STAT J535: Chapter 5: Classes of Bayesian Priors

David B. Hitchcock
E-Mail: hitchcock@stat.sc.edu

Spring 2012

The Bayesian Prior

- A prior distribution **must** be specified in a Bayesian analysis.
- The choice of prior can substantially affect posterior conclusions, especially when the sample size is not large.
- We now examine several broad methods of determining prior distributions.

Conjugate Priors

- We know that **conjugacy** is a property of a prior **along with a likelihood** that implies the posterior distribution will have the same *distributional form* as the prior (just with different parameter(s)).
- We have seen some examples of conjugate priors:
  Data/Likelihood **Prior**
  1. Bernoulli $\longrightarrow$ Beta for p
  2. Poisson $\longrightarrow$ Gamma for λ
  3. Normal $\longrightarrow$ Normal for μ
  4. Normal $\longrightarrow$ Inverse gamma for σ^2

Conjugate Priors

Other examples:

1. Multinomial $\longrightarrow$ Dirichlet for $p_1, p_2, \ldots, p_k$
2. Negative Binomial $\longrightarrow$ Beta for p
3. Uniform(0, θ) $\longrightarrow$ Pareto for upper limit
4. Exponential $\longrightarrow$ Gamma for β
5. Gamma (β unknown) $\longrightarrow$ Gamma for β
6. Pareto (α unknown) $\longrightarrow$ Gamma for α
7. Pareto (β unknown) $\longrightarrow$ Pareto for β

Conjugate Priors: Exponential Family

- Consider the family of distributions known as the **one-parameter exponential family**.
- This family consists of any distribution whose p.d.f. (or p.m.f.) can be written as:

$$f(x|\theta) = e^{[t(x)u(\theta)]}r(x)s(\theta)$$

  where $t(x)$ and $r(x)$ do not depend on the parameter θ and $u(\theta)$ and $s(\theta)$ do not depend on x.
- Note that any such density can be written as

$$f(x|\theta) = e^{\{t(x)u(\theta)+\ln[r(x)]+\ln[s(\theta)]\}}$$

Conjugate Priors: Exponential Family

- If we observe an iid sample $X_1, \ldots, X_n$, the joint density of the data is thus

$$f(\mathbf{x}|\theta) = e^{\{u(\theta)\sum_{i=1}^{n} t(x_i) + \sum_{i=1}^{n} \ln[r(x_i)] + n \ln[s(\theta)]\}}$$

- Consider a prior for θ (with the prior parameters k and γ) having the form:

$$p(\theta) = c(k, \gamma) e^{\{k u(\theta)\gamma + k \ln[s(\theta)]\}}$$

Conjugate Priors: Exponential Family

Then the posterior is

$$\begin{aligned}
\pi(\theta|\mathbf{x}) &\propto f(\mathbf{x}|\theta)p(\theta) \\
&\propto \exp\left\{ u(\theta)\sum t(x_i) + n\ln[s(\theta)] + ku(\theta)\gamma + k\ln[s(\theta)] \right\} \\
&= \exp\left\{ u(\theta)\left[\sum t(x_i) + k\gamma\right] + (n+k)\ln[s(\theta)] \right\} \\
&= \exp\left\{ (n+k)u(\theta)\left[\frac{\sum t(x_i) + k\gamma}{n+k}\right] + (n+k)\ln[s(\theta)] \right\}
\end{aligned}$$

which is of the same form as the prior, except with "k"$= n + k$ and "γ"$= \dfrac{\sum t(x_i) + k\gamma}{n+k}$.

$\Rightarrow$ If our data are iid from a one-parameter exponential family, then a conjugate prior will exist.

Conjugate Priors

- Conjugate priors are mathematically convenient.
- Sometimes they are quite flexible, depending on the specific hyperparameters we use.
- But they reflect very specific prior knowledge, so we should be wary of using them unless we truly possess that prior knowledge.

Uninformative Priors

- These priors intentionally provide very little specific information about the parameter(s).
- A classic uninformative prior is the *uniform* prior.
- A *proper* uniform prior integrates to a finite quantity.
- **Example 1**: For Bernoulli(θ) data, a uniform prior on θ is

$$p(\theta) = 1, \;\; 0 \leq \theta \leq 1.$$

- This makes sense when θ has **bounded support**.

Uninformative Priors

- **Example 2**: Consider $N(0, \sigma^2)$ data. If it is "reasonable" to assume, that, say $\sigma^2 < 100$, we could use the uniform prior

$$p(\sigma^2) = \frac{1}{100}, \;\; 0 \leq \sigma^2 \leq 100$$

  (even though σ^2 is not intrinsically bounded).
- An **improper** uniform prior integrates to ∞:
- **Example 3**: $N(\mu, 1)$ data with

$$p(\mu) = 1, \;\; -\infty < \mu < \infty.$$

- This is fine as long as the resulting **posterior** is proper.
- But be careful: Sometimes an improper prior will yield an improper posterior.