Bayesian Model for Multivariate Data

- Suppose each individual has q variables observed on it, so that $\mathbf{X}_1, \ldots, \mathbf{X}_n$ are q-dimensional *random vectors*.
- Assume the random vectors are iid *multivariate normal*, with mean vector $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Sigma}$.
- Then a set of conjugate priors for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are:

$$\boldsymbol{\mu}|\boldsymbol{\Sigma} \sim N_q\left(\boldsymbol{\delta}, \frac{1}{n_0}\boldsymbol{\Sigma}\right), \ \ \boldsymbol{\Sigma}^{-1} \sim \text{ Wishart}$$

- The Wishart distribution is a multivariate generalization of the gamma.
- n_0 is a tuning parameter that reflects confidence in the prior.
- If $\frac{n_0}{n}$ is larger, the analyst has more confidence in the prior.
- The posterior distributions are:

$$\boldsymbol{\mu}|\boldsymbol{\Sigma}, \mathbf{x} \sim N_q\left(\frac{n_0\boldsymbol{\delta} + n\bar{\mathbf{x}}}{n_0 + n}, \frac{1}{n_0 + n}\boldsymbol{\Sigma}\right), \ \ \boldsymbol{\Sigma}^{-1}|\mathbf{x} \sim \text{ another Wishart}$$

Vague Priors with Normal Data

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

$$p(\mu) = 1, \;\; -\infty < \mu < \infty \quad \text{(independent}$$
$$p(\sigma) = 1/\sigma, \;\; 0 < \sigma < \infty \quad \text{priors here)}$$

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

Vague Priors with Normal Data

- The joint posterior for μ and σ is:

$$\pi(\mu, \sigma|\mathbf{x}) \propto L(\mu, \sigma|\mathbf{x})p(\mu)p(\sigma)$$

Note

$$\begin{aligned}
L(\mu, \sigma|\mathbf{x}) &= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2} \\
&= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum[(x_i-\bar{x})-(\mu-\bar{x})]^2} \\
&= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\{\sum(x_i-\bar{x})^2-2\sum(x_i\mu-x_i\bar{x}-\bar{x}\mu+\bar{x}^2)+n(\bar{x}-\mu)^2\}} \\
&\propto \sigma^{-n} e^{-\frac{1}{2\sigma^2}[(n-1)s^2+n(\bar{x}-\mu)^2]}
\end{aligned}$$

So

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

Vague Priors with Normal Data

- To get the marginal posterior for μ, integrate out σ using the formula
$$\int_0^\infty x^{-(b+1)} e^{-\frac{a}{x^2}}\,\mathrm{d}x = \tfrac{1}{2} a^{-\frac{b}{2}} \Gamma\left(\frac{b}{2}\right)$$
- Let $x^2 = \sigma^2$, $b = n$, $a = \frac{1}{2}[(n-1)s^2 + n(\mu - \bar{x})^2]$.
Then
$$\begin{aligned}
\pi(\mu|\mathbf{x}) &= \int_0^\infty \pi(\mu, \sigma|\mathbf{x})\,\mathrm{d}\sigma \\
&\propto \tfrac{1}{2}\left\{\tfrac{1}{2}[(n-1)s^2 + n(\mu-\bar{x})^2]\right\}^{-\frac{n}{2}} \Gamma\left(\frac{n}{2}\right) \\
&= \tfrac{1}{2}[(n-1)s^2]^{-\frac{n}{2}} \left[1 + \frac{n(\mu-\bar{x})^2}{(n-1)s^2}\right]^{-\frac{n}{2}} \left(\frac{1}{2}\right)^{-\frac{n}{2}} \Gamma\left(\frac{n}{2}\right) \\
&\propto \left\{\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \left[\frac{n/s^2}{(n-1)\pi}\right]^{-1/2}\right\} \left[1 + \frac{1}{n-1}\left(\frac{\mu-\bar{x}}{s/\sqrt{n}}\right)^2\right]^{-\frac{n}{2}}
\end{aligned}$$

Vague Priors with Normal Data

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

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

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

Vague Priors with Normal Data

- To get the marginal distribution of σ^2, note

$$\begin{aligned}\pi(\sigma|\mathbf{x}) &= \int_{-\infty}^{\infty} \pi(\mu, \sigma|\mathbf{x})\, d\mu \\ &\propto \sigma^{-(n+1)} e^{-\frac{1}{2\sigma^2}(n-1)s^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2} n(\mu-\bar{x})^2}\, d\mu \\ &= \sigma^{-(n+1)} e^{-\frac{1}{2\sigma^2}(n-1)s^2} \left[\sqrt{2\pi\frac{\sigma^2}{n}}\right]\end{aligned}$$

- Including the term from the Jacobian of the transformation from σ to σ^2,

$$\begin{aligned}\pi(\sigma^2|\mathbf{x}) &\propto (\sigma^2)^{-(\frac{n+1}{2})} e^{-\frac{(n-1)s^2}{2\sigma^2}} (\sigma)\left|\tfrac{1}{2\sigma}\right| \\ &\propto (\sigma^2)^{-(\frac{n-1}{2}+1)} e^{-\frac{(n-1)s^2}{2}/\sigma^2}\end{aligned}$$

$$\Rightarrow \sigma^2|\mathbf{x} \sim \mathrm{IG}\left(\tfrac{n-1}{2}, \tfrac{(n-1)s^2}{2}\right)$$

Vague Priors with Normal Data

- Both of the posteriors (for μ and for σ^2) are **proper**.
- 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.
- Example 1(a): Midge data revisited: