**Bayesian updating**

Suppose we have a sample of observations,

We specify some model that has a parameter , and a distribution (e.g. Normal).

For example, if we are interested in the mean of then applying the CLT, we know that the mean will be approximately normally distributed, and we can show that the variance of the mean is , where is the variance of

So our sampling distribution for the sample mean is . This becomes our likelihood function, . Suppose for simplicity we know the variance, .

We specify a prior for , , where is any background information relevant for Often we choose a Normal distribution with very large variance – this approximates a uniform over a very large interval – the convenient choice for an uninformative prior.

Applying Bayes’ rule:

In our example, if the prior and the likelihood are both Normal distributions, then the posterior will be Normal distribution.

Now suppose we get a **second sample** of observations, (same variable, x, just more observations on it).

We can update the posterior from the 1st sample, using the 2nd sample.

To do this, we use the posterior from the first sample, , and use it as the prior to combine with the likelihood from the second sample, . Using Bayes’ rule:

This will be exactly the same as if we combined both samples into one,

, so observations, and then used Bayes’ rule with the original prior:

In some published paper, they report summary statistics, mean , SD, number of obs., but the data are propriety/unavailable. What do you do?

Sufficient statistics = the summary statistics needed to give you **all** the information about the distribution of the sample.

Meta-analysis – combine the results from a number of different studies to improve precision.

For linear models (regression models)

We have a Normal likelihood function

If we choose a Normal prior for everything (really an Inverted-Gamma for )

, , with prior parameters (everything with a zero subscript).

Something like, likelihood x prior for x prior for

Often choose values for prior parameters to be uninformative.

degrees of freedom, sample estimate of the variance.

Using Bayes’ rule, this gives a Normal joint distribution for the parameters

To get the marginal distribution for each parameter, we integrate out all the other parameters. It turns out, if we do that, we get

and are each Student-t distributed.

is Inverted-Gamma distributed.

You can specify the prior as either

, ,

Or

, ,

Confusing point in the literature:

Often we use “the precision” instead of the variance .

If is inverted-gamma, then is gamma distributed, and vice versa.

The gamma distribution tends to be more stable to work with numerically and is more convenient mathematically.