STSB6816 Special Test of 2022

Mathematical Statistics and Actuarial Science; University of the Free State

2022/06/23

Time: 150 minutes; Marks: 40

Instructions

Answer all questions in a single R Markdown document. Please knit to Word or PDF at the end and submit both the PDF/Word document and the .Rmd file for assessment, in that order.

Label questions clearly, as it is done on this question paper.

All results accurate to at least 1 decimal place, ensure that simulation error is small enough (by doing enough simulations). Intervals should cover 95% probability unless stated otherwise.

Show all derivations, formulas, code, sources and reasoning.

No communication software, no devices, and no communication capable websites may be accessed prior to submission. You may not (nor even appear to) attempt to communicate or pass information to another student.

Question 1

Consider the Laplace density: $f(y|\mu, b) = \frac{1}{2b}\exp\left(-\frac{|y-\mu|}{b}\right)$.

1.1) Write down the log density. **[1]**

1.2) Derive the MDI prior. **[2]**

1.3) Derive the Jeffreys prior for b assuming $\mu = 0$. **[4]**

1.4) Now consider that you have a sample of values $y_1, \dots, y_n$ that are explained by a single continuous explanatory variable $x_1, \dots, x_n$ but with Laplace residuals. Write down the log likelihood and then show that the log posterior is given by the expression below. **[3]**

$$\pi(\boldsymbol{\theta}|\mathbf{y}, \mathbf{x}) = -(n+1)\log b - \frac{1}{b}\sum_{i=1}^{n}|y_i - (\beta_0 + \beta_1 x_i)|$$

1.5) What is the theoretical impact of using a Laplace distribution for the residuals instead of a normal distribution? **[3]**

1.6) Use the code below to capture a set of data, and show that the average y value is 5.2 **[2]**

```
d <- data.frame(x = 1:20, y = c(8.6, 5.6, 7.2, 6.9, 6.8, 5.4, 6.8, 5.9, 4.4, 5.9, 4.
8, 5.2, 3.7, 5.3, 3.4, 4, 4.2, 3.1, 3.6, 3.2))
mean(d$y)
```

1.7) Fit an OLS regression through the points as given. Give a summary of the parameters. **[2]**

1.8) Fit a Bayesian linear model with Laplace errors through the points as given. Give a trace plot and summary of the parameters. **[10]**

1.9) Draw a single plot showing the data points and both model fits, as well as prediction intervals for both models. **[10]**

HINT You can use posterior mean or median estimates to produce Laplace predicted values without intervals. Leave the Laplace intervals for last, after you have already done the plot and some interpretation, as they require you to simulate new Laplace values.

1.10) Explain, in detail, which model appears to be a better fit based on your plot. **[3]**