Course EPIB-675 - Bayesian Analysis in Medicine

Assignment 7

In the first three questions we will analyse data for a meta-analysis. In question 1, we will use a very simple meta-analysis model that assumes that the effects are identical across all trials. In the second we will use a hierarchical (random effects) model, that assumes that the effects across studies are not identical, but follow a common distribution. Finally, in question 3, we will continue to use a random effects model, but we will try to see if some of the variations in effects between studies can be explained by study-level covariates, by adding a regression component to the meta-analytic hierarchical model.

The basic setup for all three questions is as follows: Eleven trials have been carried out to see if inserting stents are useful following heart attacks. Each of the 11 trials has a treatment (stent) group, and a control (placebo) group. In addition, each trial either used a plastic or a metal stent. The primary objective is to see whether subjects who receive stents have fewer future events (i.e., further heart attacks).

1. Download the data set meta.txt from the course web page. It is already in WinBUGS format. WinBUGS comes with an example called Blocker, which we have discussed in class. We will use the Blocker example in question 2 below, but here we want to use a simpler meta-analysis model, without random effects. Starting from the Blocker model, therefore, we wish to change delta[i] to a single effect delta, which also means that the line giving a normal distribution to delta should move outside of the "loop" over i, and "d" and "tau" should be constants (such as 0, and 0.001) rather than variables. The rest of the program can remain as it is, but add a line that provides an odds ratio for the overall effect, i.e., add a line like:

```
or <- exp(delta)
```

Run this meta-analytic model using the meta.txt data set, monitor all unknown parameters (including pc and pt), and report the results. Does there seem to be an effect of the stents?

2. Now, using the same data set, run the blocker model, but as it was originally. Again, add a line that gets the overall odds ratio, now using delta.new rather than delta. That is, add a line like:

```
or <- exp(delta.new)
```

Compare the odds ratio for stents in the two models. Are their means similar? What about their variances? Looking at the parameter for the SD of the effect of delta (i.e., looking at sigma), does a random effects model seem warranted (i.e., does there seem to be variations in effects across the 11 studies)?

3. Finally, again use the blocker model, but now switch to the meta.reg.txt data set. This data set is identical to the one used in the first two questions, except that is adds a variable to indicate whether the coating was plastic or metal. We will see if some of the variability in study-to-study effect can be explained by the stent type by adding a regression term to the prior distribution of delta[i]. To do this, remove the three lines in blocker (note that they are not consecutive lines in the program)

```
delta[i] ~ dnorm(d, tau)
d ~ dnorm(0.0,1.0E-6)
delta.new ~ dnorm(d, tau)
```

and replace them with:

```
delta.mean[i] <- alpha + beta*coating[i]
delta[i] ~ dnorm(delta.mean[i], tau)
alpha ~ dnorm(0.0,0.001)
beta ~ dnorm(0.0,0.001)
mean.plastic <- alpha
mean.metal <- alpha + beta
delta.plastic ~ dnorm(mean.plastic, tau)
delta.metal ~ dnorm(mean.metal, tau)
or.plastic <- exp(delta.plastic)
```

```
or.metal <- exp(delta.metal)
or.diff <- or.plastic - or.metal
```

Lines with [i]'s in them (first two lines above) go inside the loop, other lines (all the rest) go outside the loop. The last line calculates the difference in odds ratios between the two stent types.

Run this model, and report results from all parameters. Does it appear that stent type explains some of the variability in the effectiveness of stents across studies?

In the next two questions we will investigate measurement error. In question 4, we will generate a data set in R, and then deliberately add measurement error to the x variable. By comparing the estimated slope in each case (i.e., with and without measurement error), we will see the effect that measurement error has on an estimated regression line. In question 5, we will use WinBUGS to see if a model specially constructed to adjust for measurement error can reconstruct the original (correct) estimates.

4. Generate a simulated linear regression data set in R that follows the following model (sample size = 100):

$$y = 2 + 5 * x, \quad \sigma = 1, \quad x \sim normal(0, 1)$$

To do this, use lines such as:

```
x <- round(rnorm(100, mean=0, sd=1),2)
y <- round(rnorm(100, mean = 2 + 5*x, sd=1),2)
```

Note that since we are using random numbers, everyone in the class will be using a slightly different data set. I rounded everything to 2 decimal places, which makes for cleaner data sets without losing too much precision.

(a) Plot x versus y.

(b) Use R to run a standard linear regression of x versus y. Provide the estimates and 95% confidence intervals for the intercept and slope (see R class notes if you forget how to do this, and recall that approximate 95% intervals can be derived from the point estimates ± 1.96 times the standard error for each parameter). Are they close to their theoretical values (of 2 and 5, respectively)?

(c) Now we will add some measurement error to the x values. In particular, we will create a measurement error version of x using the R command

```
x.error <- round(rnorm(100, mean=x, sd=2),2 )
```

Note that the measurement error version of x is centered at the true value of x, but has random noise about the observation. This is typical of measurement error seen when data are generated by an unbiased but imprecise measuring tool. Plot x.error versus y, and note any differences from your plot in part (a).

(d) Rerun the linear regression again, but this time using $x.error$ rather than x. Compare the results (point estimates and confidence intervals) you obtain here with those obtained in part (b), and note any differences.

(e) Before leaving R, save your data sets for use in WinBUGS in problem 5. To do this, use commands such as:

```
x.list <- list(x=x, yy)
xerror.list <- list(x.error = x.error, y=y)
dput(x.list, file = "c://temp//x.txt")
dput(xerror.list, file= "c://temp//xerror.txt")
```

You will use the first data set in the (a) of question 5, and the second data set in parts (b) and (c) of question 5.

5. In this question we will analyse the same two data sets as were used in question 4, but now using WinBUGS, with and without correcting for possible measurement error.

(a) Run a straightforward WinBUGS program for the linear regression of x versus y (see class notes of simple WinBUGS programs if you do not recall how to do this). Provide the point estimates and 95% credible intervals for the intercept and slope. Compare these to your estimates in part (b) of question 4 (they should be quite similar).

(b) repeat part (a), but now using $x.error$ versus y. Provide the point estimates and 95% credible intervals for the intercept and slope. Compare these to your estimates in part (d) of question 4 (again, they should be quite similar).

(c) Now, we will modify the simple linear regression model to account for any measurement error. To the basic linear regression model (from part (a), NOT part (b), because we want to estimate the true relationship with x, not the one with measurement error variable $x.error$!), add a line such as:

```
x[i] ~ dnorm(x.error[i], tau.error)
```

You will also need to add a line for the prior for tau.error. As usual, we will define tau.error in terms of sigma.error, and put a uniform prior on sigma. Use the following lines:

```
tau.error <- 1/(sigma.error*sigma.error)
sigma.error ~ dunif(1,5)
```

to indicate that it is known that the measurement error variance is between 1 and 5 (real value, recall, was 2).

Run this model, and report the point estimates and 95% credible intervals for the intercept and slope. Compare these to your estimates in part (b) of question 5 . . . has the model correctly adjusted for the measurement error?