Other Updating Examples

James A. Chalfant

A natural choice for prior beliefs about p, the probability that $X = 1$ for any Bernoulli trial, is the Beta distribution:

$$f(p) \propto p^{\alpha-1}(1-p)^{\beta-1}$$

The expected value is $\alpha/(\alpha+\beta)$. Any combination of prior mean and prior variance are available by choosing α and β. (The variance is $\alpha\beta/(\alpha+\beta)^2(\alpha+\beta+1)$.)

The likelihood function then reveals that the Beta is the natural conjugate prior for the Bernoulli (or binomial, of course):

$$f(p|y) \propto f(p) \cdot \mathcal{L}(y|p)$$

or

$$f(p|y) \propto p^{\alpha-1}(1-p)^{\beta-1} \cdot p^{y}(1-p)^{(1-y)} = p^{\alpha-1+y}(1-p)^{\beta-1+(1-y)}$$

The posterior mean is thus

$$E(p|y) = \frac{\alpha+y}{\alpha+y+\beta+(1-y)} = \left(\frac{\alpha+\beta}{\alpha+\beta+1}\right)\frac{\alpha}{\alpha+\beta} + \left(\frac{1}{\alpha+\beta+1}\right)y$$

i.e., a linear combination of the prior mean and the sample mean.

One updates in the expected way with an observation of $y = 1$, toward a higher expected value for p, and similarly with an observation of $y = 0$, toward a lower expected value for p.

The analogy to the prior distribution being equivalent to another sample holds. The likelihood function for $n = 10$ observations, say, is

$$p^{\sum y_i}(1-p)^{10-\sum y_i}$$

and the *mle* is the mean, $\hat{p} = \sum y_i/10$. It is straightforward how we would update with one more observation:

$$\mathcal{L}(p|y_1, \ldots, y_{11}) = p^{\sum y_i}(1-p)^{10-\sum y_i} \cdot p^{y_{11}}(1-p)^{1-y_{11}}$$

and the *mle* would be $\hat{p} = (\sum y_i + y_{11})/11$. This could easily be shown to be 10/11 times the old estimate plus 1/11 times y_{11}, an updating formula that is similar to the Bayesian one.

It is straightforward to generalize to the case where we observe n Bernoulli trials; the sample evidence gets proportionally greater weight in the posterior mean. Also, a very similar exercise can be conducted with a Poisson random variable, and a prior belief for $E(Y) = \lambda$ of the Gamma form. Both examples are covered nicely in John Hey's *Data In Doubt*, which is out of print, unfortunately.