INDIAN STATISTICAL INSTITUTE

M.Stat. 2nd Year

BAYESIAN INFERENCE

ASSIGNMENT I

1. Consider the problem of estimation of a real parameter θ with the loss function

$$\begin{aligned} L(\theta, a) &= K_0(\theta - a) && \text{if } \theta - a \geq 0, \\ & K_1(a - \theta) && \text{if } \theta - a < 0. \end{aligned}$$

Show that the Bayes estimate is given by the quantile of order $K_0/(K_0 + K_1)$ of the posterior distribution (assume, for simplicity, uniqueness of the quantile).

2. Given $0 < \theta < 1$, let $X_1, \ldots, X_n$ be i.i.d. $Bin(1, \theta)$. Consider the Jeffreys prior for θ. Find by simulation the frequentist coverage of θ by the two-tailed 95% credible interval for $\theta = \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, \frac{7}{8}$. Do the same for the usual frequentist interval $\hat{\theta} \pm z_{0.025}\sqrt{\hat{\theta}(1-\hat{\theta})/n}$ where $\hat{\theta} = \sum X_i/n$.

3. Let $X_1, \ldots, X_n$ be i.i.d. $\sim f(x|\theta)$, $\theta \in R$ and $\pi(\theta)$ be a prior density of θ. Use a result proved in the class (see proof of asymptotic normality of posterior distribution) to rigorously prove that

$$\begin{aligned} \log \int_R \prod_1^n f(X_i|\theta)\pi(\theta)d\theta &= \sum_{i=1}^n \log f(X_i|\hat{\theta}_n) - \frac{1}{2}\log n + \frac{1}{2}\log(2\pi) \\ & \quad -\frac{1}{2}\log I(\theta_0) + \log \pi(\theta_0) + o_p(1) \end{aligned}$$

as $n \to \infty$, where $\hat{\theta}_n$ is the MLE and $I(\theta_0)$ is Fisher information number at θ_0.

4. Show that the result on asymptotic normality of the posterior distribution of $\sqrt{n}(\theta - \hat{\theta}_n)$, proved in the class, implies consistency of the posterior distribution of θ at θ_0.

5. Show that Condition (A4), used to prove asymptotic normality of posterior distribution, holds when $X_1, \ldots, X_n$ are i.i.d. $N(\theta, 1)$.

6. Let $X_1, \ldots, X_n$ be i.i.d. with a Cauchy density

$$f(x|\theta) = \frac{1}{\pi[1 + (x - \theta)^2]}, \quad -\infty < x < \infty, \ \theta > 0.$$

We want to find the $100(1 - \alpha)\%$ credible set for θ.

Draw n (your choice) observations from this distribution with a chosen θ. Based on these observations, find 95% and 99% HPD credible sets for θ. Do this for three (or more) different values of n (small, moderately large, large/very large). Also describe your algorithm for finding the HPD credible sets.

7. Let $X_1, \ldots, X_n$ be i.i.d. $\sim N(\mu, \sigma^2)$, μ, σ^2 both unknown. Consider the set up of Jeffreys test (see class note).

(a) Show that if $\bar{X} \to \infty$ and s^2 is bounded, BF_{01} goes to zero for the Cauchy prior but does not go to zero for normal prior.

(b) Consider Cauchy prior for this problem. Using the representation of the Cauchy density $g_1(\mu|\sigma)$ as a scale mixture of normals, express the integrated likelihood under H_1 as a one-dimensional integral (over the mixing variable τ).

8. (a) *Welch's paradox.* Let X_1, X_2 be i.i.d. $\sim U(\theta - 1/2, \theta + 1/2)$, $\theta \in R$. A frequentist 95% confidence interval is $(\bar{X} - 0.3882, \bar{X} + 0.3882)$ where $\bar{X} = (X_1 + X_2)/2$. Show that if X_1 and X_2 are sufficiently apart, say $X_1 - X_2 > d$ (find d) then θ must be in this confidence interval (but a frequentist reports the confidence level as only 95%).

Calculate P(the interval $\bar{X} \mp 0.3882$ covers θ $|X_1 - X_2)$. Also find the posterior distribution of θ with the objective prior $\pi(\theta) \equiv 1$ and find an appropriate 95% credible interval for θ.

(b) Let X_1, X_2 be i.i.d. with a common density belonging to a location parameter family of densities with a location parameter θ. Assume without loss of generality that $E_\theta X_1 = \theta$. One can find a frequentist 95% confidence interval of the form $(\bar{X} - c, \bar{X} + c)$. Suppose now that $X_1 - X_2$ is known and one calculates P(the interval $\bar{X} \mp c$ covers θ $|X_1 - X_2)$. When can Welch's paradox occur in such a scenario?

Can Welch's paradox occur if X_1, X_2 are i.i.d. $N(\theta, 1)$? (Explain.)

9. Let $X_1, \ldots, X_m$ and Let $Y_1, \ldots, Y_n$ be two independent random samples from $N(\mu_1, \sigma^2)$ and $N(\mu_2, \sigma^2)$ respectively. Assume that the prior distribution of $(\mu_1, \mu_2, \log \sigma^2)$ is improper uniform where (μ_1, μ_2, σ^2) are independent. Find the posterior distribution of $\mu_1 - \mu_2$.

10. Let $X_1, \ldots, X_n$ be i.i.d. $\sim N(\theta, \sigma^2)$, σ^2 known. Assume $\sigma^2 = 1$.

(a) Consider the problem of testing $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$. We reject H_0 if $T = \sqrt{n}(\bar{X} - \theta_0)$ is large. A classical (frequentist) measure of evidence against H_0 is the P-value defined by

$$P = \sup_{\theta \leq \theta_0} P_\theta[\sqrt{n}(\bar{X} - \theta_0) > t]$$

where t is the observed value of T (We reject H_0 at level α if $P \leq \alpha$). Find the P-value (in terms of t).

Consider now the uniform prior $\pi(\theta) \equiv 1$. Find the posterior probability of H_0 (note that it is the same as the P-value).

(b) Suppose we want to test $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$. We reject H_0 if $T = |\sqrt{n}(\bar{X} - \theta_0)|$ is large. Here P-value $= P_{\theta_0}[|\sqrt{n}(\bar{X} - \theta_0)| > t]$ where t is the observed value of T. Find the P-value in terms of t.

Consider now a $N(\theta_0, 1)$ prior for θ under H_1 and find the Bayes factor BF_{01}. Assuming prior probabilities $P(H_0) = P(H_1) = \frac{1}{2}$, find the posterior probability $P(H_0|X_1, \ldots, X_n)$. When $n = 50, t = 1.960$, show that P-value $= 0.05$, $BF_{01} = 1.08$ and $P(H_0|X_1, \ldots, X_n) = 0.52$. (This shows a conflict between frequentist and Bayesian answers.)

11. Let the sample space be $\{1, 2, \ldots, k\}$ and $P = (p_1, \ldots, p_k)$ be a random probability distribution on this sample space. Let $X_1, \ldots, X_n$ be i.i.d. $\sim P$, and $P \sim$ Dirichlet$(\alpha_1, \ldots, \alpha_k)$, $\alpha_i > 0\ \forall i$. Show that for any subset A of the sample space, the posterior mean of $P(A)$ is a weighted average of its prior mean and $P_n(A)$ where P_n denotes the empirical distribution of $X_1, \ldots, X_n$.