inverse-gamma parameters, and we set $\mu_0 = 0$ and $\kappa = 1$ for the conditional normal part.**

## 1.2 Marginal posterior for $\mu$

The marginal posterior density for $\mu$ is the $t$ distribution, with mean 4.6188, scale parameter 0.1868, and degrees of freedom 16.

```
n = length(Y)
my_kappa = 1
mu_0 = 0
my_alpha = 3
my_beta = 2

n_0 = my_alpha * 2
sigmasq_0 = my_beta * 2 / n_0

# Mean
post_mean = (my_kappa*mu_0 + n*mean(Y)) / (my_kappa + n)

# Scale
post_scale = (n_0*sigmasq_0 + (n-1)*var(Y) + (my_kappa*n*(mean(Y) - mu_0)^2)/(my_kappa +
  ((my_kappa + n)*(n_0 + n))

# Degrees of Freedom
post_df = n + n_0
```

## 1.3 95% Credible interval for $\mu$

The posterior 95% equal-tailed credible interval for $\mu$ is equal to $(3.702, 5.535)$.

```
(post_mean + qt(c(0.025, 0.975), post_df) * sqrt(post_scale)) %>% round(3)
```

## 1.4 Marginal posterior for $\sigma^2$

The marginal posterior distribution for $\sigma^2$ is Inverse-gamma, with parameters shape=8, scale=16.445.

```
# Shape
my_alpha + n/2

# Scale
my_beta + (n-1)/2*var(Y) + (my_kappa*n*(mean(Y) - mu_0)^2) / (2*(my_kappa+n))
```

### 1.5 95% Credible interval for $\sigma^2$

The posterior 95% equal-tailed credible interval for $\sigma^2$ is equal to (1.1402, 4.761).

```
1 / qgamma(c(0.975, 0.025), post_shape, post_scale)
```

## 2 Standard noninformative prior

### 2.1 Marginal posterior for $\mu$

The marginal posterior distribution for $\mu$ is the $t$-distribution, with parameters mean $\overline{Y} = 5.0807$, scale parameter $\frac{s^2}{n} = 0.06025$, and degrees of freedom $n-1 = 9$.

```
post_mean = mean(Y)
post_scale = var(Y) / n
post_df = n-1
```

### 2.2 95% Credible interval for $\mu$

```
(post_mean + qt(c(0.025, 0.975), post_df) * sqrt(post_scale)) %>% round(3)
```

The posterior 95% equal-tailed credible interval for $\mu$ is equal to $(4.525, 5.636)$.

### 2.3 Marginal posterior for $\sigma^2$

The marginal posterior distribution for $\sigma^2$ is Inverse-gamma, with parameters shape=4.5, scale=2.711.

```
post_shape = (n-1) / 2
post_scale = (n-1)*var(Y) / 2
```

### 2.4 95% Credible interval for $\sigma^2$

The posterior 95% equal-tailed credible interval for $\sigma^2$ is equal to (0.285, 2.008).

```
1 / qgamma(c(0.975, 0.025), post_shape, post_scale)
```

# 3 GRADUATE ONLY - Deriving Jeffrey's prior

## 3.1 Information matrix derivations

$$lnL(\mu, \sigma^2 \mid Y) = -\frac{n}{2}ln(2\pi) - \frac{n}{2}ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \mu)^2 \tag{5}$$

$$\frac{d}{d\sigma^2}lnL(\mu, \sigma^2 \mid Y) = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(Y_i - \mu)^2 \tag{6}$$

$$\frac{d}{d\mu}\frac{d}{d\sigma^2}lnL(\mu, \sigma^2 \mid Y) = \frac{d}{d\mu}\left(\frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(Y_i^2 - 2\mu Y_i + \mu^2)\right) \tag{7}$$

$$= \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(-2Y_i + 2\mu) \tag{8}$$

$$\mathbf{E}\left[\frac{d^2}{d\mu d\sigma^2}lnL(\mu, \sigma^2 \mid Y)\right] = \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}\mathbf{E}\left(-2Y_i + 2\mu\right) \tag{9}$$

$$= 0 \tag{10}$$

The last line follows because $\mathbf{E}(Y_i) = \mu$ by definition.

## 3.2 Jeffrey's prior derivation

In part (a) we found that $I_{12}(\mu, \sigma^2) = 0$. Therefore the determinant of the information matrix is only the product of the diagonal elements $I_{11}, I_{22}$.

$$\det(I) = I_{11}I_{22} = \frac{n}{\sigma^2}\frac{n}{2(\sigma^2)^2} = \frac{n^2}{2(\sigma^2)^3} \tag{11}$$

Therefore the Jeffreys prior density is proportional to $\sqrt{\det(I)} = \frac{n}{\sqrt{2}(\sigma^2)^{3/2}}$