STAT40380/STAT40390/STAT40850
Bayesian Analysis

Dr Niamh Russell

School of Mathematics and Statistics
University College Dublin

niamh.russell@ucd.ie

February 2016

Bayesian inference for the normal distribution

So far we have met the situation where:

$$\begin{aligned} \theta &\sim N(\theta_0, \phi_0) \qquad \text{(prior)} \\ x_i|\theta &\sim N(\theta, \phi) \qquad \text{(likelihood)} \end{aligned}$$

for $i = 1, \ldots, n$
When $n = 1$ we have posterior $\theta|x \sim N(\theta_1, \phi_1)$, where:

$$\phi_1 = \left(\frac{1}{\phi_0} + \frac{1}{\phi} \right)^{-1}, \quad \theta_1 = \phi_1 \left(\frac{\theta_0}{\phi_0} + \frac{x}{\phi} \right)$$

For general n we have posterior $\theta|\mathbf{x} \sim N(\theta_2, \phi_2)$, where:

$$\phi_2 = \left(\frac{1}{\phi_0} + \frac{n}{\phi} \right)^{-1}, \quad \theta_2 = \phi_2 \left(\frac{\theta_0}{\phi_0} + \frac{\sum x_i}{\phi} \right)$$

Notes for the $n = 1$ case

- Remember that precision = 1/variance
- The posterior precision is equal to the prior precision + the likelihood precision
- Thus the posterior variance is going to be *smaller* than both the prior variance and the likelihood variance
- The posterior mean is a *weighted* version of the prior mean and the data observation
- The weight on the prior mean is $\frac{\phi}{\phi+\phi_0}$ and the weight on the observation is $\frac{\phi_0}{\phi+\phi_0}$
- So if the likelihood variance is small (ie ϕ is small) compared to the prior variance (ϕ_0), the posterior mean will be mostly influenced by the observation x
- However, if the likelihood variance is relatively large compared to the prior variance, the posterior mean is mostly influenced by the prior mean θ_0

Notes for general n case

- Recall that $\bar{x} \sim N(\theta, \phi/n)$, ie the sample mean is normally distributed with variance ϕ/n
- The posterior mean and variance can be rewritten as:

$$\phi_2 = \left(\frac{1}{\phi_0} + \left[\frac{\phi}{n} \right]^{-1} \right)^{-1}, \quad \theta_2 = \phi_2 \left(\frac{\theta_0}{\phi_0} + \frac{\bar{x}}{(\phi/n)} \right)$$

- Now the posterior mean can be seen as a combination of the prior on θ, *and an observation of the mean $\bar{x}$ from the above distribution*
- As before, the posterior precision is now equal to the prior precision plus the precision of the sample mean,. . .
- . . . and the posterior mean is a weighted average of the prior mean and the observed mean

Example 1: Bayesian inference on the precision

Example

Note: unrealistic example

Suppose that the likelihood is $x_i \sim N(\theta, \phi^{-1})$ for $i = 1, \ldots, n$ with θ known. A prior is suggested for the precision as $\phi \sim Ga(\alpha, \beta)$. What is the posterior distribution of $\phi | x$?

Example 2: Weights of rats

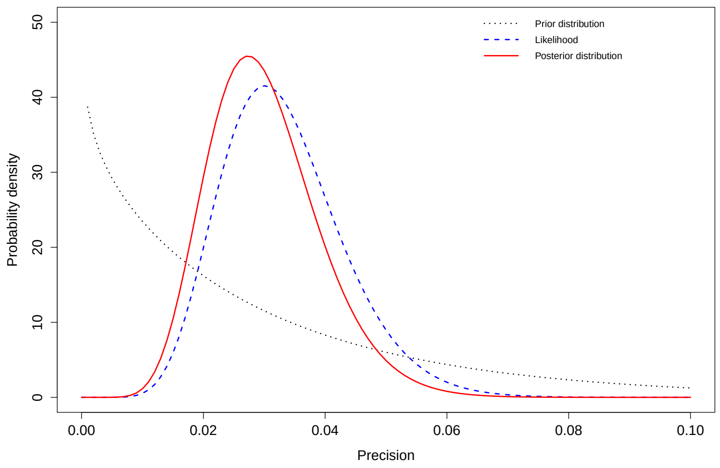
Example

The uterine weights of 20 rats in milligrammes follow a normal or Gaussian distribution. We give the following summary statistics:

$$\sum x_i = 420, \quad \sum (x_i - \bar{x})^2 = 664$$

Previous studies have suggested a $Ga(0.9, 30)$ distribution for the precision of the weights. Assuming that the mean of rats' uterine weight is known at 21.0mg, find the posterior distribution of the precision.

Example 2: picture



Tricks for finding posterior distributions

- So far we have assumed a prior distribution and likelihood, and calculated the resulting posterior.
- We have used two tricks to make things easier:
  1. Ignoring other variables that we weren't interested in by use of the proportionality term.
  2. Spotting that the posterior distribution has been of a similar form to the prior.
- The proportionality trick is something we will use throughout the course. We will discuss more on the computational and theoretical properties of the proportionality constant in later lectures.
- Conversely, spotting that the posterior distribution is of the same form as the prior occurs only occasionally with specific prior/likelihood combinations.
- When the posterior distribution is of the same form as the prior, we term the prior and likelihood *conjugate* distributions.

Example 3: The Poisson and Gamma distributions

Example

Suppose that the likelihood is $x_i \sim P(\lambda)$ for $i = 1, \ldots, n$. A prior is suggested for the rate as $\lambda \sim Ga(\alpha, \beta)$.
What is the posterior distribution of $\lambda | x$?

Example 4: Misprints

Example

The number of misprints in the first few pages of a book is found to be:

$$3, 4, 2, 1, 2, 3$$

A previous book by the same author was found to have misprints distributed as $Ga(9, 6)$. What is the posterior distribution for the rate at which mistakes occur in this author's books?