ADVANCED BAYESIAN MODELING

ROBUST INFERENCE: **A COMPUTATIONAL TRICK**

Can we implement a Gibbs sampler for a model that uses the t distribution?

Direct natural conjugacy is not available.

Partial conjugacy is possible using a special representation ...

Scale Mixture Representation

For fixed μ, σ^2, ν, if

$$y \mid V \sim \mathrm{N}(\mu, V) \qquad \text{and} \qquad V \sim \text{Inv-}\chi^2(\nu, \sigma^2)$$

then

$$y \sim t_\nu(\mu, \sigma^2)$$

This represents the t distribution as a *scale mixture of normals*.

May be useful, since the normal and scaled inverse chi-square both have natural conjugacy.

Thus, in a Bayesian model, replace

$$y_1, \ldots, y_n \mid \mu, \sigma^2, \nu \quad \sim \quad \text{iid } t_\nu(\mu, \sigma^2)$$

with

$$\begin{aligned} y_1, \ldots, y_n \mid \mu, V_1, \ldots, V_n &\quad \sim \quad \text{indep. } \mathrm{N}(\mu, V_i) \\ V_1, \ldots, V_n \mid \sigma^2, \nu &\quad \sim \quad \text{iid Inv-}\chi^2(\nu, \sigma^2) \end{aligned}$$

This is an example of *data augmentation*:

The **auxiliary variables** V_i are present only to assist in the computation.

Gibbs Sampler

If ν is fixed, and the prior is

$$\begin{aligned} p(\mu) &\propto 1 & -\infty < \mu < \infty \\ p(\sigma^2) &\propto (\sigma^2)^{-1} & \sigma^2 > 0 \end{aligned}$$

then the conditional posterior distributions are standard and easily sampled – see BDA3, Sec. 12.1.

Thus, a Gibbs sampler is easily implemented.

(The same is true with some proper conjugate priors: normal for μ, and gamma for σ^2.)

What if ν is a parameter (not fixed)?

There is no natural conjugacy for ν, but because it is one-dimensional, it can be sampled from its conditional posterior by many methods.

For example, you could use Metropolis or Metropolis-Hastings.

This would be an example of *Metropolis-within-Gibbs*.

Note: Unless you are programming the Gibbs sampler yourself, you probably do not need this computational trick.

For example, JAGS lets you directly specify a t distribution in your model, and then it handles the computational details.

Application to Normal Sample Model

Recall: For a two-parameter normal sample with the usual improper prior, the posterior can be expressed as

$$\begin{aligned} \mu \mid \sigma^2, y &\sim \mathrm{N}(\bar{y},\, \sigma^2/n) \\ \sigma^2 \mid y &\sim \text{Inv-}\chi^2(n-1,\, s^2) \end{aligned}$$

Since it can be shown that

$$\sigma^2/n \mid y \;\sim\; \text{Inv-}\chi^2(n-1,\, s^2/n)$$

it follows from the t representation that the marginal posterior for μ is

$$\mu \mid y \;\sim\; t_{n-1}(\bar{y},\, s^2/n)$$

Thus, simulation is actually not needed to get posterior information about μ.

For example, it follows that

$$\left.\frac{\mu - \bar{y}}{s/\sqrt{n}} \right| y \quad \sim \quad t_{n-1}(0,1)$$

from which it follows that the 95% central posterior interval is the same as the classical t-interval:

$$\bar{y} \quad \pm \quad t_{n-1,0.025} \cdot s/\sqrt{n}$$