Noninformative ("Default") Bayes: A Quick Review

Rebecca C. Steorts

Bayesian Methods and Modern Statistics: STA 360/601

Lecture 7

Exam I

- Exam Thursday, Feb 11th in class. Be early to class so that you can start you exam on time.
- You will need pencil and paper. No calculators, no computers, no cell phones, etc permitted. No notes permitted.
- The exam will cover material through Module 4. This includes all readings.
- Assignment 2 solutions have been posted.
- Assignment 3 has been posted.
- There was an optional homework problem with Module 3, Part I. The solutions have been posted.
- Lab this week: Review sessions to prepare for the exam.

Exam I

- Intro to Bayes. What is it and why do we use it?
- Decision theory - loss, risk (all three of them).
- Hierarchical modeling - conjugacy, priors, posteriors, likelihood.
- Consistency, posterior predictive, credible intervals.
- Objective Bayes

Exam I: Expect 5 problems. You will need to really know the material to get through this exam.

Today's menu

- This is all material that was covered last class (we will do a quick review).
- Review of invariance.
- Review of transformations (Uniform and Jeffreys').
- Tying back into invariance.
- Next: Module 5: Monte Carlo!

Invariant in distribution

Let θ be our parameter of interest.

Transform to $g(\theta)$.

The transformation is said to be invariant (in distribution) if the distributions of θ and $g(\theta)$ have the same form (Normal, Beta, etc).

Furthermore, we could have invariance of parameters:

- location (mean),
- the scale (variance),
- or both (mean and variance).

Why is it nice to have such a property?

Suppose θ is the true population height in inches!

However, we receive some data from Europe and the data is now in cm.

Instead of reformatting the data, we could just transform the parameter.

Also, we would hope that our prior is not sensitive to a slight change in our parameter (inches, cm).

Recall the uniform prior on θ

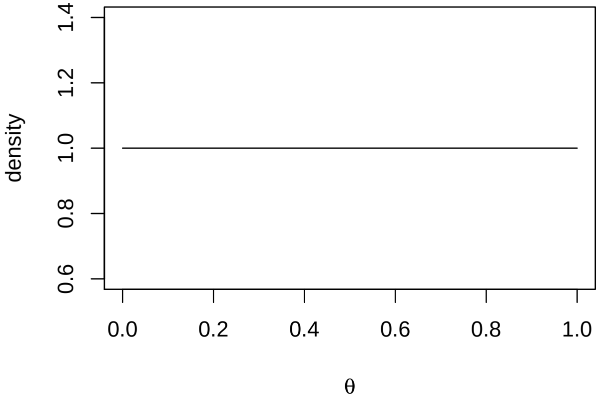

Figure 1: Unif(0,1) prior on θ.

What happens if we consider though the transformation to $\phi = \frac{1}{\theta}$.

Think we're moving from the variance to the precision!

$$\phi = \frac{1}{\theta} \implies \theta = \frac{1}{\phi}. \tag{1}$$

Then using a change of variables transformation,

$$|\frac{\partial \theta}{\partial \phi}| = \frac{1}{\phi^2}. \tag{2}$$

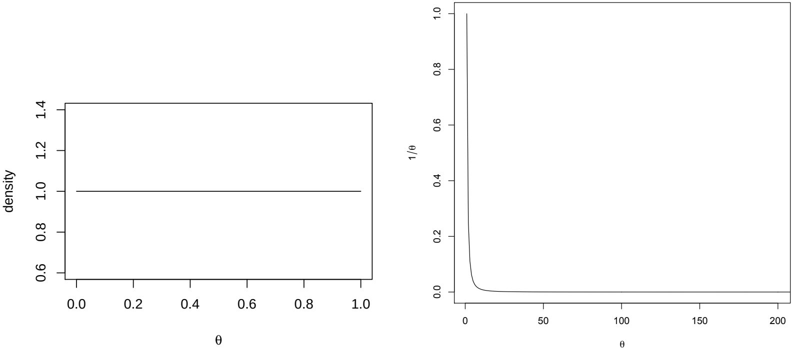

Figure 2: Comparison of the Uniform prior and the transformed prior on θ.

Recall Jeffrey's prior

Let's go back over Jeffreys' when $X \sim \text{Bin}(n, \theta)$.

How do we calculate Jeffreys' prior?

Specifically, we showed that

$$p_J(\theta) \propto \text{Beta}(1/2, 1/2).$$

Recall Jeffrey's prior

Recall

$$p_J(\theta) = \sqrt{\frac{n}{\theta(1-\theta)}}.$$

What happens if we consider though the transformation to $\phi = \dfrac{1}{\theta}$?

$$\begin{aligned}
p_J(\phi) &= p_J(1/\phi) \times \left|\frac{\partial \theta}{\partial \phi}\right| && (3)\\
&= \sqrt{\frac{n}{\frac{1}{\phi}(1-\frac{1}{\phi})}} \times \frac{1}{\phi^2} && (4)\\
&\propto \frac{\phi}{\sqrt{\phi-1}} \times \frac{1}{\phi^2} && (5)\\
&\propto \frac{1}{\phi\sqrt{\phi-1}}, \;\; \text{where } \; 1 \leq \phi < \infty. && (6)
\end{aligned}$$

Is the resulting transformation invariant? Yes!

It's invariant because $\phi = \dfrac{1}{\theta}$ is a one-to-one function.

Now consider the transformation $\theta = \frac{1}{\sqrt{\phi}}$.

$$|\frac{\partial\theta}{\partial\phi}| = \frac{1}{2\phi^{3/2}}. \tag{7}$$

$$\begin{align}
p_J(1/\sqrt{\phi}) &= p_J(1/\sqrt{\phi}) \times |\frac{\partial\theta}{\partial\phi}| \tag{8}\\
&= (\phi)^{-1/4}\left(\frac{\phi^{1/2}-1}{\phi^{1/2}}\right)^{-1/2} \times \frac{1}{2\phi^{3/2}} \tag{9}\\
&\propto \phi^{3/2}(\phi^{1/2}-1)^{-1/2} \text{ where } 1 \leq \phi < \infty. \tag{10}
\end{align}$$

Is the resulting transformation invariant? Why or why not? (It's also invariant because the transformation is a one-to-one function).

Take home message with Jeffreys'

Consider parameter θ and transformation $g(\theta)$.

Jeffreys' prior is invariant under parameterization means that Jeffreys' prior corresponds to $g(\theta)$ is the *same* as applying a change of "measure" or distribution to the Jeffreys' prior for θ.

Said differently, if $\theta = g(\theta)$, then 1 and 2 are the same below. Let J represent a Jeffreys' prior.

1. $$J_\phi = J_\theta \times |\frac{\partial g^{-1}(\theta)}{\partial \theta}|$$

2. $$J_\phi = \sqrt{I(\phi)},$$

where $I(\phi)$ is the Fisher information of ϕ.

The proof of this is omitted. (I will post it time permitting).[1]

[1] Slides 10–13 will not be on midterm 1. You may be asked about other slides or easier questions about objective Bayes and invariance.