Chapter 6: Bayesian Inference for Gaussian Distribution

강의 목표

- 정규분포를 중심으로 베이지안 추론의 이해
- Parameter Estimation (모수 추정)
  - Point Estimation (점추정)
  - Confidence Interval (구간추정)
- Prediction (예측)

Why Normal Models?

- Approximately normally distributed quantities appear often in nature.

Why Normal Models?

- Approximately normally distributed quantities appear often in nature.
- CLT tells us any variable that is basically a sum of independent components should be approximately normal.

Why Normal Models?

- Approximately normally distributed quantities appear often in nature.
- CLT tells us any variable that is basically a sum of independent components should be approximately normal.
- Note $\bar{X}$ and S^2 are independent when sampling from a normal population - so if beliefs about the mean are independent of beliefs about the variance, a normal model may be appropriate.

Why Normal Models?

- The normal model is analytically convenient.

Why Normal Models?

- The normal model is analytically convenient.
- Inference about the population mean based on a normal model will be correct as $n \to \infty$ even if the data are truly non-normal.

Conjugate analysis with Normal Data (known variance)

- Assume data $X_1, ..., X_n$ are iid $N(\mu, \sigma^2)$, with μ unknown and σ^2 known.

Conjugate analysis with Normal Data (known variance)

- Assume data $X_1, ..., X_n$ are iid $N(\mu, \sigma^2)$, with μ unknown and σ^2 known.
- We will make inference about μ.

Conjugate analysis with Normal Data (known variance)

- Assume data $X_1, ..., X_n$ are iid $N(\mu, \sigma^2)$, with μ unknown and σ^2 known.
- We will make inference about μ.
- The likelihood is following:

Conjugate analysis with Normal Data (known variance)

- Assume data $X_1, ..., X_n$ are iid $N(\mu, \sigma^2)$, with μ unknown and σ^2 known.
- We will make inference about μ.
- The likelihood is following:

$$L(\mu \mid X) = \prod_{i=1}^{n} (2\pi\sigma^2)^{-1/2} e^{-\frac{1}{2\sigma^2}(x_i - \mu)^2}$$

Conjugate analysis with Normal Data (known variance)

- Assume data $X_1, ..., X_n$ are iid $N(\mu, \sigma^2)$, with μ unknown and σ^2 known.
- We will make inference about μ.
- The likelihood is following:

$$L(\mu \mid X) = \prod_{i=1}^{n} (2\pi\sigma^2)^{-1/2} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2}$$

- A conjugate prior for μ is $\mu \sim N(\delta, \tau^2)$:

Conjugate analysis with Normal Data (known variance)

- Assume data $X_1, ..., X_n$ are iid $N(\mu, \sigma^2)$, with μ unknown and σ^2 known.
- We will make inference about μ.
- The likelihood is following:

$$L(\mu \mid X) = \prod_{i=1}^{n} (2\pi\sigma^2)^{-1/2} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2}$$

- A conjugate prior for μ is $\mu \sim N(\delta, \tau^2)$:

$$p(\mu) = \frac{1}{\sqrt{(2\pi\tau^2)}} e^{-\frac{1}{2\tau^2}(\mu-\delta)^2}.$$

Conjugate analysis with Normal Data (known variance)

- The posterior is:

$$\pi(\mu \mid x) \quad \propto \quad L(\mu \mid x)p(\mu)$$

Conjugate analysis with Normal Data (known variance)

- The posterior is:

$$
\begin{aligned}
\pi(\mu \mid x) &\propto L(\mu \mid x)p(\mu) \\
&\propto \prod_{i=1}^{n} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2} e^{-\frac{1}{2\tau^2}(\mu-\delta)^2}
\end{aligned}
$$

Conjugate analysis with Normal Data (known variance)

- The posterior is:

$$\begin{aligned}
\pi(\mu \mid x) &\propto L(\mu \mid x) p(\mu) \\
&\propto \prod_{i=1}^{n} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2} e^{-\frac{1}{2\tau^2}(\mu-\delta)^2} \\
&= \exp\left\{-\frac{1}{2}\left(\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 + \frac{1}{\tau^2}(\mu-\delta)^2\right)\right\}
\end{aligned}$$

Conjugate analysis with Normal Data (known variance)

- The posterior is:

$$
\begin{aligned}
\pi(\mu \mid x) &\propto L(\mu \mid x) p(\mu) \\
&\propto \prod_{i=1}^{n} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2} e^{-\frac{1}{2\tau^2}(\mu-\delta)^2} \\
&= \exp\left\{-\frac{1}{2}\left(\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 + \frac{1}{\tau^2}(\mu-\delta)^2\right)\right\} \\
&= \exp\left\{-\frac{1}{2}\left(\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i^2 - 2x_i\mu + \mu^2) + \frac{1}{\tau^2}(\mu^2 - 2\mu\delta + \delta^2)\right)\right\}
\end{aligned}
$$

Conjugate analysis with Normal Data (known variance)

$$\propto \quad \exp\left\{-\frac{1}{2\sigma^2\tau^2}\left(\tau^2\sum x_i^2 - 2\tau^2\mu n\bar{x} + n\mu^2\tau^2 + (\sigma^2\mu^2 - 2\sigma^2\mu\delta + \sigma^2\delta^2)\right)\right\}$$

Conjugate analysis with Normal Data (known variance)

$$
\begin{aligned}
\propto \quad & \exp\left\{-\frac{1}{2\sigma^2\tau^2}\left(\tau^2\sum x_i^2 - 2\tau^2\mu n\bar{x} + n\mu^2\tau^2 + (\sigma^2\mu^2 - 2\sigma^2\mu\delta + \sigma^2\delta^2)\right)\right\} \\
= \quad & \exp\left\{-\frac{1}{2\sigma^2\tau^2}\left(\mu^2(\sigma^2 + n\tau^2) - 2\mu(\delta\sigma^2 + \tau^2 n\bar{x}) + (\sigma^2\delta^2\tau^2\sum x_i^2)\right)\right\}
\end{aligned}
$$

Conjugate analysis with Normal Data (known variance)

$$
\begin{aligned}
&\propto \exp\left\{-\frac{1}{2\sigma^2\tau^2}\left(\tau^2\sum x_i^2 - 2\tau^2\mu n\bar{x} + n\mu^2\tau^2 + (\sigma^2\mu^2 - 2\sigma^2\mu\delta + \sigma^2\delta^2)\right)\right\} \\
&= \exp\left\{-\frac{1}{2\sigma^2\tau^2}\left(\mu^2(\sigma^2 + n\tau^2) - 2\mu(\delta\sigma^2 + \tau^2 n\bar{x}) + (\sigma^2\delta^2\tau^2\sum x_i^2)\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\mu^2\left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right) - 2\mu\left(\frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2} + k\right)\right\}
\end{aligned}
$$

Conjugate analysis with Normal Data (known variance)

$$
\begin{aligned}
&\propto \exp\left\{-\frac{1}{2\sigma^2\tau^2}\left(\tau^2\sum x_i^2 - 2\tau^2\mu n\bar{x} + n\mu^2\tau^2 + (\sigma^2\mu^2 - 2\sigma^2\mu\delta + \sigma^2\delta^2)\right)\right\} \\
&= \exp\left\{-\frac{1}{2\sigma^2\tau^2}\left(\mu^2(\sigma^2 + n\tau^2) - 2\mu(\delta\sigma^2 + \tau^2 n\bar{x}) + (\sigma^2\delta^2\tau^2\sum x_i^2)\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\mu^2\left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right) - 2\mu\left(\frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2} + k\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)\left(\mu^2 - 2\mu\left(\frac{\frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}}\right) + k\right)\right\}
\end{aligned}
$$

Conjugate analysis with Normal Data (known variance)

$$
\begin{aligned}
&\propto \exp\left\{-\frac{1}{2\sigma^2\tau^2}\left(\tau^2\sum x_i^2 - 2\tau^2\mu n\bar{x} + n\mu^2\tau^2 + (\sigma^2\mu^2 - 2\sigma^2\mu\delta + \sigma^2\delta^2)\right)\right\} \\
&= \exp\left\{-\frac{1}{2\sigma^2\tau^2}\left(\mu^2(\sigma^2 + n\tau^2) - 2\mu(\delta\sigma^2 + \tau^2 n\bar{x}) + (\sigma^2\delta^2\tau^2\sum x_i^2)\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\mu^2\left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right) - 2\mu\left(\frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2} + k\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)\left(\mu^2 - 2\mu\left(\frac{\frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}}\right) + k\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)\left(\mu - \frac{\frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}}\right)^2\right\}
\end{aligned}
$$

where k is a constant.

Conjugate analysis with Normal Data (known variance)

- The posterior expectation:

Conjugate analysis with Normal Data (known variance)

- The posterior expectation:

$$\frac{\frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}}$$

Conjugate analysis with Normal Data (known variance)

- The posterior expectation:

$$\frac{\frac{\delta}{\tau^2}+\frac{n\bar{x}}{\sigma^2}}{\frac{1}{\tau^2}+\frac{n}{\sigma^2}} = \frac{\sigma^2}{\sigma^2+n\tau^2}\delta + \frac{n\tau^2}{\sigma^2+n\tau^2}\bar{x}.$$

Conjugate analysis with Normal Data (known variance)

- The posterior expectation:

$$\frac{\frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}} = \frac{\sigma^2}{\sigma^2 + n\tau^2}\delta + \frac{n\tau^2}{\sigma^2 + n\tau^2}\bar{x}.$$

- The posterior variance:

$$\left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{-1}$$

Conjugate analysis with Normal Data (known variance)

- The posterior expectation:

$$\frac{\frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}} = \frac{\sigma^2}{\sigma^2 + n\tau^2}\delta + \frac{n\tau^2}{\sigma^2 + n\tau^2}\bar{x}.$$

- The posterior variance:

$$\left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{-1}$$

- The precision is the reciprocal of the variance.

Conjugate analysis with Normal Data (known variance)

- The posterior expectation:

$$\frac{\frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}} = \frac{\sigma^2}{\sigma^2 + n\tau^2}\delta + \frac{n\tau^2}{\sigma^2 + n\tau^2}\bar{x}.$$

- The posterior variance:

$$\left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{-1}$$

- The precision is the reciprocal of the variance.
- Prior precision: $\frac{1}{\tau^2}$.

Conjugate analysis with Normal Data (known variance)

- The posterior expectation:

$$\frac{\frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}} = \frac{\sigma^2}{\sigma^2 + n\tau^2}\delta + \frac{n\tau^2}{\sigma^2 + n\tau^2}\bar{x}.$$

- The posterior variance:

$$\left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{-1}$$

- The precision is the reciprocal of the variance.
- Prior precision: $\frac{1}{\tau^2}$.
- Data precision: $\frac{n}{\sigma^2}$.

Conjugate analysis with Normal Data (known variance)

- The posterior expectation:

$$\frac{\frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}} = \frac{\sigma^2}{\sigma^2 + n\tau^2}\delta + \frac{n\tau^2}{\sigma^2 + n\tau^2}\bar{x}.$$

- The posterior variance:

$$\left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{-1}$$

- The precision is the reciprocal of the variance.
- Prior precision: $\frac{1}{\tau^2}$.
- Data precision: $\frac{n}{\sigma^2}$.
- Posterior precision: $\frac{1}{\tau^2} + \frac{n}{\sigma^2}$.

Conjugate analysis with Normal Data (known variance)

- The posterior mean $E(\mu \mid x)$:

Conjugate analysis with Normal Data (known variance)

- The posterior mean $E(\mu \mid x)$:

$$\frac{\frac{1}{\tau^2}}{\frac{1}{\tau^2}+\frac{n}{\sigma^2}}\delta+\frac{\frac{n}{\sigma^2}}{\frac{1}{\tau^2}+\frac{n}{\sigma^2}}\bar{x}$$

Conjugate analysis with Normal Data (known variance)

- The posterior mean $E(\mu \mid x)$:

$$\frac{\frac{1}{\tau^2}}{\frac{1}{\tau^2}+\frac{n}{\sigma^2}}\delta + \frac{\frac{n}{\sigma^2}}{\frac{1}{\tau^2}+\frac{n}{\sigma^2}}\bar{x}$$

- The mean is a weighted average of the **prior mean** and the **sample mean**.

Conjugate analysis with Normal Data (known variance)

- The posterior mean $E(\mu \mid x)$:

$$\frac{\frac{1}{\tau^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}}\delta + \frac{\frac{n}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}}\bar{x}$$

- The mean is a weighted average of the **prior mean** and the **sample mean**.
- If the prior is highly precise, the weight is large on δ.

Conjugate analysis with Normal Data (known variance)

- The posterior mean $E(\mu \mid x)$:

$$\frac{\frac{1}{\tau^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}}\delta + \frac{\frac{n}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}}\bar{x}$$

- The mean is a weighted average of the **prior mean** and the **sample mean**.
- If the prior is highly precise, the weight is large on δ.
- If the data are highly precise (e.g., when n is large), the weight is large on $\bar{x}$.

Conjugate analysis with Normal Data (known variance)

- The posterior mean $E(\mu \mid x)$:

$$\frac{\frac{1}{\tau^2}}{\frac{1}{\tau^2}+\frac{n}{\sigma^2}}\delta + \frac{\frac{n}{\sigma^2}}{\frac{1}{\tau^2}+\frac{n}{\sigma^2}}\bar{x}$$

- The mean is a weighted average of the **prior mean** and the **sample mean**.
- If the prior is highly precise, the weight is large on δ.
- If the data are highly precise (e.g., when n is large), the weight is large on $\bar{x}$.
- Clearly as $n \to \infty$, $E(\mu \mid x) \asymp \bar{x}$ and **var**$(\mu \mid x) \asymp \frac{\sigma^2}{n}$ if we choose a large prior variance τ^2.

Conjugate analysis with Normal Data (known variance)

- The posterior mean $E(\mu \mid x)$:

Conjugate analysis with Normal Data (known variance)

- The posterior mean $E(\mu \mid x)$:

$$\frac{\frac{1}{\tau^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}}\delta + \frac{\frac{n}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}}\bar{x}$$

Conjugate analysis with Normal Data (known variance)

- The posterior mean $E(\mu \mid x)$:

$$\frac{\frac{1}{\tau^2}}{\frac{1}{\tau^2}+\frac{n}{\sigma^2}}\delta + \frac{\frac{n}{\sigma^2}}{\frac{1}{\tau^2}+\frac{n}{\sigma^2}}\bar{x}$$

- This implies that for τ^2 large and n large, Bayesian and Frequentist inference about μ will be nearly identical.

Example for Normal Data (known variance)

과거의 자료로부터 $\theta \sim N(370, 21^2)$ 사전분포를 가정하고,

$\bar{X}\bar{\theta} \sim N(\theta, 8^2)$ 분포를 따르는 $\bar{X}$의 관측치는 $x = 421$이었다.

Example for Normal Data (known variance)

과거의 자료로부터 $\theta \sim N(370, 21^2)$ 사전분포를 가정하고, $\bar{X}\bar{\theta} \sim N(\theta, 8^2)$ 분포를 따르는 $\bar{X}$의 관측치는 $x = 421$이었다. 이 경우 θ의 사후분포는 평균과 분산으로 각각

Example for Normal Data (known variance)

과거의 자료로부터 $\theta \sim N(370, 21^2)$ 사전분포를 가정하고, $\bar{X}\bar{\theta} \sim N(\theta, 8^2)$ 분포를 따르는 $\bar{X}$의 관측치는 $x = 421$이었다. 이 경우 θ의 사후분포는 평균과 분산으로 각각

$$\mu_\pi \quad = \quad \frac{1/8^2 \times 421 + 1/21^2 \times 370}{1/8^2 + 1/21^2} = 413$$

Example for Normal Data (known variance)

과거의 자료로부터 $\theta \sim N(370, 21^2)$ 사전분포를 가정하고, $\bar{X}\bar{\theta} \sim N(\theta, 8^2)$ 분포를 따르는 $\bar{X}$의 관측치는 $x = 421$이었다. 이 경우 θ의 사후분포는 평균과 분산으로 각각

$$\mu_\pi = \frac{1/8^2 \times 421 + 1/21^2 \times 370}{1/8^2 + 1/21^2} = 413$$

$$\sigma_\pi^2 = \frac{1}{1/8^2 + 1/21^2} = 7^2.$$

을 갖는 정규분포이다.

Example for Normal Data (known variance)

과거의 자료로부터 $\theta \sim N(370, 21^2)$ 사전분포를 가정하고, $\bar{X}\bar{\theta} \sim N(\theta, 8^2)$ 분포를 따르는 $\bar{X}$의 관측치는 $x = 421$이었다. 이 경우 θ의 사후분포는 평균과 분산으로 각각

$$
\begin{aligned}
\mu_\pi &= \frac{1/8^2 \times 421 + 1/21^2 \times 370}{1/8^2 + 1/21^2} = 413 \\
\sigma_\pi^2 &= \frac{1}{1/8^2 + 1/21^2} = 7^2.
\end{aligned}
$$

을 갖는 정규분포이다. 정규분포는 symmetric이므로 95% HPD interval은

Example for Normal Data (known variance)

과거의 자료로부터 $\theta \sim N(370, 21^2)$ 사전분포를 가정하고, $\bar{X}\bar{\theta} \sim N(\theta, 8^2)$ 분포를 따르는 $\bar{X}$의 관측치는 $x = 421$이었다. 이 경우 θ의 사후분포는 평균과 분산으로 각각

$$
\begin{aligned}
\mu_\pi &= \frac{1/8^2 \times 421 + 1/21^2 \times 370}{1/8^2 + 1/21^2} = 413 \\
\sigma_\pi^2 &= \frac{1}{1/8^2 + 1/21^2} = 7^2.
\end{aligned}
$$

을 갖는 정규분포이다. 정규분포는 symmetric이므로 95% HPD interval은

$$413 \pm 1.96 \times 7.$$

Example for Normal Data (known variance)

과거의 자료로부터 $\theta \sim N(370, 50^2)$ 사전분포를 가정하고,

$\bar{X} \mid \theta \sim N(\theta, 8^2)$ 분포를 따르는 $\bar{X}$의 관측치는 $x = 421$이었다.

Example for Normal Data (known variance)

과거의 자료로부터 $\theta \sim N(370, 50^2)$ 사전분포를 가정하고, $\bar{X} \mid \theta \sim N(\theta, 8^2)$ 분포를 따르는 $\bar{X}$의 관측치는 $x = 421$이었다. 이 경우 θ의 사후분포는 평균과 분산으로 각각

Example for Normal Data (known variance)

과거의 자료로부터 $\theta \sim N(370, 50^2)$ 사전분포를 가정하고,

$\bar{X} \mid \theta \sim N(\theta, 8^2)$ 분포를 따르는 $\bar{X}$의 관측치는 $x = 421$이었다. 이 경우 θ의 사후분포는 평균과 분산으로 각각

$$\mu_\pi = \frac{1/8^2 \times 421 + 1/50^2 \times 370}{1/8^2 + 1/50^2} = 419$$

Example for Normal Data (known variance)

과거의 자료로부터 $\theta \sim N(370, 50^2)$ 사전분포를 가정하고, $\bar{X} \mid \theta \sim N(\theta, 8^2)$ 분포를 따르는 $\bar{X}$의 관측치는 $x = 421$이었다. 이 경우 θ의 사후분포는 평균과 분산으로 각각

$$
\begin{aligned}
\mu_\pi &= \frac{1/8^2 \times 421 + 1/50^2 \times 370}{1/8^2 + 1/50^2} = 419 \\
\sigma_\pi^2 &= \frac{1}{1/8^2 + 1/50^2} = 8^2.
\end{aligned}
$$

Example for Normal Data (known variance)

과거의 자료로부터 $\theta \sim N(370, 50^2)$ 사전분포를 가정하고, $\bar{X} \mid \theta \sim N(\theta, 8^2)$ 분포를 따르는 $\bar{X}$의 관측치는 $x = 421$이었다. 이 경우 θ의 사후분포는 평균과 분산으로 각각

$$\begin{aligned} \mu_\pi &= \frac{1/8^2 \times 421 + 1/50^2 \times 370}{1/8^2 + 1/50^2} = 419 \\ \sigma_\pi^2 &= \frac{1}{1/8^2 + 1/50^2} = 8^2. \end{aligned}$$

을 갖는 정규분포이다.

Example for Normal Data (known variance)

과거의 자료로부터 $\theta \sim N(370, 50^2)$ 사전분포를 가정하고, $\bar{X} \mid \theta \sim N(\theta, 8^2)$ 분포를 따르는 $\bar{X}$의 관측치는 $x = 421$이었다. 이 경우 θ의 사후분포는 평균과 분산으로 각각

$$
\begin{aligned}
\mu_\pi &= \frac{1/8^2 \times 421 + 1/50^2 \times 370}{1/8^2 + 1/50^2} = 419 \\
\sigma_\pi^2 &= \frac{1}{1/8^2 + 1/50^2} = 8^2.
\end{aligned}
$$

을 갖는 정규분포이다. 이 경우 prior variance가 sample variance에 비해 매우 크므로, posterior mean이 sample mean에 더욱 가까워진다.

Prediction Distribution (known variance)

- Prediction Dist:

$$f(x_{n+1} \mid x_1, ..., x_n) = \int f(x_{n+1} \mid \mu, x_1, ..., x_n)\pi(\mu \mid x_1, ..., x_n)d\mu$$
$$= \int f(x_{n+1} \mid \mu)\pi(\mu \mid x_1, x_2, ..., x_n)d\mu$$

Prediction Distribution (known variance)

- Prediction Dist:

$$\begin{aligned} f(x_{n+1} \mid x_1, ..., x_n) &= \int f(x_{n+1} \mid \mu, x_1, ..., x_n)\pi(\mu \mid x_1, ..., x_n)d\mu \\ &= \int f(x_{n+1} \mid \mu)\pi(\mu \mid x_1, x_2, ..., x_n)d\mu \end{aligned}$$

- Since we know $f(x_{n+1} \mid \theta) \sim$ Normal distribution, and its posterior is also Normal,

Prediction Distribution (known variance)

- Prediction Dist:

$$
\begin{aligned}
f(x_{n+1} \mid x_1, \ldots, x_n) &= \int f(x_{n+1} \mid \mu, x_1, \ldots, x_n)\pi(\mu \mid x_1, \ldots, x_n)d\mu \\
&= \int f(x_{n+1} \mid \mu)\pi(\mu \mid x_1, x_2, \ldots, x_n)d\mu
\end{aligned}
$$

- Since we know $f(x_{n+1} \mid \theta) \sim$ Normal distribution, and its posterior is also Normal, the prediction distribution is Normal.

Prediction Distribution (known variance)

- Using the independent property, we have

$$\begin{aligned} X_{n+1} \mid x_1, ..., x_n &=_d X_{n+1} - \mu + \mu \mid x_1, ..., x_n \\ X_{n+1} - \mu \mid \mu, x_1, ..., x_n &\sim N(0, \sigma^2) \\ X_{n+1} \mid x_1, ..., x_n &\sim N(\mu_\pi, \sigma^2 + \sigma_\pi^2) \end{aligned}$$

Prediction Distribution (known variance)

- Using the independent property, we have

$$\begin{aligned} X_{n+1} \mid x_1, ..., x_n &=_d X_{n+1} - \mu + \mu \mid x_1, ..., x_n \\ X_{n+1} - \mu \mid \mu, x_1, ..., x_n &\sim N(0, \sigma^2) \\ X_{n+1} \mid x_1, ..., x_n &\sim N(\mu_\pi, \sigma^2 + \sigma_\pi^2) \end{aligned}$$

- Note that

$$\begin{aligned} P(X_{n+1} - \mu \mid B) &= \frac{\int_\mu P(X_{n+1}-\mu|\mu,B)P(\mu,B)d\mu}{P(B)} \\ &= P(X_{n+1} - \mu \mid \mu, B). \end{aligned}$$

Example for Normal Data (known variance)

우리나라 전체 가구의 월수입을 로그변환하면 자료의 분포가 대략적으로 대칭적인 정규분포 모양을 가진다. 따라서 가구 당 로그 월수입을 변수 X라 한다면 X의 분포를 정규분포로 가정해도 큰 무리가 없다. 우리의 관심은 어떤 특정 도시 A의 가구당 월평균 수입을 알아보는 것이다. 정규성 가정이 편리하므로 로그 월수입을 X라 하고, X의 평균 μ를 추정한 뒤 이를 지수 변환하여 원하는 가구당 월 평균 수입을 얻도록 하자.

Example for Normal Data (known variance)

θ 에 대한 사전정보를 얻기 위하여 알아보니 해당 연도 전체 한국 국민의 가구당 월수입은 평균이 294만 원, 표준편차가 200만 원이었다. 이를 로그변환하면 로그 월수입은 평균이 $\log 294 = 5.68$ 만 원, 표준편차는 $\log 200 = 5.30$이다. 이로부터 $N(5.68, 5.30^2)$를 θ의 사전분포로 사용한다.

Example for Normal Data (known variance)

실제로 도시 A에서 50개의 랜덤 표본을 뽑아 조사하니 로그 월수입의 평균은 5.5, 표준편차는 $s = 4.8$이었다. X의 실제 표준편차는 모르지만 표본의 수가 작지 않으므로 표 본 표준편차 $s = 4.8$을 실제 모집단의 표준편차 $\sigma = 4.8$로 사용하도록 하자. 이에 따라 $X_1, ..., X_{50}$의 평균 $\bar{X}$는 $N(\mu, 4.8^2/50)$ 분포를 따르며 관측치는 $\bar{x} = 5.5$가 된다. 이 들을 앞의 식에 대입하면, μ의 사후분포는 $N(\mu_\pi, \sigma_\pi^2)$으로

Example for Normal Data (known variance)

실제로 도시 A에서 50개의 랜덤 표본을 뽑아 조사하니 로그 월수입의 평균은 5.5, 표준편차는 $s = 4.8$이었다. X의 실제 표준편차는 모르지만 표본의 수가 작지 않으므로 표 본 표준편차 $s = 4.8$을 실제 모집단의 표준편차 $\sigma = 4.8$로 사용하도록 하자. 이에 따라 $X_1, ..., X_{50}$의 평균 $\bar{X}$는 $N(\mu, 4.8^2/50)$ 분포를 따르며 관측치는 $\bar{x} = 5.5$가 된다. 이 들을 앞의 식에 대입하면, μ의 사후분포는 $N(\mu_\pi, \sigma_\pi^2)$으로

$$\mu_\pi = \frac{50/4.8^2 \times 5.5 + 1/5.3^2 \times 5.68}{50/4.8^2 + 1/5.3^2} = 5.503$$

Example for Normal Data (known variance)

실제로 도시 A에서 50개의 랜덤 표본을 뽑아 조사하니 로그 월수입의 평균은 5.5, 표준편차는 $s = 4.8$이었다. X의 실제 표준편차는 모르지만 표본의 수가 작지 않으므로 표 본 표준편차 $s = 4.8$을 실제 모집단의 표준편차 $\sigma = 4.8$로 사용하도록 하자. 이에 따라 $X_1, ..., X_{50}$의 평균 $\bar{X}$는 $N(\mu, 4.8^2/50)$ 분포를 따르며 관측치는 $\bar{x} = 5.5$가 된다. 이 들을 앞의 식에 대입하면, μ의 사후분포는 $N(\mu_\pi, \sigma_\pi^2)$으로

$$
\begin{aligned}
\mu_\pi &= \frac{50/4.8^2 \times 5.5 + 1/5.3^2 \times 5.68}{50/4.8^2 + 1/5.3^2} = 5.503 \\
\sigma_\pi^2 &= \frac{1}{50/4.8^2 + 1/5.3^2} = 0.4533.
\end{aligned}
$$

Example for Normal Data (known variance)

μ는 로그 월수입의 평균이므로 이를 지수변환하면 A도시의 가구당 월평균 수입을 추정할 수 있는데, 추정치는 $e^{5.503} = 245.42$만 원이고, 이 추정치의 표준오차는 $e^{\sqrt{5.503}} = 1.96$이다.

Example for Normal Data (known variance)

μ는 로그 월수입의 평균이므로 이를 지수변환하면 A도시의 가구당 월평균 수입을 추정할 수 있는데, 추정치는 $e^{5.503} = 245.42$만 원이고, 이 추정치의 표준오차는 $e^{\sqrt{5.503}} = 1.96$이다.

$$\mu_\pi \quad = \quad \frac{50/4.8^2 \times 5.5 + 1/5.3^2 \times 5.68}{50/4.8^2 + 1/5.3^2} = 5.503$$

Example for Normal Data (known variance)

μ는 로그 월수입의 평균이므로 이를 지수변환하면 A도시의 가구당 월평균 수입을 추정할 수 있는데, 추정치는 $e^{5.503} = 245.42$만 원이고, 이 추정치의 표준오차는 $e^{\sqrt{5.503}} = 1.96$이다.

$$\begin{aligned}
\mu_\pi &= \frac{50/4.8^2 \times 5.5 + 1/5.3^2 \times 5.68}{50/4.8^2 + 1/5.3^2} = 5.503 \\
\sigma_\pi^2 &= \frac{1}{50/4.8^2 + 1/5.3^2} = 0.4533.
\end{aligned}$$

Example for Normal Data (known variance)

μ는 로그 월수입의 평균이므로 이를 지수변환하면 A도시의 가구당 월평균 수입을 추정할 수 있는데, 추정치는 $e^{5.503} = 245.42$만 원이고, 이 추정치의 표준오차는 $e^{\sqrt{5.503}} = 1.96$이다.

$$
\begin{aligned}
\mu_\pi &= \frac{50/4.8^2 \times 5.5 + 1/5.3^2 \times 5.68}{50/4.8^2 + 1/5.3^2} = 5.503 \\
\sigma_\pi^2 &= \frac{1}{50/4.8^2 + 1/5.3^2} = 0.4533.
\end{aligned}
$$

이 예에서는 사전분포 μ의 분산이 표본평균 $\bar{x}$의 분산에 비해 상당히 큰 값으로 사전분포의 영향이 거의 없는 경우이다

Example for Normal Data (known variance)

새로운 관측치 X_{new}의 예측 분포도 앞의 공식을 통해 찾을 수 있다.

Example for Normal Data (known variance)

새로운 관측치 X_{new}의 예측 분포도 앞의 공식을 통해 찾을 수 있다.

앞의 식에서 예측 분포는 $N(\mu_\pi, \sigma^2 + \sigma_\pi^2)$, 즉

Example for Normal Data (known variance)

새로운 관측치 X_{new}의 예측 분포도 앞의 공식을 통해 찾을 수 있다.

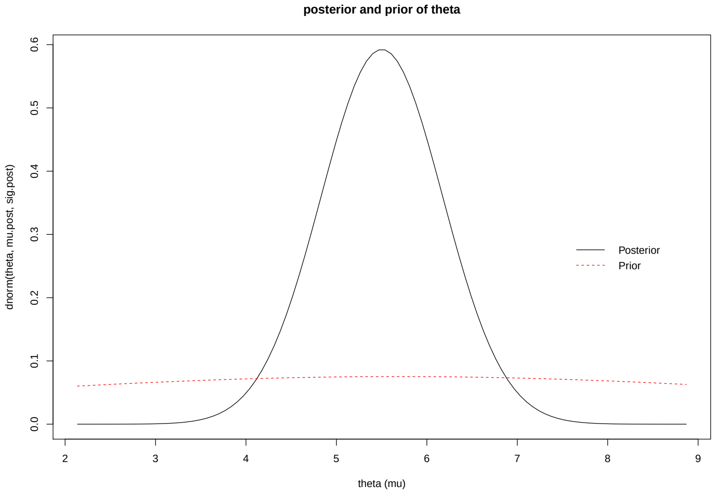
앞의 식에서 예측 분포는 $N(\mu_\pi, \sigma^2 + \sigma_\pi^2)$, 즉

$$X_{new} \mid x_1, ..., x_n \sim N(5.503, 4.8^2 + 0.4533).$$

Posterior vs. Prior

```
mu0 = 5.68; sig0 = 5.30; n = 50; xbar = 5.5; s = 4.8
#posterior
c= n/s^2 ; c0=1/sig0^2; w = c/(c+c0)
mu.post = w * xbar + (1-w)* mu0; sig.post = sqrt( 1/(c+c0) )
theta=seq(mu.post-5*sig.post, mu.post+5*sig.post, length=100)
plot(theta, dnorm(theta,mu.post, sig.post), type= "l",
main= "posterior and prior of theta", xlab = "theta (mu)" )
lines(theta,dnorm(theta,mu0,sig0),lty=2, col = 2)
legend(7.5,0.3,legend= c( "Posterior", "Prior"),lty= c(1,2),
col= c(1,2), bty= "n")
```

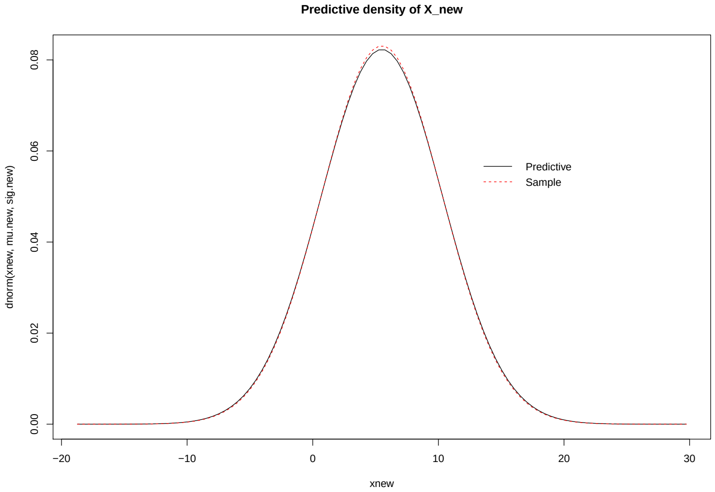
Posterior v.s. Prior



Predictive vs. Data

```
#predictive
mu.new = mu.post; sig.new = sqrt( s^2 + sig.post^2 )
xnew=seq(mu.new - 5* sig.new , mu.new + 5* sig.new, length=100)
plot(xnew, dnorm(xnew, mu.new, sig.new),type= "l",
main= "Predictive density of X_new")
lines(xnew,dnorm(xnew,xbar,s),lty=2, col = 2)
legend(12.5,0.06,legend= c( "Predictive", "Sample"),lty= c(1,2),
col= c(1,2), bty= "n")
```

Predictive vs. Data



Conjugate analysis with Normal Data (known μ)

- $X_1, ..., X_n \sim N(\mu, \sigma^2)$ with μ known and σ^2 unknown.

Conjugate analysis with Normal Data (known μ)

- $X_1, ..., X_n \sim N(\mu, \sigma^2)$ with μ known and σ^2 unknown.
- We will make inference about σ^2.

Conjugate analysis with Normal Data (known μ)

- $X_1, ..., X_n \sim N(\mu, \sigma^2)$ with μ known and σ^2 unknown.
- We will make inference about σ^2.
- The likelihood of σ^2:

Conjugate analysis with Normal Data (known μ)

- $X_1, ..., X_n \sim N(\mu, \sigma^2)$ with μ known and σ^2 unknown.
- We will make inference about σ^2.
- The likelihood of σ^2:

$$L(\sigma^2 \mid x) \propto (\sigma^2)^{-\frac{n}{2}} \exp\left\{ \frac{n}{2\sigma^2} \left[\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2 \right] \right\}$$

- Let W denote the sufficient statistic: $\frac{1}{n} \sum_{i=1}^{n} (X_i - \mu)^2$.

Conjugate analysis with Normal Data (known μ)

- $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ with μ known and σ^2 unknown.
- We will make inference about σ^2.
- The likelihood of σ^2:

$$L(\sigma^2 \mid x) \propto (\sigma^2)^{-\frac{n}{2}} \exp\left\{ \frac{n}{2\sigma^2} \left[\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2 \right] \right\}$$

- Let W denote the sufficient statistic: $\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$.
- The conjugate prior for σ^2 is the inverse gamma distribution.

Conjugate analysis with Normal Data (known μ)

- The conjugate prior for σ^2 is the inverse gamma distribution.

Conjugate analysis with Normal Data (known μ)

- The conjugate prior for σ^2 is the inverse gamma distribution.
- If a r.v. $Y \sim$ gamma, then $Y^{-1} \sim$ inverse gamma (IG).

Conjugate analysis with Normal Data (known μ)

- The conjugate prior for σ^2 is the inverse gamma distribution.
- If a r.v. $Y \sim$ gamma, then $Y^{-1} \sim$ inverse gamma (IG).
- The prior for σ^2:

Conjugate analysis with Normal Data (known μ)

- The conjugate prior for σ^2 is the inverse gamma distribution.
- If a r.v. $Y \sim$ gamma, then $Y^{-1} \sim$ inverse gamma (IG).
- The prior for σ^2:

$$p(\sigma^2) = \frac{\beta^\alpha}{\Gamma(\alpha)}(\sigma^2)^{-(\alpha+1)} e^{-\left(\frac{\beta}{\sigma^2}\right)} \quad \text{for} \quad \sigma^2 > 0$$

where $\alpha, \beta > 0$.

Conjugate analysis with Normal Data (known μ)

- The prior mean:

Conjugate analysis with Normal Data (known μ)

- The prior mean:

$$\mathbb{E}(\sigma^2) = \frac{\beta}{\alpha - 1} \quad \text{provided that} \quad \alpha > 1.$$

Conjugate analysis with Normal Data (known μ)

- The prior mean:

$$\mathbb{E}(\sigma^2) = \frac{\beta}{\alpha - 1} \quad \text{provided that} \quad \alpha > 1.$$

- The prior variance:

Conjugate analysis with Normal Data (known μ)

- The prior mean:

$$\mathbb{E}(\sigma^2) = \frac{\beta}{\alpha - 1} \quad \text{provided that} \quad \alpha > 1.$$

- The prior variance:

$$\textbf{var}(\sigma^2) = \frac{\beta^2}{(\alpha - 1)^2(\alpha - 2)} \quad \text{provided that} \quad \alpha > 2.$$

Conjugate analysis with Normal Data (known μ)

- The prior mean:

$$\mathbb{E}(\sigma^2) = \frac{\beta}{\alpha - 1} \quad \text{provided that} \quad \alpha > 1.$$

- The prior variance:

$$\textbf{var}(\sigma^2) = \frac{\beta^2}{(\alpha - 1)^2(\alpha - 2)} \quad \text{provided that} \quad \alpha > 2.$$

- The posterior for σ^2:

Conjugate analysis with Normal Data (known μ)

- The prior mean:

$$\mathbb{E}(\sigma^2) = \frac{\beta}{\alpha - 1} \quad \text{provided that} \quad \alpha > 1.$$

- The prior variance:

$$\textbf{var}(\sigma^2) = \frac{\beta^2}{(\alpha - 1)^2(\alpha - 2)} \quad \text{provided that} \quad \alpha > 2.$$

- The posterior for σ^2:

$$\begin{aligned}
\pi(\sigma^2 \mid x) &\propto L(\sigma^2 \mid x)p(\sigma^2) \\
&\propto (\sigma^2)^{-\frac{n}{2}} e^{-\frac{n}{2\sigma^2} w} (\sigma^2)^{-(\alpha+1)} e^{-(\beta/\sigma^2)} \\
&= (\sigma^2)^{-(\alpha + \frac{n}{2} + 1)} e^{-\frac{\beta + \frac{n}{2} w}{\sigma^2}}.
\end{aligned}$$

Conjugate analysis with Normal Data (known μ)

- Conjugate: The posterior is an $IG(\alpha + \frac{n}{2}, \beta + \frac{n}{2}w)$ distribution, where $w = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$.

A Conjugate analysis with Normal Data (known μ)

- How to choose the prior parameters α and β?

A Conjugate analysis with Normal Data (known μ)

- How to choose the prior parameters α and β?
- Note that common estimators are

$$\alpha = \frac{\mathbb{E}(\sigma^2)^2}{var(\sigma^2)} + 2$$

and

$$\beta = \mathbb{E}(\sigma^2)\left\{\frac{\mathbb{E}(\sigma^2)^2}{var(\sigma^2)} + 1\right\}.$$

A Conjugate analysis with Normal Data (known μ)

- How to choose the prior parameters α and β?
- Note that common estimators are

$$\alpha = \frac{\mathbb{E}(\sigma^2)^2}{var(\sigma^2)} + 2$$

and

$$\beta = \mathbb{E}(\sigma^2)\left\{\frac{\mathbb{E}(\sigma^2)^2}{var(\sigma^2)} + 1\right\}.$$

- So we could make guesses about $\mathbb{E}(\sigma^2)$ and $var(\sigma^2)$ and use these to determine α and β.

Model for Normal Data (both unknown)

- $X_1, ..., X_n$ are iid $N(\mu, \sigma^2)$ with **unknown** μ and σ^2.

Model for Normal Data (both unknown)

- $X_1, ..., X_n$ are iid $N(\mu, \sigma^2)$ with **unknown** μ and σ^2.
- The **conjugate prior** for the mean explicitly **depends on** the variance:

$$p(\sigma^2) \quad \propto \quad (\sigma^2)^{-(\alpha+1)} e^{\beta/\sigma^2},$$

Model for Normal Data (both unknown)

- $X_1, ..., X_n$ are iid $N(\mu, \sigma^2)$ with **unknown** μ and σ^2.
- The **conjugate prior** for the mean explicitly **depends on** the variance:

$$
\begin{aligned}
p(\sigma^2) &\propto (\sigma^2)^{-(\alpha+1)} e^{\beta/\sigma^2}, \\
p(\mu \mid \sigma^2) &\propto (\sigma^2)^{-\frac{1}{2}} e^{\frac{1}{2\sigma^2/s_0}(\mu-\delta)^2}.
\end{aligned}
$$

- The prior parameter s_0 measures the analyst's confidence in the prior specification.

Model for Normal Data (both unknown)

- $X_1, ..., X_n$ are iid $N(\mu, \sigma^2)$ with **unknown** μ and σ^2.
- The **conjugate prior** for the mean explicitly **depends on** the variance:

$$\begin{aligned} p(\sigma^2) &\propto (\sigma^2)^{-(\alpha+1)} e^{\beta/\sigma^2}, \\ p(\mu \mid \sigma^2) &\propto (\sigma^2)^{-\frac{1}{2}} e^{\frac{1}{2\sigma^2/s_0}(\mu-\delta)^2}. \end{aligned}$$

- The prior parameter s_0 measures the analyst's confidence in the prior specification.
- When s_0 is large, we strongly believe in our prior.

Joint Posterior for (μ, σ^2)

Joint Posterior for (μ, σ^2)

$\pi(\mu, \sigma^2 \mid x) \propto L(\mu, \sigma^2 \mid x) \cdot p(\sigma^2) \cdot p(\mu \mid \sigma^2)$

Joint Posterior for (μ, σ^2)

$$\pi(\mu, \sigma^2 \mid x) \propto L(\mu, \sigma^2 \mid x) \cdot p(\sigma^2) \cdot p(\mu \mid \sigma^2)$$

$$\propto (\sigma^2)^{-\alpha-\frac{n}{2}-\frac{3}{2}} \exp\left\{-\frac{\beta}{\sigma^2} - \frac{1}{\sigma^2} - \sum_{i=1}^{n}(x_i - \mu)^2 - \frac{1}{2\sigma^2/s_0}(\mu - \delta)^2\right.$$

Joint Posterior for (μ, σ^2)

$$\pi(\mu, \sigma^2 \mid x) \propto L(\mu, \sigma^2 \mid x) \cdot p(\sigma^2) \cdot p(\mu \mid \sigma^2)$$

$$\propto (\sigma^2)^{-\alpha - \frac{n}{2} - \frac{3}{2}} \exp\left\{ -\frac{\beta}{\sigma^2} - \frac{1}{\sigma^2} - \sum_{i=1}^{n} (x_i - \mu)^2 - \frac{1}{2\sigma^2 / s_0} (\mu - \delta)^2 \right\}$$

$$= (\sigma^2)^{-\alpha - \frac{n}{2} - \frac{3}{2}} \exp\left\{ -\frac{\beta}{\sigma^2} - \frac{1}{\sigma^2} - \sum_{i=1}^{n} (x_i^2 - 2x_i\mu + \mu^2) - \frac{1}{2\sigma^2 / s_0} (\mu^2 - 2\mu\delta + \delta^2) \right\}$$

$$= \left\{ (\sigma^2)^{-\alpha - \frac{n}{2} - 1} \exp\left\{ -\frac{\beta}{\sigma^2} - \frac{1}{2\sigma^2} \left(\sum_{i=1}^{n} (x_i^2 - n\bar{x}) \right) \right\} \right\}$$

$$\times \left\{ (\sigma^{-1} \exp\left\{ -\frac{1}{2\sigma^2} ((n + s_0)\mu^2 - 2(n\bar{x} + \delta s_0)\mu + (n\bar{x} + s_0\delta^2)) \right\} \right\}.$$

Note the second part is simply **a normal kernel** for μ.

Posterior for σ^2

Posterior for σ^2

To get the posterior for σ^2, we integrate out μ:

Posterior for σ^2

To get the posterior for σ^2, we integrate out μ:

$$\begin{aligned}
\pi(\sigma^2 \mid x) &= \int_{-\infty}^{\infty} \pi(\mu, \sigma^2 \mid x) d\mu \\
&\propto (\sigma^2)^{-\alpha - \frac{n}{2} - 1} \exp\left\{ -\frac{1}{\sigma^2}(\beta + \frac{1}{2}(\sum x_i^2 - n\bar{x}) + \frac{1}{2}\frac{s_0 n}{s_0 + n}(\bar{x} - \delta)^2 \right\}.
\end{aligned}$$

Posterior for σ^2

To get the posterior for σ^2, we integrate out μ:

$$\begin{aligned}\pi(\sigma^2 \mid x) &= \int_{-\infty}^{\infty} \pi(\mu, \sigma^2 \mid x) d\mu \\ &\propto (\sigma^2)^{-\alpha-\frac{n}{2}-1} \exp\left\{-\frac{1}{\sigma^2}(\beta + \frac{1}{2}(\sum x_i^2 - n\bar{x}) + \frac{1}{2}\frac{s_0 n}{s_0 + n}(\bar{x} - \delta)^2\right\}.\end{aligned}$$

Hence, we see the posterior for σ^2 is inverse gamma:

Posterior for σ^2

To get the posterior for σ^2, we integrate out μ:

$$
\begin{aligned}
\pi(\sigma^2 \mid x) &= \int_{-\infty}^{\infty} \pi(\mu, \sigma^2 \mid x) d\mu \\
&\propto (\sigma^2)^{-\alpha - \frac{n}{2} - 1} \exp\left\{ -\frac{1}{\sigma^2}(\beta + \frac{1}{2}(\sum x_i^2 - n\bar{x}) + \frac{1}{2}\frac{s_0 n}{s_0 + n}(\bar{x} - \delta)^2 \right\}.
\end{aligned}
$$

Hence, we see the posterior for σ^2 is inverse gamma:

$$\sigma^2 \mid x \sim IG\left(\alpha + \frac{n}{2}, \beta + \frac{1}{2}\left(\sum (x_i - \bar{x})^2 + \frac{s_0 n}{s_0 + n}(\bar{x} - \delta)^2\right)\right).$$

Posterior for μ

Posterior for μ

$$\pi(\mu \mid \sigma^2, x) = \frac{\pi(\mu, \sigma^2 \mid x)}{\pi(\sigma^2 \mid x)}.$$

Posterior for μ

$$\pi(\mu \mid \sigma^2, x) = \frac{\pi(\mu, \sigma^2 \mid x)}{\pi(\sigma^2 \mid x)}.$$

- After some cancellations,

$$\pi(\mu \mid \sigma^2, x) \propto \sigma^{-2} \exp\left\{-\frac{1}{2\sigma^2}((n + s_0)\mu^2 - 2(n\bar{x} + \delta s_0)\mu + (n\bar{x} + s_0\delta^2))\right\}$$

Posterior for μ

$$\pi(\mu \mid \sigma^2, x) = \frac{\pi(\mu, \sigma^2 \mid x)}{\pi(\sigma^2 \mid x)}.$$

- After some cancellations,

$$\pi(\mu \mid \sigma^2, x) \propto \sigma^{-2} \exp\left\{-\frac{1}{2\sigma^2}((n+s_0)\mu^2 - 2(n\bar{x} + \delta s_0)\mu + (n\bar{x} + s_0\delta^2))\right\}$$

$$= \sigma^{-2} \exp\left\{-\frac{1}{2\sigma^2/(n+s_0)}\left(\mu^2 - 2\frac{n\bar{x} + \delta s_0}{n+s_0}\mu + \frac{n\bar{x}^2 + s_0\delta^2}{n+s_0}\right)\right\}.$$

- Hence, $\pi(\mu \mid \sigma^2, x)$ is **Normal**.

$$\mu \mid \sigma^2, x \sim N\left(\frac{n\bar{x} + \delta s_0}{n+s_0}, \frac{\sigma^2}{n+s_0}\right).$$

Vague Prior: $s_0 \to 0$

- Note as $s_0 \to 0$, (the level of belief of prior information is very weak)

Vague Prior: $s_0 \to 0$

- Note as $s_0 \to 0$, (the level of belief of prior information is very weak)

$$\mu \mid \sigma^2,\ x \sim N\left(\bar{x}, \frac{\sigma^2}{n}\right).$$

Vague Prior: $s_0 \to 0$

- Note as $s_0 \to 0$, (the level of belief of prior information is very weak)

$$\mu \mid \sigma^2,\ x \sim N\left(\bar{x}, \frac{\sigma^2}{n}\right).$$

- Note also the conditional posterior mean is

Vague Prior: $s_0 \to 0$

- Note as $s_0 \to 0$, (the level of belief of prior information is very weak)

$$\mu \mid \sigma^2,\ x \sim N\left(\bar{x}, \frac{\sigma^2}{n}\right).$$

- Note also the conditional posterior mean is

$$\frac{n}{n+s_0}\bar{x} + \frac{s_0}{n+s_0}\delta.$$

Vague Prior: $s_0 \to 0$

- Note as $s_0 \to 0$, (the level of belief of prior information is very weak)

$$\mu \mid \sigma^2,\ x \sim N\left(\bar{x}, \frac{\sigma^2}{n}\right).$$

- Note also the conditional posterior mean is

$$\frac{n}{n+s_0}\bar{x} + \frac{s_0}{n+s_0}\delta.$$

- The relative sizes of n and s_0 determine the weighting of the sample mean $\bar{x}$ and the prior mean δ.

Marginal Posterior for μ

- The marginal posterior for μ is:

$$\begin{aligned}\pi(\mu \mid x) &= \int_0^\infty \pi(\mu, \sigma^2 \mid x) d\sigma^2 \\ &= \int_0^\infty (\sigma^2)^{-\alpha-\frac{n}{2}-\frac{3}{2}} \exp\left\{\frac{2\beta + (s_0+n)(\mu-\delta)^2}{2\sigma^2}\right\} d\sigma^2.\end{aligned}$$

- Letting $A = 2\beta + (s_0+n)(\mu-\delta)^2$, $z = \frac{A}{2\sigma^2}$.
- Then $\sigma^2 = \frac{A}{2z}$ and $d\sigma^2 = -\frac{A}{2z^2} dz$

A Model for Normal Data (both unknown)

- The marginal posterior for μ is:

$$\begin{aligned}
\pi(\mu \mid x) &\propto \int_0^\infty \left(\frac{A}{2z}\right)^{-\alpha-\frac{n}{2}-\frac{3}{2}} \frac{A}{2z^2} e^{-z} dz \\
&= \int_0^\infty \left(\frac{A}{2z}\right)^{-\alpha-\frac{n}{2}-\frac{1}{2}} \frac{1}{z} e^{-z} dz \\
&\propto A^{-\alpha-\frac{n}{2}-\frac{1}{2}} \int_0^\infty z^{-\alpha-\frac{n}{2}-\frac{1}{2}-1} e^{-z} dz
\end{aligned}$$

- This integrand is the kernel of a **gamma density** and thus the integral is a constant.

A Model for Normal Data (both unknown)

- The marginal posterior for μ is:

$$
\begin{aligned}
\pi(\mu \mid x) &\propto A^{-\alpha-\frac{n}{2}-\frac{1}{2}} \\
&= \left(2\beta + (s_0+n)(\mu-\delta)^2\right)^{\frac{-2\alpha+n+1}{2}} \\
&\propto \left(1 + \frac{(s_0+n)(\mu-\delta)^2}{2\beta}\right)^{\frac{-2\alpha+n+1}{2}}
\end{aligned}
$$

- which is a (scaled) noncentral $t-$kernel having noncentrality parameter δ and degrees of freedom $n+2\alpha$.

Example

서울시 거주 학생들의 2009년 1인당 월평균 사교육비 지출액수를 추정하고자 한다. 2009년 서울시 거주 학생들의 1인당 월별 사교육비 지출액이 정규분포를 따르고 평균이 μ, 표준편차가 σ 라고하자. 2008년 100명을 대상으로 한 조사에 따르면 학생 1인당 사교육비 지출은 월평균 22.5만원이고, 표준편차는 20.4 만원이었다. 2008년의 자료 는 100명에 대한 조사지만 과거의 자료를 현재 자료와 동등하게 취급하는 것은 문제가 있으므로, 과거의 자료를 약 5개의 현재자료와 동등한 정보를 준다고 간주하자.

Example

μ와 σ^2의 사전분포를 $\sigma^2 \sim IG(\alpha, \beta)$, $\mu \mid \sigma^2 \sim N(\delta, \sigma^2/s_0)$ 로 가정하고, 과거 자료를이용하여 사전분포의 모수를결정한다. μ에 대한 사전 분포에서 $\delta = 22.5$, $s_0 = 5$로 놓는 것이 적당할 것으로 보인다. α와 β의 경우 일반화된 estimator는 없으나 posteior 분포를 통해 $\frac{\text{관측치수}}{2} = \frac{s_0}{2} = \frac{5}{2}$라고 할 수 있다. β의 경우에도 $\frac{3}{2} \times 20.4^2$라고 하자

Example

종합하면, 다음과 같이 가정할 수 있다.

$$\sigma^2 \sim IG(\frac{5}{2}, \frac{3}{2}20.4^2),$$

Example

종합하면, 다음과 같이 가정할 수 있다.

$$\sigma^2 \sim IG(\frac{5}{2}, \frac{3}{2}20.4^2),$$

$$\mu \mid \sigma^2 \sim N(22.5, \sigma^2/5)$$

Example

2009년 서울시 거주 학부모 20명을 대상으로 자녀의 사교육비 현황을 조사한 결과 학생 1 인당 월평균 27.4만 원을 지출하였다고 답하였고 표준편차 $s = 22.5$만 원이었다. 이로 부터 사후분포를 구하면,

Example

2009년 서울시 거주 학부모 20명을 대상으로 자녀의 사교육비 현황을 조사한 결과 학생 1 인당 월평균 27.4만 원을 지출하였다고 답하였고 표준편차 $s = 22.5$만 원이었다. 이로 부터 사후분포를 구하면,

$$
\begin{aligned}
\mu_\pi &= \frac{s_0\delta + n\bar{x}}{s_0 + n} = \frac{5 \times 22.5 + 20 \times 27.4}{5 + 20} = 26.42. \\
\alpha_\pi &= \frac{n}{2} + \alpha = \frac{20}{2} + \frac{5}{2} \\
\beta_\pi &= \frac{1}{2}\{9618.75 + \frac{5 \times 20}{5 + 20}(27.4 - 22.5)^2 + \frac{3}{2} \times 20.4^2 = 74.03^2
\end{aligned}
$$

Example

종합하면 다음과 같은 사후 분포를 얻을 수 있다.

Example

종합하면 다음과 같은 사후 분포를 얻을 수 있다.

$$\sigma^2 \mid x_1, \ldots, x_n \sim IG\left(\frac{5}{2}, \frac{3}{2} \times 20.4^2\right),$$

Example

종합하면 다음과 같은 사후 분포를 얻을 수 있다.

$$\sigma^2 \mid x_1, ..., x_n \sim IG\left(\frac{5}{2}, \frac{3}{2} \times 20.4^2\right),$$

$$\mu \mid \sigma^2, \mid x_1, ..., x_n \sim N\left(22.5, \sigma^2/5\right)$$

Example

종합하면 다음과 같은 사후 분포를 얻을 수 있다.

$$\sigma^2 \mid x_1, ..., x_n \sim IG\left(\frac{5}{2}, \frac{3}{2} \times 20.4^2\right),$$

$$\mu \mid \sigma^2, \mid x_1, ..., x_n \sim N\left(22.5, \sigma^2/5\right)$$

즉,

$$\mu, \sigma^2 \mid x_1, ..., x_n \sim N\left(22.5, \sigma^2/5\right) \times IG\left(\frac{5}{2}, \frac{3}{2} \times 20.4^2\right),$$

R-code for Multivariate Normal Distribution

```
## joint posterior distribution
install.packages("MCMCpack")
library(MCMCpack)
#prior
mu0=22.5; s0=5; sig0=20.4
a=s0/2; b=(a-1)*sig0^2
#data
xbar=27.4; n=20; s=22.5
#parameters
mu.theta.post=(s0* mu0+n*xbar)/(s0+n)
a.post=n/2+a
b.post=1/2*((n-1)*s^2+s0*n/(s0+n)*(xbar-mu0)^2)+b
```

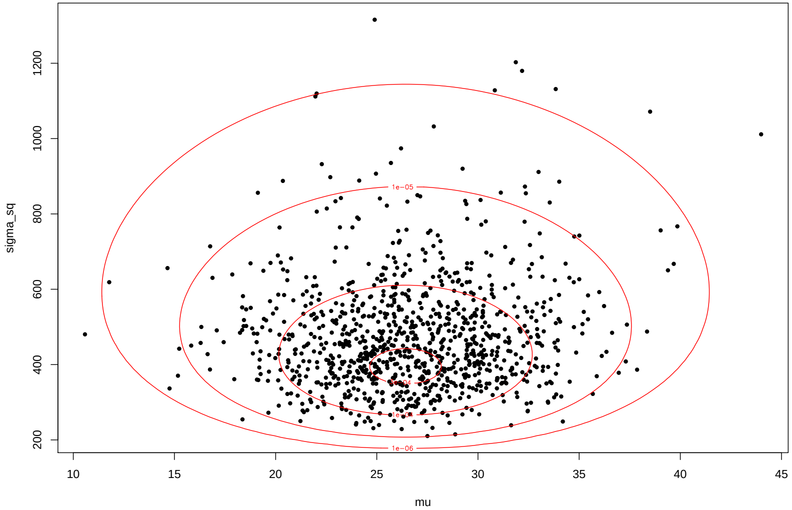
R-code for Multivariate Normal Distribution

```
f=function(theta,sigsq,a.post,b.post,mu.theta.post,n,s0){
    f= dinvgamma(sigsq,a.post,scale=b.post)*
    dnorm(theta,mu.theta.post, sqrt ( sigsq/ ( n+s0)) )
}


post.joint=
outer(theta.grid,sigsq.grid,f,a.post,b.post,mu.theta.post,n,s0)
#Monte Carlo using 1000 samples
Nsim=1000
sigsq.sim=1/rgamma(Nsim,a.post,b.post)
theta.sim=rnorm(Nsim,mu.theta.post,sqrt(sigsq.sim/(s0+n)))
```

R-code for Multivariate Normal Distribution

```
#Contour Plot

par(mfrow=c(1,1))

plot(theta.sim, sigsq.sim, xlab="mu",ylab="sigma_sq", pch =20)

contour(theta.grid,sigsq.grid,post.joint,

level= c(1.e-6,1.e-5,1.e-4,3.e-4),add= T, col = "red")
```

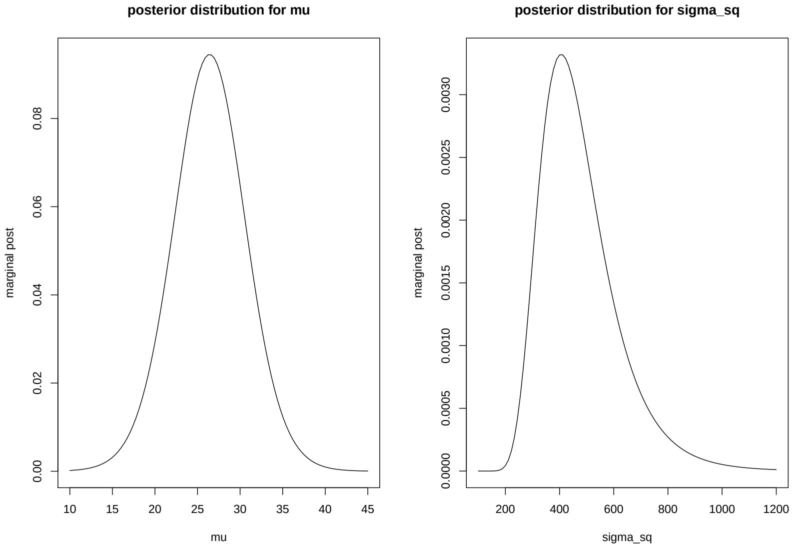
Multivariate Normal Distribution



R-code for Multivariate Normal Distribution

```
#marginal posterior.
post.sigsq= dinvgamma(sigsq.grid,a.post, scale=b.post)
post.theta=theta.grid*0
for(j in 1:length(theta.grid) ){
post.theta[j]= mean( dnorm(theta.grid[j],mu.theta.post,
sqrt(sigsq.sim/(n+s0)) ))
}

par(mfrow=c(1,2))
plot(theta.grid, post.theta, type="l",xlab="mu",ylab="marginal post")
title("posterior distribution for mu")
plot(sigsq.grid, post.sigsq, type="l",xlab="sigma_sq",ylab="marginal post")
title("posterior distribution for sigma_sq")
```

R-code for Multivariate Normal Distribution



Noninformative (Vague) Priors with Normal Data

- The conjugate priors we have discussed include a certain amount of **subjective** prior information.

Noninformative (Vague) Priors with Normal Data

- The conjugate priors we have discussed include a certain amount of **subjective** prior information.
- Another approach is to use a noninformative or vague prior.

Noninformative (Vague) Priors with Normal Data

- The conjugate priors we have discussed include a certain amount of **subjective** prior information.
- Another approach is to use a noninformative or vague prior.
- Consider $(X_1, ..., X_n) \sim_{iid} N(\mu, \sigma^2)$, with μ, σ^2 unknown.

Noninformative (Vague) Priors with Normal Data

- The conjugate priors we have discussed include a certain amount of **subjective** prior information.
- Another approach is to use a noninformative or vague prior.
- Consider $(X_1, ..., X_n) \sim_{iid} N(\mu, \sigma^2)$, with μ, σ^2 unknown.
- We can use the vague priors for μ and σ^2:

Noninformative (Vague) Priors with Normal Data

- The conjugate priors we have discussed include a certain amount of **subjective** prior information.
- Another approach is to use a noninformative or vague prior.
- Consider $(X_1, ..., X_n) \sim_{iid} N(\mu, \sigma^2)$, with μ, σ^2 unknown.
- We can use the vague priors for μ and σ^2:

$$\begin{aligned} p(\mu) &\propto 1, \text{ if } -1 < \mu < 1 \text{ independent priors here} \\ p(\sigma) &= 1/\sigma, \quad 0 < \sigma < 1 \end{aligned}$$

Noninformative (Vague) Priors with Normal Data

- The conjugate priors we have discussed include a certain amount of **subjective** prior information.
- Another approach is to use a noninformative or vague prior.
- Consider $(X_1, ..., X_n) \sim_{iid} N(\mu, \sigma^2)$, with μ, σ^2 unknown.
- We can use the vague priors for μ and σ^2:

$$\begin{array}{rcl} p(\mu) & \propto & 1, \text{ if } -1 < \mu < 1 \text{ independent priors here} \\ p(\sigma) & = & 1/\sigma, \quad 0 < \sigma < 1 \end{array}$$

- Clearly these priors are improper since they integrate to ∞ and thus are **not** true densities.

Noninformative (Vague) Priors with Normal Data

- The conjugate priors we have discussed include a certain amount of **subjective** prior information.
- Another approach is to use a noninformative or vague prior.
- Consider $(X_1, \ldots, X_n) \sim_{iid} N(\mu, \sigma^2)$, with μ, σ^2 unknown.
- We can use the vague priors for μ and σ^2:

$$\begin{aligned} p(\mu) &\propto 1, \text{ if } -1 < \mu < 1 \text{ independent priors here} \\ p(\sigma) &= 1/\sigma, \quad 0 < \sigma < 1 \end{aligned}$$

- Clearly these priors are improper since they integrate to ∞ and thus are **not** true densities.
- This is OK, as long as the resulting posteriors are proper densities.

Joint Posterior for μ and σ^2

- The joint posterior for μ and σ^2 is:

Joint Posterior for μ and σ^2

- The joint posterior for μ and σ^2 is:

$$\pi(\mu, \sigma^2 \mid x) \propto L(\mu, \sigma^2 \mid x) p(\mu) p(\sigma^2).$$

Joint Posterior for μ and σ^2

- The joint posterior for μ and σ^2 is:

$$\pi(\mu, \sigma^2 \mid x) \propto L(\mu, \sigma^2 \mid x) p(\mu) p(\sigma^2).$$

- Note that

$$L(\mu, \sigma^2 \mid x) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\}$$

Joint Posterior for μ and σ^2

- The joint posterior for μ and σ^2 is:

$$\pi(\mu, \sigma^2 \mid x) \propto L(\mu, \sigma^2 \mid x) p(\mu) p(\sigma^2).$$

- Note that

$$L(\mu, \sigma^2 \mid x) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\}$$

$$= (\sqrt{2\pi}\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2} \sum [(x_i - \bar{x}) - (\mu - \bar{x})]^2\}$$

Joint Posterior for μ and σ^2

- The joint posterior for μ and σ^2 is:

$$\pi(\mu, \sigma^2 \mid x) \propto L(\mu, \sigma^2 \mid x) p(\mu) p(\sigma^2).$$

- Note that

$$L(\mu, \sigma^2 \mid x) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\}$$

$$= (\sqrt{2\pi}\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2} \sum [(x_i - \bar{x}) - (\mu - \bar{x})]^2\}$$

$$= (\sqrt{2\pi}\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2} \{\sum (x_i - \bar{x})^2 - 2 \sum (x_i\mu - x_i\bar{x} - \bar{x}\mu + \bar{x}^2) + n(\bar{x} - \mu)^2\}\}$$

Joint Posterior for μ and σ^2

- The joint posterior for μ and σ^2 is:

$$\pi(\mu, \sigma^2 \mid x) \propto L(\mu, \sigma^2 \mid x) p(\mu) p(\sigma^2).$$

- Note that

$$L(\mu, \sigma^2 \mid x) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\}$$

$$= (\sqrt{2\pi}\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2} \sum [(x_i - \bar{x}) - (\mu - \bar{x})]^2\}$$

$$= (\sqrt{2\pi}\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2} \{\sum (x_i - \bar{x})^2 - 2 \sum (x_i\mu - x_i\bar{x} - \bar{x}\mu + \bar{x}^2) + n(\bar{x} - \mu)^2\}\}$$

$$\propto (\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}[(n-1)s^2 + n(\bar{x} - \mu)^2]\right\}$$

Joint Posterior for μ and σ^2

- The joint posterior for μ and σ^2 is:

$$\pi(\mu, \sigma^2 \mid x) \propto L(\mu, \sigma^2 \mid x)p(\mu)p(\sigma^2).$$

- Note that

$$L(\mu, \sigma^2 \mid x) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\}$$

$$= (\sqrt{2\pi}\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2}\sum[(x_i - \bar{x}) - (\mu - \bar{x})]^2\}$$

$$= (\sqrt{2\pi}\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2}\{\sum(x_i - \bar{x})^2 - 2\sum(x_i\mu - x_i\bar{x} - \bar{x}\mu + \bar{x}^2) + n(\bar{x} - \mu)^2\}\}$$

$$\propto (\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}[(n-1)s^2 + n(\bar{x} - \mu)^2]\right\}$$

- Hence, the posterior is:

Joint Posterior for μ and σ^2

- The joint posterior for μ and σ^2 is:

$$\pi(\mu, \sigma^2 \mid x) \propto L(\mu, \sigma^2 \mid x) p(\mu) p(\sigma^2).$$

- Note that

$$L(\mu, \sigma^2 \mid x) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\}$$

$$= (\sqrt{2\pi}\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2} \sum [(x_i - \bar{x}) - (\mu - \bar{x})]^2\}$$

$$= (\sqrt{2\pi}\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2} \{\sum (x_i - \bar{x})^2 - 2 \sum (x_i\mu - x_i\bar{x} - \bar{x}\mu + \bar{x}^2) + n(\bar{x} - \mu)^2\}\}$$

$$\propto (\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2} [(n-1)s^2 + n(\bar{x} - \mu)^2]\right\}$$

- Hence, the posterior is:

$$\pi(\mu, \sigma^2 \mid x) \propto L(\mu, \sigma^2 \mid x)(1)\frac{1}{\sigma^2} \propto (\sigma^2)^{-\frac{n}{2}-1} \exp\left\{-\frac{1}{2\sigma^2} [(n-1)s^2 + n(\mu - \bar{x})^2]\right\}$$

Marginal Posterior for μ

Marginal Posterior for μ

- To get the marginal posterior for μ, integrate out σ^2 using the formula:

$$\int_0^\infty x^{-(b+1)} \exp\left\{-\frac{a}{x^2}\right\} dx = \frac{1}{2} a^{-\frac{b}{2}} \Gamma\left(\frac{b}{2}\right).$$

Marginal Posterior for μ

- To get the marginal posterior for μ, integrate out σ^2 using the formula:

$$\int_0^{\infty} x^{-(b+1)} \exp\left\{-\frac{a}{x^2}\right\} dx = \frac{1}{2} a^{-\frac{b}{2}} \Gamma\left(\frac{b}{2}\right).$$

- For ease of calculation, let us use the

$$\pi(\mu, \sigma \mid x) \propto L(\mu, \sigma \mid x)(1)\frac{1}{\sigma} \propto \sigma^{-n-1} \exp\left\{-\frac{1}{2\sigma^2}[(n-1)s^2 + n(\mu - \bar{x})^2]\right\}$$

Marginal Posterior for μ

- To get the marginal posterior for μ, integrate out σ^2 using the formula:

$$\int_0^\infty x^{-(b+1)}\exp\left\{-\frac{a}{x^2}\right\}dx = \frac{1}{2}a^{-\frac{b}{2}}\Gamma\left(\frac{b}{2}\right).$$

- For ease of calculation, let us use the

$$\pi(\mu,\sigma \mid x) \propto L(\mu,\sigma \mid x)(1)\frac{1}{\sigma} \propto \sigma^{-n-1}\exp\left\{-\frac{1}{2\sigma^2}[(n-1)s^2 + n(\mu-\bar{x})^2]\right\}$$

$$\begin{aligned}
\pi(\mu \mid x) &= \int_0^\infty \pi(\mu,\sigma \mid x)d\sigma \\
&\propto \frac{1}{2}\left\{\frac{1}{2}\left[(n-1)s^2 + n(\mu-\bar{x})^2\right]\right\}^{-\frac{n}{2}}\Gamma\left(\frac{n}{2}\right) \\
&= \frac{1}{2}[(n-1)s^2]^{-\frac{n}{2}}\left[1+\frac{n(\mu-\bar{x})^2}{(n-1)s^2}\right]^{-\frac{n}{2}}\left(\frac{1}{2}\right)^{-\frac{n}{2}}\Gamma\left(\frac{n}{2}\right) \\
&\propto \left\{\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\left[\frac{n/s^2}{(n-1)\pi}\right]^{-\frac{1}{2}}\right\}\left[1+\frac{1}{n-1}\left(\frac{\mu-\bar{x}}{s/\sqrt{n}}\right)^2\right]^{-\frac{n}{2}}.
\end{aligned}$$

Posterior Distribution for μ

- Making the transformation $t = \frac{\mu - \bar{x}}{s/\sqrt{n}}$ with Jacobian $J = \frac{s}{\sqrt{n}}$:

Posterior Distribution for μ

- Making the transformation $t = \frac{\mu - \bar{x}}{s/\sqrt{n}}$ with Jacobian $J = \frac{s}{\sqrt{n}}$:

$$\pi(t \mid x) = \frac{\Gamma(\frac{n-1+1}{2})}{\Gamma(\frac{n-1}{2})} \frac{1}{[(n-1)\pi]^{1/2}\left[1 + \frac{t^2}{n-1}\right]^{\frac{n-1+1}{2}}}$$

Posterior Distribution for μ

- Making the transformation $t = \frac{\mu - \bar{x}}{s/\sqrt{n}}$ with Jacobian $J = \frac{s}{\sqrt{n}}$:

$$\pi(t \mid x) = \frac{\Gamma(\frac{n-1+1}{2})}{\Gamma(\frac{n-1}{2})} \frac{1}{[(n-1)\pi]^{1/2} \left[1 + \frac{t^2}{n-1}\right]^{\frac{n-1+1}{2}}}$$

- This is clearly a t-distribution with $n-1$ degrees of freedom.

Posterior Distribution for μ

- Making the transformation $t = \frac{\mu - \bar{x}}{s/\sqrt{n}}$ with Jacobian $J = \frac{s}{\sqrt{n}}$:

$$\pi(t \mid x) = \frac{\Gamma(\frac{n-1+1}{2})}{\Gamma(\frac{n-1}{2})} \frac{1}{[(n-1)\pi]^{1/2}\left[1 + \frac{t^2}{n-1}\right]^{\frac{n-1+1}{2}}}$$

- This is clearly a t-distribution with $n-1$ degrees of freedom.
- Since the posterior distribution for μ is t-distribution, Bayes inference and classical inference are the same.

Marginal Distribution for σ^2

$$\pi(\sigma^2 \mid x) = \int_{-\infty}^{\infty} \pi(\mu, \sigma^2 \mid x) d\mu$$

Marginal Distribution for σ^2

$$\pi(\sigma^2 \mid x) = \int_{-\infty}^{\infty} \pi(\mu, \sigma^2 \mid x) d\mu$$

$$\propto (\sigma^2)^{\frac{-(n+1)}{2}} \exp\{-\frac{1}{2\sigma^2}(n-1)s^2\} \int_{-\infty}^{\infty} \exp\{-\frac{1}{2}\sigma^2 n(\mu - \bar{x})^2\} d\mu$$

Marginal Distribution for σ^2

$$\pi(\sigma^2 \mid x) = \int_{-\infty}^{\infty} \pi(\mu, \sigma^2 \mid x) d\mu$$

$$\propto (\sigma^2)^{\frac{-(n+1)}{2}} \exp\{-\frac{1}{2\sigma^2}(n-1)s^2\} \int_{-\infty}^{\infty} \exp\{-\frac{1}{2}\sigma^2 n(\mu - \bar{x})^2\} d\mu$$

$$= (\sigma^2)^{\frac{-(n+1)}{2}} \exp\{-\frac{1}{2\sigma^2}(n-1)s^2\} \left[\sqrt{2\pi\frac{\sigma^2}{n}}\right].$$

Marginal Distribution for σ^2

$$\pi(\sigma^2 \mid x) = \int_{-\infty}^{\infty} \pi(\mu, \sigma^2 \mid x) d\mu$$

$$\propto (\sigma^2)^{\frac{-(n+1)}{2}} \exp\{-\frac{1}{2\sigma^2}(n-1)s^2\} \int_{-\infty}^{\infty} \exp\{-\frac{1}{2}\sigma^2 n(\mu - \bar{x})^2\} d\mu$$

$$= (\sigma^2)^{\frac{-(n+1)}{2}} \exp\{-\frac{1}{2\sigma^2}(n-1)s^2\} \left[\sqrt{2\pi \frac{\sigma^2}{n}}\right].$$

- Hence, the posterior distribution for σ^2 is:

Marginal Distribution for σ^2

$$\pi(\sigma^2 \mid x) = \int_{-\infty}^{\infty} \pi(\mu, \sigma^2 \mid x) d\mu$$

$$\propto (\sigma^2)^{\frac{-(n+1)}{2}} \exp\{-\frac{1}{2\sigma^2}(n-1)s^2\} \int_{-\infty}^{\infty} \exp\{-\frac{1}{2}\sigma^2 n(\mu - \bar{x})^2\} d\mu$$

$$= (\sigma^2)^{\frac{-(n+1)}{2}} \exp\{-\frac{1}{2\sigma^2}(n-1)s^2\} \left[\sqrt{2\pi\frac{\sigma^2}{n}}\right].$$

- Hence, the posterior distribution for σ^2 is:

$$\sigma^2 \mid x \sim \mathrm{IG}\left(\frac{n-1}{2}, \frac{(n-1)s^2}{2}\right).$$

Summary for the Vague Priors with Normal Data

- Compared to the posteriors in the conjugate analyses, they are more diffuse (spread).

Summary for the Vague Priors with Normal Data

- Compared to the posteriors in the conjugate analyses, they are more diffuse (spread).
- This is because we had vague prior information.

Summary for the Vague Priors with Normal Data

- Compared to the posteriors in the conjugate analyses, they are more diffuse (spread).
- This is because we had vague prior information.
- For a large sample size, there is little difference between the conjugate analysis and the "noninformative" analysis.