

SYST/STAT 664: Homework Assignment 7

due April 3, 2023

Please make sure you mark clearly which question you are answering and that you explain how you arrived at your answer. Your response will be graded for correctness and clarity. Points may be deducted if you do not provide information on how you arrived at your answer. Upload your responses to Gradescope. Please submit R code either as a separate attachment on Blackboard or in your main submission.

1. This assignment continues with the pollution data from Assignment 6: HCB measurements taken at the bottom and the surface from Wolf River in Tennessee:

Surface	Bottom
3.74	5.44
4.61	6.88
4.00	5.37
4.67	5.44
4.87	5.03
5.12	6.48
4.52	3.89
5.29	5.85
5.74	6.85
5.48	7.16

As we did in Assignment 6, assume the observations are independent normal random variables with unknown depth-specific means Θ_s and Θ_b and precisions P_s and P_b . Assume independent improper reference priors for the surface and bottom parameters:

$$g(\theta_s, \theta_b, \rho_s, \rho_b) = g(\theta_s, \rho_s) g(\theta_b, \rho_b) \propto \rho_s^{-1} \rho_b^{-1}.$$

This prior can be treated as the product of two normal-gamma priors with $\mu_s = \mu_b = 0$, $k_s = k_b = 0$, $\alpha_s = \alpha_b = -1/2$, and $\beta_s = \beta_b = \infty$. (These are not valid normal-gamma distributions, but you can use the usual Bayesian conjugate updating rule to find the posterior distribution.) In Assignment 6, you found the joint posterior distribution for the parameters $(\theta_s, \theta_b, \rho_s, \rho_b)$. (Solutions to that assignment will be posted by mid-week for you to use in this assignment.) Find the posterior predictive distribution for the sample mean of a future sample of size 10 from the surface and a future sample of size 10 from the bottom. Find 95% credible intervals on the sample mean of each future sample. Repeat for future samples of size 40. Compare your results and discuss.

2. Use direct Monte Carlo to estimate the predictive distribution for the difference in the two sample means for 10 future surface and bottom samples. Clearly describe your procedure for drawing samples. Plot a kernel density estimator for the density function for the difference in means. Find a 95% credible interval for the difference in the two sample means. Repeat for future samples of 40 surface and 40 bottom observations. Comment on your results.
3. Repeat Problem 1, but use a model in which the standard deviation is known and equal to the sample standard deviation, and the depth-specific means Θ_s and Θ_b have a uniform prior distribution. Compare the 95% credible intervals for the future sample means for the known and unknown standard deviation models. Discuss.