Final Exam
of
STAT 7330 Bayesian Data Analysis

Due: 11:59 pm

December 8 (Friday), 2023

DON'T TURN THE PAGE UNTIL YOU'RE READY TO BEGIN THE EXAM.

Instructions:

1. You will have one week to complete this exam.
2. You must use LaTeX or related tools (e.g. `Rmarkdown`) to type your answers and submit the paper in pdf format.
3. Open text and open notes. NO INTERNET.
4. Statistical software packages are permitted.
5. All necessary code should be provided.
6. Discussion with anyone excluding the instrctor is forbidden
7. You must state and sign the UTD Honor Code.
8. Under the Honor Code, you must NOT DISCUSS this exam with any other person except the instructor.

The Pledge: As a Comet, I pledge honesty, integrity, and service in all that I do.

Signature:

Printed name:

Problem 1: Weibull model with conjugate prior

The Weibull distribution is a distribution that is often used for lifetimes of equipment/parts. It actually has two parameters but for the moment let's assume that one of them is fixed. The Weibull(2) density is

$$f(x|\theta) = 2\theta x \exp(-\theta x^2)$$

for $x > 0$. The parameter θ is something like the "inverse lifetime" parameter. Large θ means short lifetimes, while small θ means long lifetimes. The mean of the distribution is $0.886\theta^{-\frac{1}{2}}$. Suppose we observe data $x_1, \ldots, x_n$ as independent samples from the Weibull(2) distribution.

Question 1.1 (8 points)

Show that the Gamma distribution, i.e. Ga(a, b), is the conjugate prior distribution and derive the posterior distribution. Use the p.d.f. $\pi(\theta) = \frac{b^a}{\Gamma(a)}\theta^{a-1}\exp(-b\theta)$.

Question 1.2 (8 points)

Derive the marginal distribution $\pi(x_1, \ldots, x_n)$. (Hint: There are two related ways to obtain the marginal, through $\pi(x_1, \ldots, x_n) = \int f(x_1, \ldots, x_n|\theta)\pi(\theta)d\theta$ or $\pi(x_1, \ldots, x_n) = f(x_1, \ldots, x_n|\theta)\pi(\theta)/\pi(\theta|x_1, \ldots, x_n)$)

Question 1.3 (4 points)

In one application, the lifetime of a kind of gear was measured in $1,000$ hours. Let $x = 1$ means the part lasted $1,000$ hours and $x = 1/2$ means the part lasted 500 hours. Gears tend to last between 500 and $5,000$ hours, with $1,000$ being a typical lifetime. This suggests a Gamma prior distribution for θ with $a = 1.4$ and $b = 2$. Obtain a graph of this prior density and argue that this choice of prior is reasonable given the information provided.

Question 1.4 (8 points)

Suppose that we observed $n = 10$ with

$$x = (0.25, 0.52, 0.60, 0.91, 0.97, 1.00, 1.07, 1.09, 1.18, 1.38).$$

Evidently this gear is fairly typical with lifetimes around $1,000$. Find and graph the posterior distribution. Give the posterior mean and variance. Give a 95% posterior interval for θ, using both analytical and numerical approaches.

Problem 2: Change-point model

Consider the following hierarchical change-point model for the number of occurrence Y_i of some event during time interval i:

$$Y_i \sim \begin{cases} \text{Poi}(\theta) & i = 1, \ldots, k \\ \text{Poi}(\lambda) & i = k+1, \ldots, n \end{cases},$$

with the following prior formulation: $\theta \sim \text{Ga}(a_1, \beta_1)$ and $\lambda \sim \text{Ga}(a_2, \beta_2)$ with θ and λ independent. We further assume $\beta_1 \sim \text{IG}(c_1, d_1)$ and $\beta_2 \sim \text{IG}(c_2, d_2)$. The p.d.f.'s of those priors are defined as below:

$$\begin{aligned}
\pi(\theta|a_1, \beta_1) =& \frac{1}{\beta_1^{a_1}\Gamma(a_1)}\theta^{a_1-1}\exp\left(-\frac{\theta}{\beta_1}\right)\\
\pi(\lambda|a_2, \beta_2) =& \frac{1}{\beta_2^{a_2}\Gamma(a_2)}\lambda^{a_2-1}\exp\left(-\frac{\lambda}{\beta_2}\right)\\
\pi(\beta_1|c_1, d_1) =& \frac{1}{d_1^{c_1}\Gamma(c_1)}\left(\frac{1}{\beta_1}\right)^{c_1+1}\exp\left(-\frac{1}{d_1\beta_1}\right)\\
\pi(\beta_2|c_2, d_2) =& \frac{1}{d_2^{c_2}\Gamma(c_2)}\left(\frac{1}{\beta_2}\right)^{c_2+1}\exp\left(-\frac{1}{d_2\beta_2}\right).
\end{aligned}$$

Question 2.1 (10 points)

Fit this model the "coal-data.txt", which gives counts of coal mining disasters in Great Britain by year from 1851 to 1962, where disaster is defined as an accident resulting in the deaths of 10 or more minors. Set the hyperparameters to be reasonably non-informative to: $a_1 = a_2 = 0.5$, $c_1 = c_2 = 1$, and $d_1 = d_2 = 1$. Assume $k = 40$ (corresponding to year 1890). Derive full conditional distributions for θ and λ.

Question 2.2 (10 points)

Implement a Gibbs sampler to derive the posterior distributions of θ, λ, and $R = \theta/\lambda$. Provide the histograms and kernel density estimates along with the posterior sample summaries: means, standard deviations, quantiles etc. What is meant by "convergence diagnosis"? Describe some tools you might use to assist in this regard. Present some evidence that the MCMC has reasonably converged.

Question 2.3 (10 points)

Now assume that k is unknown and adopt the following prior:

$$k \sim \text{Discrete-Unif}(1, n),$$

which should be independent of θ and λ. Now add k into the sampling chain, and obtain the marginal posterior density estimate for it. What is the effect on the posterior for $R = \theta/\lambda$?

Question 2.4 (10 points)

Now replace the third-stage priors given above with, $\beta_1 \sim \text{Ga}(c_1, d_1)$ and $\beta_2 \sim \text{Ga}(c_2, d_2)$ with β_1 and β_2 being independent, thus destroying the conjugacy for these two conditionals. Resort to a Metropolis Hastings subsampling for these two components instead, using the following hyperparameter values $c_1 = c_2 = 1$ and $d_1 = d_2 = 100$. What is the effect on the posterior distribution for β_1 and β_2. For $R = \theta/\lambda$? For k?

Problem 3: Model comparison on regression models

For the following data we consider two competing models: H_1: a linear vs. H_2: a quadratic regression.

x_i	-1.9	-0.39	0.79	-0.20	0.42	-0.35	0.67	0.63	-0.024	1.2
y_i	-1.7	-0.23	0.50	-0.66	1.97	0.10	0.60	1.13	-0.943	2.6

Model H_1:

$$\begin{aligned} y_i =& \beta_1 + \beta_2 x_i + \epsilon_i, \quad i = 1, \dots, n \\ \beta_1 \sim& \mathrm{N}(0,1) \\ \beta_2 \sim& \mathrm{N}(1,1) \\ \epsilon_i \sim& \mathrm{N}(0,1) \end{aligned}$$

with β_1 and β_2 *a priori* independent.

Model H_2:

$$\begin{aligned} y_i =& \gamma_1 + \gamma_2 x_i + \gamma_3 x_i^2 + \epsilon_i, \quad i = 1, \dots, n \\ \gamma_1, \gamma_3 \sim& \mathrm{N}(0,1) \\ \gamma_2 \sim& \mathrm{N}(1,1) \\ \epsilon_i \sim& \mathrm{N}(0,1) \end{aligned}$$

with γ_1, γ_2, and γ_3 *a priori* independent.

Question 3.1 (7 points)

Find the marginal distributions $\pi(y_1, \dots, y_n | H_1) = \int f(y_1, \dots, y_n | \beta_1, \beta_2) \pi(\beta_1) \pi(\beta_2) d\beta_1 d\beta_2$

Question 3.2 (7 points)

Find the marginal distributions $\pi(y_1, \dots, y_n | H_2) = \int f(y_1, \dots, y_n | \gamma_1, \gamma_2, \gamma_3) \pi(\gamma_1) \pi(\gamma_2) \pi(\gamma_3) d\gamma_1 d\gamma_2 d\gamma_3$

Question 3.3 (8 points)

Write down the Bayes factor $B = f(y_1, \dots, y_n | H_2) / f(y_1, \dots, y_n | H_1)$ for comparing the two models and evaluate it for the given data set.

Question 3.4 (10 points)

We now replace the prior distribution by improper constant priors: $\pi(\beta_1) = \pi(\beta_2) = c_1$ in model H_1 and $\pi(\gamma_1) = \pi(\gamma_2) = \pi(\gamma_3) = c_2$ in model H_2. We can still obtain the marginal distributions and define a Bayes factor. Show that the value of the Bayes factor depends on both c_1 and c_2.