Homework 1

You will collect some data yourself. Data may not come from any already existing data set. Sample size should range from 4 to 10 observations. Data should be acceptably modeled as normal. Possible examples are travel times to (or from) school in the morning (or afternoon, but not both!), heart rate or blood pressure measurements from a blood pressure machine at the drug store, jogging times for a set distance.

1. Explain what your measurements will be.
2. Before you collect the data, decide on your prior. Please use a normal density, and specify your prior mean μ_0 and standard deviation τ. So μ_0 is your best guess at the average of all your blood pressure (or other) measures, and τ is an estimate of the standard deviation that the true value may differ from your guess. (Note: you get better at this with practice, you won't be penalized for being too ridiculous in your guessing, within reason.) Explain your reasoning (1 or 2 sentences).
3. Report the data and the sample mean and variance $(n-1)$ denominator.
4. Now specify the sampling standard deviation σ. Since we are doing a one parameter model, and since σ is usually *not* known, we need to do something because we are working with such a simple model. You may either
   (a) Estimate σ yourself, or
   (b) Set σ to the sample sd of your data set.
   (c) Specify the exact value for σ that you use in all your calculations (i.e. sqrt(2), 1.41, 1.414, 1.4)

   Either way, this is commonly known as *cheating*; we often do this (Bayesians less often perhaps) in complicated models where treating the parameter as unknown complicates things substantially. If there were a later analysis, as we learn more about modeling and computation, we would relax the assumption of σ^2 known. Give your method of setting σ.
5. Calculate the posterior mean $\bar{\mu}$, variance V, and sd. Show the formulas with your data values in place of the symbols. Remember that in the likelihood, $\bar{y} \sim N(\mu, \sigma^2/n)$.

6. The *prior predictive* density is the density that you predict for a single observation before seeing any data. In this model, the prior predictive for a single observation is $y \sim N(\mu_0, \sigma^2 + \tau^2)$.

7. Construct a table with means, sds and vars for the (i) posterior for μ, (ii) the prior for μ, (iii) the prior predictive for y, and (iv) the likelihood of μ.

8. Plot on a single plot the (i) posterior for μ, (ii) the prior for μ, (iii) the prior predictive for y, and (iv) the likelihood of μ (suitably normalized so it looks like a density, ie a normal with mean $\bar{y}$ and variance σ^2/n) all on the same graph. *Interpret the plot.*

9. Write R/WinBUGS programs to sample from the posterior of μ.

10. Adapt your BUGS program to sample from the prior and prior predictive. Do this by not loading your data, rather, in loading the initial values, move the data y over to the init list instead. There is an example at the end of Homework 2 for a Poisson-gamma likelihood/prior. [Helpful step: set keyword DIC=F in the call to bugs, as WinBUGS can not calculate DIC for prior predictions.]

11. Adapt your BUGS program to sample from the likelihood.

12. Report your WinBUGS models and R code, data, and inits. Use at least samples of size 10000.

13. Construct a table with means, sds and vars for the (i) posterior for μ, (ii) the prior for μ, (iii) the prior predictive for y, and (iv) the likelihood of μ from the WinBUGS output.

14. Plot on a single plot the (i) posterior for μ, (ii) the prior for μ, (iii) the prior predictive for y, and (iv) the likelihood of μ (suitably normalized so it looks like a density, ie a normal with mean $\bar{y}$ and variance σ^2/n) all on the same graph. All from the WinBUGS output. *Interpret the plot.*

15. (no response needed). Your WinBUGS program and the exact calculations should give exactly the same results, (up to rounding error). If they are not, you've done something wrong and need to fix it.

16. (no credit) Estimate the amount of time spent specifically on this homework (don't include class, lab or reading time).