ST 740: Prior Predictive Checking

Alyson Wilson

Department of Statistics
North Carolina State University

November 13, 2013

Model Assessment

As we have through the class, we will use *model* to describe the sampling distribution, the prior distribution, any hierarchical structure, and issues such as which explanatory variables have been included in a regression.

Key Questions

1. How can I tell if my model is providing an adequate fit to the data? (*goodness of fit*)
2. What aspects of reality are not captured by my model? Are these important given the use I would like to make of the model? (*model checking*)
3. How can I tell if any of the modeling choices I have made are having an undue impact on my results? (*sensitivity analysis*)
4. Which model (or models) should I ultimately choose for the final presentation of my results? (*model selection*)

Prior/Data Conflict

Suppose that we have a model with

$$\begin{aligned} Y \mid \theta &\sim \text{Normal}(\theta, 1) \\ \theta &\sim \text{Normal}(0, 1) \end{aligned}$$

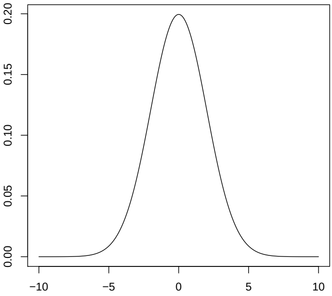
The posterior distribution for θ given one observation y is

$$\theta \mid y \sim \text{Normal}(\frac{y}{2}, 0.5)$$

Suppose that we observe $y = 10$, so our posterior distribution is Normal(5,0.5).

Prior Predictive Distribution

The prior predictive distribution for this problem is $\tilde{Y} \sim \text{Normal}(0, 2)$. Our observation of $y = 10$ is extremely unlikely under this model.



We need to think about whether we have an outlier and/or whether we may have mis-specified our prior.

Mixture Priors

Suppose that we change our prior so that we think there's a 50% chance that $\theta \sim$ Normal$(0, 1)$ and a 50% chance that $\theta \sim$ Normal$(5, 9)$.

$$\pi(\theta) = \frac{0.5}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\theta^2\right) + \frac{0.5}{3\sqrt{2\pi}} \exp\left(-\frac{1}{18}(\theta - 5)^2\right)$$

Using our likelihood $Y \sim$ Normal$(\theta, 1)$, what is our posterior distribution?

Mixture Priors

$$
\begin{aligned}
\pi(\theta \mid y) &\propto \frac{0.5}{2\pi} \exp\left(-\frac{1}{2}(\theta^2 + y^2 - 2\theta y + \theta^2)\right) + \\
&\quad \frac{0.5}{6\pi} \exp\left(-\frac{1}{18}(\theta^2 - 10\theta + 25) - \frac{1}{2}(y^2 - 2\theta y + \theta^2)\right) \\
&\propto \frac{0.5 \exp(-0.25 y^2)}{2\pi} \exp\left(-(\theta^2 - \frac{2\theta y}{2} + \frac{y^2}{4})\right) + \\
&\quad \frac{0.5}{6\pi} \exp(-\frac{25}{18} - \frac{y^2}{2} + \frac{(5+9y)^2}{180}) \\
&\quad \exp\left(-\frac{10}{18}(\theta^2 - 2\theta\frac{5+9y}{10} + (\frac{5+9y}{10})^2)\right)
\end{aligned}
$$

Mixture Priors

$$
\begin{aligned}
\pi(\theta \mid y) &\propto \frac{0.5\exp(-0.25y^2)}{2\pi}\exp\left(-(\theta-\frac{y}{2})^2\right) + \\
&\quad \frac{0.5}{6\pi}\exp(-\frac{25}{18}-\frac{y^2}{2}+\frac{(5+9y)^2}{180}) \\
&\quad \exp\left(-\frac{10}{(2)(9)}(\theta-\frac{5+9y}{10})^2\right) \\
&\propto \frac{\exp(-0.25y^2)}{2\sqrt{\pi}}\text{Normal}\left(\frac{y}{2},\frac{1}{2}\right) + \\
&\quad \frac{1}{4\sqrt{5\pi}}\exp\left(-\frac{25}{18}-\frac{y^2}{2}+\frac{(5+9y)^2}{180}\right)\text{Normal}\left(\frac{5+9y}{10},\frac{9}{10}\right)
\end{aligned}
$$

Mixture Priors

In order for the posterior to integrate to 1, we need the mixture probabilities to add to 1. Let

$$\gamma = \frac{\frac{0.5\exp(-0.25y^2)}{2\pi}}{\frac{0.5\exp(-0.25y^2)}{2\pi} + \frac{1}{4\sqrt{5\pi}}\exp\left(-\frac{25}{18} - \frac{y^2}{2} + \frac{(5+9y)^2}{180}\right)}$$

Then

$$\pi(\theta \mid y) = \gamma \mathrm{Normal}\left(\frac{y}{2}, \frac{1}{2}\right) + (1-\gamma)\mathrm{Normal}\left(\frac{5+9y}{10}, \frac{9}{10}\right)$$

If $y = 10$, $\gamma \approx 0$. If $y = 2.1$, $\gamma = 0.43$.

Mixture Priors

Posterior distribution when $y = 2.1$