Penn State STAT 540
Homework #3, due Thursday, Nov 15, 2018

What you have to submit in a Canvas submission folder: (i) Your `R` code in a file titled PSUemailidHW1.R (e.g. muh10HW1.R), (ii) pdf file that contains a clear writeup for the questions below named PSUemailidHW1.pdf (e.g. muh10HW1.pdf). Note that your code should be readily usable, without any modifications.

1. Follow the bulbs example from class. Let the number of bulbs in rooms A and B be $m = 400, n = 25$ respectively. You may assume that all bulb lifetimes are independent, and distributed according to an exponential distribution with expectation θ. The number of bulbs that survived until time $\tau = 15$ in room A, S, is 254. You can find the lifetime of bulbs in Room B, $B_1, \ldots, B_n$ here `personal.psu.edu/muh10/540/bulbsB.dat`.

   (a) Find the maximum likelihood estimator of $\theta, \hat{\theta_{MLE}}$ using the Expectation-Maximization (E-M) algorithm. Write out clear pseudocode for this algorithm, with enough details that someone not familiar with the E-M algorithm could code it up. Explain also how you obtained starting values, and your stopping criteria for the algorithm. (Note that for this simple problem, you do not need an E-M algorithm as you can construct the observed data log likelihood easily. If you like, you can use the observed data log likelihood to find the MLE, and then compare that result to the answer you obtain from the E-M algorithm.)

   (b) Approximate the standard error of $\hat{\theta_{MLE}}$ using a nonparametric bootstrap, and construct a 95% confidence interval for θ. Write out pseudocode as above.

   (c) Repeat the above, but with a parametric bootstrap.

   (d) Assume that θ has a prior distribution that is Uniform(0,100). Write out clear, detailed pseudocode (including conditional distributions, how to construct Metropolis-Hastings updates etc.) for an auxiliary variable MCMC algorithm for carrying out inference for the above problem. That is, this code should be useful for approximating $\pi(\theta|B_1, \ldots, B_m)$. Before you provide the pseudocode, clearly show what the target posterior distribution is for your auxiliary variable algorithm (recall from class that this target will include the auxiliary random variables). State clearly how you would approximate $E_\pi(\theta|B_1, \ldots, B_m, S)$ using the Markov chain you have constructed.

   (e) (Optional) Write code for the above auxiliary variable MCMC algorithm and report the posterior 95% credible interval for θ and plot the marginal posterior pdf for θ.

2. Return to the above bulbs problem, except assume that the model for the bulbs is now Unif($(0, \kappa)$). The number of bulbs that survived until time

$\tau = 15$ in room A, S, is 167. You can find the lifetime of bulbs in Room B, $B_1, \ldots, B_n$ here `personal.psu.edu/muh10/540/bulbsBProb2.dat`.

(a) Find the MLE of κ using an EM algorithm. Provide pseudocode and details of the algorithm as above. Report your estimate.

(b) Find a 95% confidence interval for κ. You may use any method you like, just explain it clearly and point out any potential problems.

(c) Compare your point estimate and 95% confidence interval for κ from the previous two parts with the estimates you would obtain if you just used the data from Room B (that is, if you discarded the data from Room A).