Conjugate Priors

For a given likelihood, if prior leads to a posterior in the same family of distributions, then the prior is called a conjugate prior.

Example (Geometric):

Suppose Y_i are i.i.d. Geometric(p), $0 \le p \le 1$

$$P(Y_i = y) = p(1-p)^{y-1}, \quad y = 1, 2, \ldots$$

The likelihood is

$$\prod_{i=1}^{n} P(Y_i = y_i) = \prod_{i=1}^{n} p(1-p)^{y_i - 1}$$

$$= p^n (1-p)^{\sum_{i=1}^{n} y_i - n}$$

A conjugate prior should have sample space $[0,1]$ and be of the form $p^{\text{something}} (1-p)^{\text{something else}}$

We use Beta(α, β) for a conjugate prior.

$$\cancel{PDF} \quad f(p) = p^{\alpha-1}(1-p)^{\beta-1}, \quad 0 \le p \le 1$$

$$f(p|\underline{y}) \propto p^n (1-p)^{\sum_{i=1}^n y_i - n} \times p^{\alpha-1}(1-p)^{\beta-1}$$

$$= p^{n+\alpha-1}(1-p)^{\sum_{i=1}^n y_i - n + \beta - 1}$$

posterior distribution $\sim Beta(n+\alpha, \sum_{i=1}^n y_i - n + \beta)$

Advantages

1. All you have to do is calculate new paramaters
2. The posterior could become a new prior if you have additional information.

Uninformative Priors

An uninformative prior assumes all values of a parameter are equally likely. The prior in the battery example was ~~uninformative~~ uninformative. An uninformative prior on a Poisson(λ) parameter would be $\lambda \sim Uniform(0,\infty)$

We would $f(\lambda) = \frac{1}{\infty - 0} = 0$

But $\int_0^\infty 0 \, d\lambda = 0$ so this is not a valid prior.

Nevertheless, some statisticians define an uninformative prior by

$$f(\lambda) = 1$$

In which case $f(\lambda | \underline{y}) \propto f(\underline{y} | \lambda)$

Loss Functions

To find an estimator of a parameter from the posterior, we use a loss function. Different loss functions result in different estimators.

~~A loss function~~ Suppose we want estimate a function of $\phi(\theta)$. A loss function is a function of a statistic $T(\underline{Y})$ and ϕ such that

1. It is 0 when $T(\underline{Y}) = \phi$
2. It is positive when $T(\underline{Y}) \neq \phi$ and is non-decreasing as $T(\underline{Y})$ gets further away from ϕ.

Given a posterior and loss function the Bayesian estimator $T(\underline{Y})$ is the $T(\underline{Y})$ which which minimizes the expected loss with respect to the posterior, where the expected posterior loss (EPL) is defined to be

$$E[L(T(\underline{y}), \underline{\theta})]$$

where E is taken with respect to the posterior on θ. (data is $\underline{y}$).

$$E[L(T(\underline{y}), \underline{\theta})] = \int L(T(\underline{y}), \underline{\theta}) f(\underline{\theta}|\underline{y}) d\underline{\theta}$$

We will be taking $p=1$, θ only parameter.