# A Conjugate analysis with Normal Data (variance known)

- Hence the posterior for $\mu$ is simply a normal distribution with mean
$$\frac{\frac{\delta}{\tau^2} + \frac{n\bar{x}}{\sigma^2}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}}$$
and variance
$$\left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{-1} = \frac{\tau^2\sigma^2}{\sigma^2 + n\tau^2}$$
- The **precision** is the reciprocal of the **variance**.
- Here, $\frac{1}{\tau^2}$ is the **prior precision** ...
- $\frac{n}{\sigma^2}$ is the **data precision** ...
- ... and $\frac{1}{\tau^2} + \frac{n}{\sigma^2}$ is the **posterior precision**.

# A Conjugate analysis with Normal Data (variance known)

- Note the posterior mean $E[\mu|\mathbf{x}]$ is simply

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

  a combination of the **prior mean** and the **sample mean**.
- If the prior is highly precise, the weight is large on $\delta$.
- If the data are highly precise (e.g., when $n$ is large), the weight is large on $\bar{x}$.
- Clearly as $n \rightarrow \infty$, $E[\mu|\mathbf{x}] \approx \bar{x}$, and $var[\mu|\mathbf{x}] \approx \frac{\sigma^2}{n}$ if we choose a large prior variance $\tau^2$.
- This implies that for $\tau^2$ large and $n$ large, Bayesian and frequentist inference about $\mu$ will be nearly identical.

# A Conjugate analysis with Normal Data (mean known)

- Now suppose $X_1, \ldots, X_n$ are iid $N(\mu, \sigma^2)$ with $\mu$ known and $\sigma^2$ unknown.
- We will make inference about $\sigma^2$.
- Our likelihood

$$L(\sigma^2|\mathbf{x}) \propto (\sigma^2)^{-\frac{n}{2}} e^{-\frac{n}{2\sigma^2}[\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2]}$$

- Let $W$ denote the sufficient statistic $\frac{1}{n}\sum(X_i - \mu)^2$.
- The conjugate prior for $\sigma^2$ is the **inverse gamma** distribution.
- If a r.v. $Y \sim$ gamma, then $1/Y \sim$ inverse gamma (IG).
- The prior for $\sigma^2$ is

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

where $\alpha > 0, \beta > 0$.

# A Conjugate analysis with Normal Data (mean known)

- Note the prior mean and variance are

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

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

- So the posterior for $\sigma^2$ is:

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

- Hence the posterior is clearly an $\text{IG}(\alpha + \frac{n}{2}, \beta + \frac{n}{2}w)$ distribution, where $w = \frac{1}{n}\sum(x_i - \mu)^2$. **Conjugate!**

# A Conjugate analysis with Normal Data (mean known)

- How to choose the prior parameters $\alpha$ and $\beta$?
- Note

$$\alpha = \frac{[E(\sigma^2)]^2}{var(\sigma^2)} + 2 \text{ and } \beta = E(\sigma^2)\left\{\frac{[E(\sigma^2)]^2}{var(\sigma^2)} + 1\right\}$$

so we could make guesses about $E(\sigma^2)$ and $var(\sigma^2)$ and use these to determine $\alpha$ and $\beta$.

# A Model for Normal Data (mean and variance both unknown)

- When $X_1, \ldots, X_n$ are iid $N(\mu, \sigma^2)$ with both $\mu$, $\sigma^2$ **unknown**, the conjugate prior for the mean explicitly depends on the variance:

$$p(\sigma^2) \propto (\sigma^2)^{-(\alpha+1)} e^{-\beta/\sigma^2}$$
$$p(\mu|\sigma^2) \propto (\sigma^2)^{-\frac{1}{2}} e^{-\frac{1}{2\sigma^2/s_0}(\mu-\delta)^2}$$

- 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.

# A Model for Normal Data (mean and variance both unknown)

The joint posterior for $(\mu, \sigma^2)$ is:

$$
\begin{aligned}
\pi(\mu, \sigma^2|\mathbf{x}) &\propto L(\mu, \sigma^2|\mathbf{x})p(\sigma^2)p(\mu|\sigma^2) \\
&\propto (\sigma^2)^{-\alpha-\frac{n}{2}-\frac{3}{2}} e^{-\frac{\beta}{\sigma^2}-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2-\frac{1}{2\sigma^2/s_0}(\mu-\delta)^2} \\
&= (\sigma^2)^{-\alpha-\frac{n}{2}-\frac{3}{2}} e^{-\frac{\beta}{\sigma^2}-\frac{1}{2\sigma^2}(\sum x_i^2-2n\bar{x}\mu+n\mu^2)-\frac{1}{2\sigma^2/s_0}(\mu^2-2\mu\delta+\delta^2)} \\
&= \left[(\sigma^2)^{-\alpha-\frac{n}{2}-\frac{1}{2}} e^{-\frac{\beta}{\sigma^2}-\frac{1}{2\sigma^2}(\sum x_i^2-n\bar{x}^2)}\right] \\
&\quad \times \left[(\sigma^2)^{-1} e^{-\frac{1}{2\sigma^2}\{(n+s_0)\mu^2-2(n\bar{x}+\delta s_0)\mu+(n\bar{x}^2+s_0\delta^2)\}}\right]
\end{aligned}
$$

Note the second part is simply a **normal kernel** for $\mu$.

# A Model for Normal Data (mean and variance both unknown)

- To get the posterior for $\sigma^2$, we integrate out $\mu$:

$$\pi(\sigma^2|\mathbf{x}) = \int_{-\infty}^{\infty} p(\mu, \sigma^2|\mathbf{x})\, d\mu$$
$$\propto (\sigma^2)^{-\alpha-\frac{n}{2}-\frac{1}{2}} e^{-\frac{1}{\sigma^2}[\beta+\frac{1}{2}(\sum x_i^2 - n\bar{x}^2)]}$$

  since the second piece (which depends on $\mu$) just integrates to a normalizing constant.

- Hence since $-\alpha - \frac{n}{2} - \frac{1}{2} = -(\alpha + \frac{n}{2} - \frac{1}{2}) - 1$, we see the posterior for $\sigma^2$ is inverse gamma:

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

# A Model for Normal Data (mean and variance both unknown)

- Note that
$$\pi(\mu|\sigma^2, \mathbf{x}) = \frac{\pi(\mu, \sigma^2|\mathbf{x})}{\pi(\sigma^2|\mathbf{x})}$$
- After lots of cancellation,
$$\begin{aligned}\pi(\mu|\sigma^2, \mathbf{x}) &\propto \sigma^{-2} \exp\{-\tfrac{1}{2\sigma^2}[(n + s_0)\mu^2 - 2(n\bar{x} + \delta s_0)\mu \\ &\quad + (n\bar{x}^2 + s_0\delta^2)]\} \\ &= \sigma^{-2} \exp\left\{-\tfrac{1}{2\sigma^2/(n+s_0)}\left[\mu^2 - 2\tfrac{n\bar{x}+\delta s_0}{n+s_0}\mu + \tfrac{n\bar{x}^2+s_0\delta^2}{n+s_0}\right]\right\}\end{aligned}$$
- Clearly $\pi(\mu|\sigma^2, \mathbf{x})$ is **normal**:
$$\mu|\sigma^2, \mathbf{x} \sim N\left(\frac{n\bar{x} + \delta s_0}{n + s_0}, \frac{\sigma^2}{n + s_0}\right)$$

# A Model for Normal Data (mean and variance both unknown)

- Note as $s_0 \to 0$, $\mu|\sigma^2, \mathbf{x} \dot{\sim} N\left(\bar{x}, \frac{\sigma^2}{n}\right)$.
- Note also the posterior mean is

$$\left(\frac{n}{n+s_0}\right)\bar{x} + \left(\frac{s_0}{n+s_0}\right)\delta.$$

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

# Example 1: Midge Data

- **Example 1**: $X_1, \ldots, X_9$ are a random sample of midge wing lengths (in mm). Assume the $X_i's \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$.
- Example 1(a): If we know $\sigma^2 = 0.01$, make inference about $\mu$.
- Example 1(a): Make inference about $\mu$ **and** $\sigma^2$, both **unknown**.