Connor McCurley

AGR 6932 **Homework 3** Fall 2019

Question 1

Question 1 demonstrated the effects of partial pooling in hierarchical models. Given the 8 Schools model, expressed as:

$$y_i \sim Normal(\theta_i, \sigma_i)$$
$$\theta_i \sim Normal(\nu, \tau)$$

we were asked to analyze the partially pooled estimates of treatment for each school, θ_i, as a function of the uncertainty of the global effect, τ. Two cases were considered. In the first case (Figure 1) the prior on τ was considered as "uninformative". The diagram demonstrates the evolution of the estimates for fixed values of τ. (Each line represents a single θ_i estimate.) As can be seen from Figure 1, the θ_i values converge to their assumed values as the variance, τ, increases. As discussed in class, this demonstrates the trend from complete pooling (for small τ values) to no pooling.

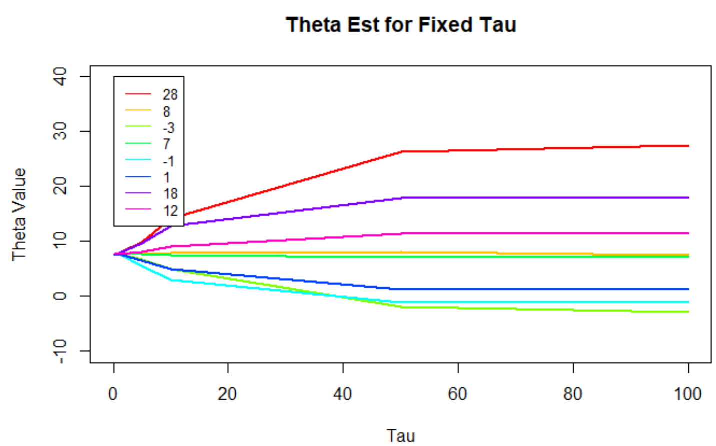


Figure 1: Θ_i as a function of its variance with an uninformative prior. Each line represents a particular school from the 8 Schools model. Complete pooling is demonstrated for small values of τ, while no pooling is observed for larger values.

A half-normal prior was then placed on τ and the simulations were repeated for different amounts of variation. As can be seen in Figure 2, the same trend can be observed as with Figure 1. However, it is assumed that more samples are needed for the model to average over the prior.

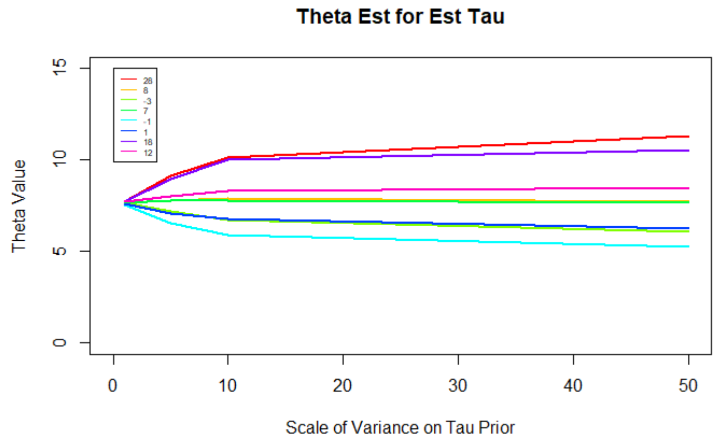


Figure 2: Θ_i as a function of its variance with a half-normal prior. Each line represents a particular school from the 8 Schools model. Complete pooling is demonstrated for small values of the variance on τ, while no pooling is observed for larger values.

Question 2

Question 2 then asked about varying logistic growth rate processes. Figure 3 shows logistic growth trajectories for 3 different rate parameters, R. It can be observed that the processes evolve smoothly as N increases. Figure 4 then shows the same trajectories with added normal observation error. As can be observed, the trajectories are significantly more chaotic.

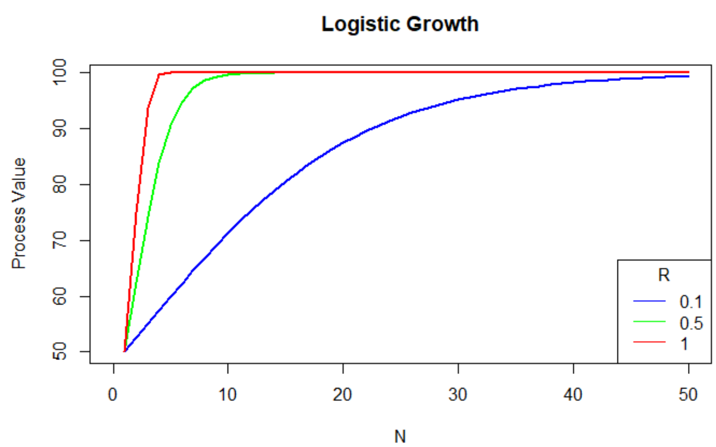


Figure 3: Logistic growth trajectories for three rate values, 0.1, 0.5, and 1.

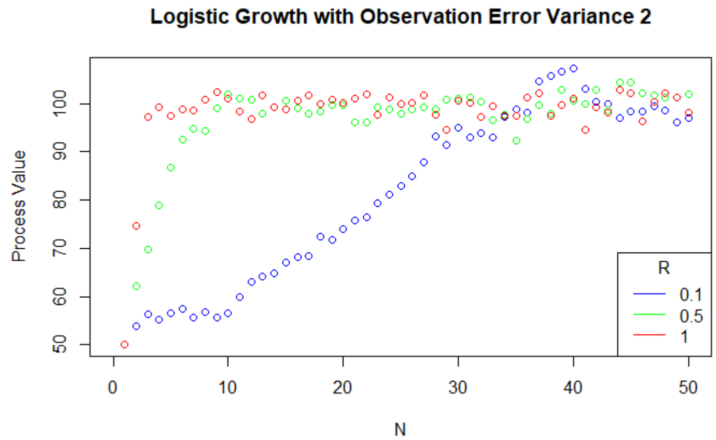
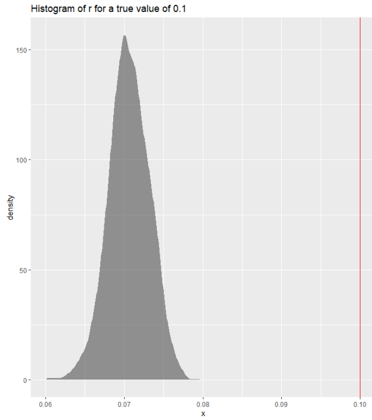


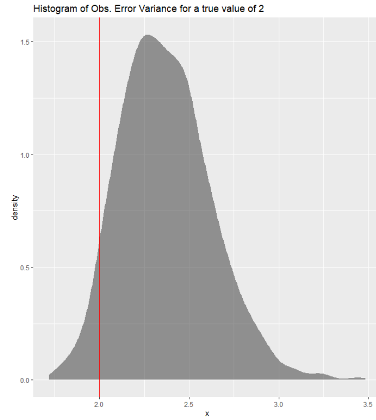
Figure 4: Logistic growth trajectories for three rate values, 0.1, 0.5, and 1 with additional normally-distributed observation error. The error variance was specified as $\sigma_{obs} = 2$.

Stan models were then implemented in order to estimate model parameters for the rate, r, and observation noise variance, σ_{obs}. The following plots demonstrate sample paths for 3 different r values, as well as histograms for the estimates of r and σ_{obs}. Simulations were conducted using 6 chains and 1000 samples each.

Figure 5: Logistic growth sample path for $r = 0.1$ and $\sigma_{obs} = 2$.



(a) Histogram of r estimates. The histogram's mean is around 0.07, while the true value is 0.1



(b) Histogram of σ_{obs} estimates. The histogram's mean is around 2.5, while the true value is 2.0

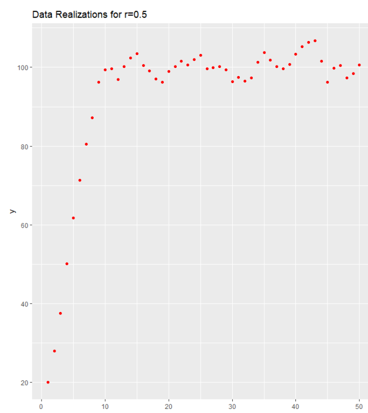
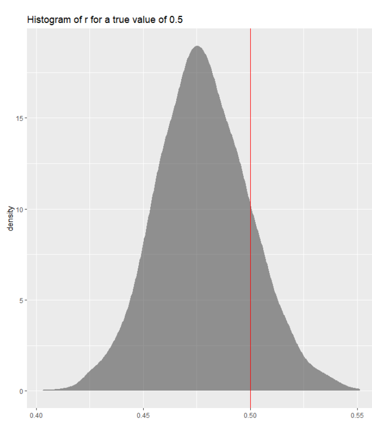
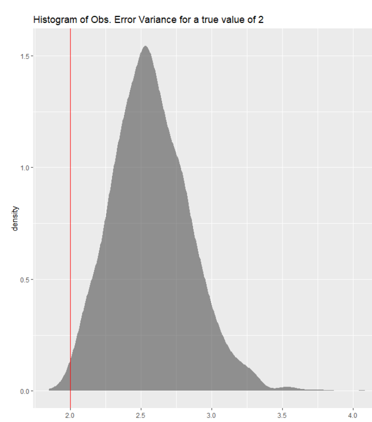


Figure 6: Logistic growth sample path for $r = 0.5$ and $\sigma_{obs} = 2$.



(a) Histogram of r estimates. The histogram's mean is around 0.47, while the true value is 0.5



(b) Histogram of σ_{obs} estimates. The histogram's mean is around 2.6, while the true value is 2.0

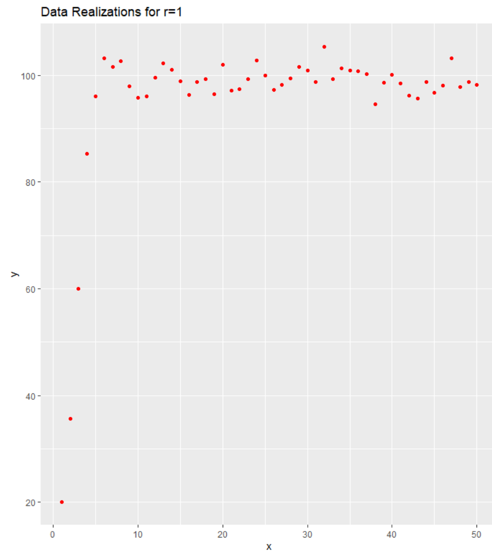
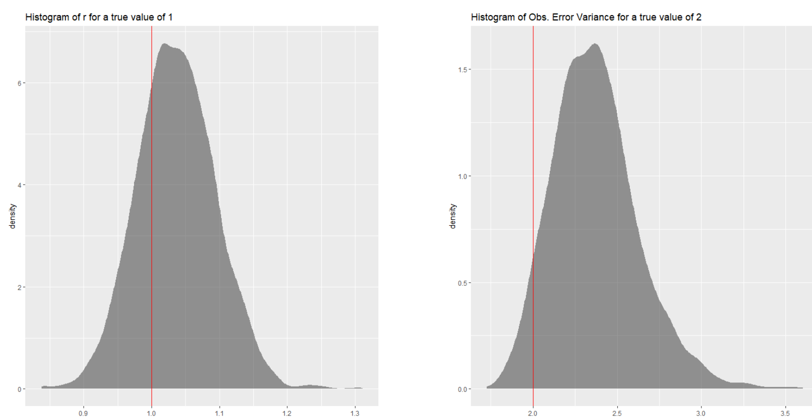


Figure 7: Logistic growth sample path for $r = 1$ and $\sigma_{obs} = 2$.



(a) Histogram of r estimates. The histogram's mean is around 0.105, while the true value is 1

(b) Histogram of σ_{obs} estimates. The histogram's mean is around 2.4, while the true value is 2.0

It can be observed that the estimates for r became more precise as r increased.

Question 3

Question 3 was an extension of question 2, where gamma-distributed process error was added to the models. Figures 8, 9, and 10 demonstrate 6 different sample paths for the logistic growth with process error only, process error and observation error, and process error with a fixed observation error. The process error was kept small, as the model would blow-up otherwise.

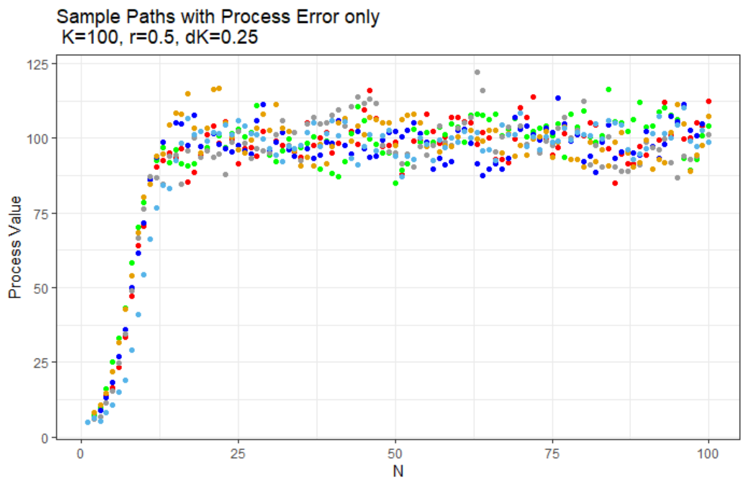


Figure 8: Logistic growth with gamma-distributed process error. The parameters of the model were set as $r = 0.5$, $K = 100$, and $dK = 0.25$ (dispersion constant).

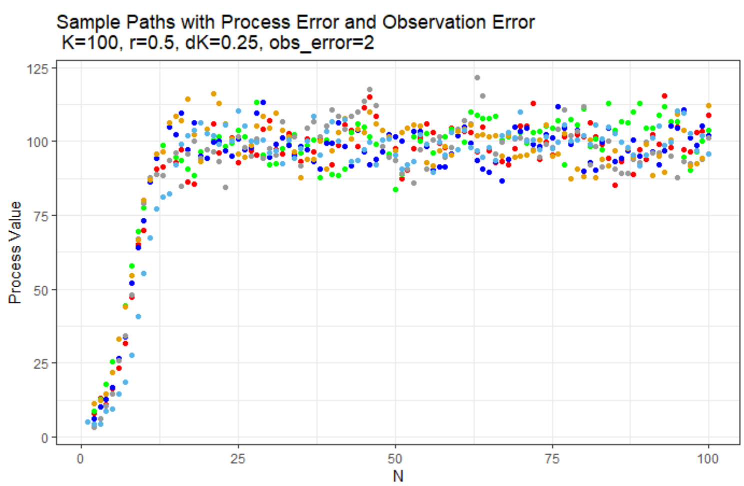


Figure 9: Logistic growth with gamma-distributed process error and normally-distributed observation error. The parameters of the model were set as $r = 0.5$, $K = 100$, $dK = 0.25$ (dispersion constant), and $\sigma_{obs} = 2$.

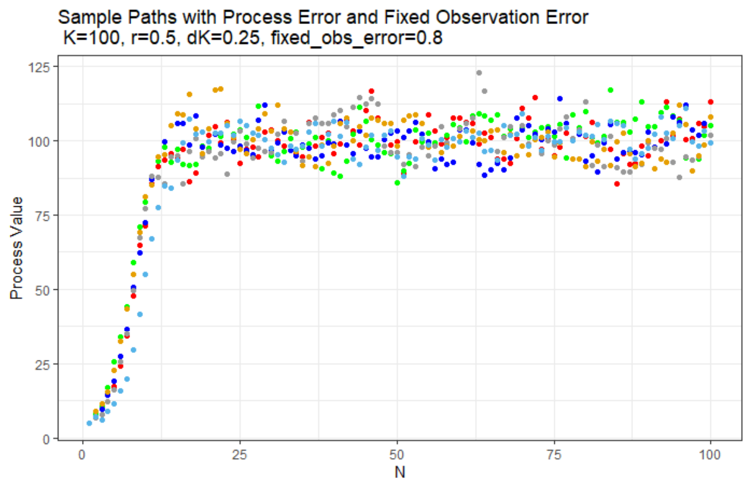


Figure 10: Logistic growth with gamma-distributed process error and normally-distributed observation error. The parameters of the model were set as $r = 0.5$, $K = 100$, $dK = 0.25$ (dispersion constant), and fixed $\sigma_{obs} = 0.8$.

At this point, I mostly implemented the code to run the Stan models. However, I had difficulty getting it to actually sample. (Unfortunately, I did not have a lot of time to investigate due to my upcoming project status meeting.)