There are *N* streams and each streams has  observations. Let  be the observation for the *k*’th site in steam *i* and **D** denote the entire data set. The parameters of the model are the mean (), the between-stream variation in average density (), the residual variation (), and the random effects . The likelihood for a random effects model integrates over the random effects, i.e.:

 (1)

Now, each data point is associated with only one random effect so we can simplify the integral from one six dimension integral to six one dimensional integral (see discussion of why we should do this in lecture 2), i.e.:

 (2)

where  is the data set for stream *i* and  is the random effect for stream *i*. Now, the likelihood of each data point is normal so:

 (3)

Given that  is also normal, we can write Equation 2 as:

 (4)

You need to write a function to do each of the N integrations included in the RHS of this equation, a function which computes Equation 4 using that function and code to use the mle function in R to find the best estimates for , , and .

Two hints:

1. The integral in Equation 4 is from  and  and if you use Simpsons rule (see the equation at the bottom of slide 10 of Friday’s lecture) you will need many many function calls (trust me). However, a trick is to scale the  by  so that Equation 4 becomes:

 (5)

You can now get away with a range for  from about -5 to 5 (but chose a small step size anyway).

1. Run the streams data set through lme (with the method=ML option) to get the correct values for the parameters to check your answers.