Problem class 2

Lecturer: Georgios Karagiannis georgios.karagiannis@durham.ac.uk

Exercise 1. ($\star\star$)Consider the Bayesian model

$$\begin{cases} x_i|\theta & \overset{\text{iid}}{\sim} \text{Ga}(\alpha,\beta),\ \forall i=1,...,n \\ (\alpha,\beta) & \sim \Pi(\alpha,\beta) \end{cases}$$

where $\text{Ga}(a,\beta)$ is the Gamma distribution with expected value α/β. Specify a Jeffrey's prior for $\theta=(\alpha,\beta)$.

Hint-1: Gamma distr.: $x \sim \text{Ga}(a,b)$ has pdf $f(x) = \frac{b^a}{\Gamma(a)} x^{a-1}\exp(-bx)1_{(0,+\infty)}(x)$, and Expected value $\text{E}_{\text{Ga}}(x|a,b)=\frac{a}{b}$

Hint-2: You may also need that the second derivative of the logarithm of a Gamma function is the 'polygamma function of order 1'. Ie,

- $F^{(0)}(\alpha) = \frac{\text{d}}{\text{d}\alpha}\log(\Gamma(a))$
- $F^{(1)}(\alpha) = \frac{\text{d}^2}{\text{d}\alpha^2}\log(\Gamma(a))$

Hint-3: You may leave your answer in terms of function $F^{(1)}(\alpha)$.

Exercise 2. ($\star\star$)Consider observables $x=(x_1,...,x_n)$. Consider the Bayesian model

$$\begin{cases} x_i|\theta & \overset{\text{IID}}{\sim} \text{N}(\theta,1), \quad i=1,...,n \\ \theta & \sim \Pi(\theta) \end{cases}$$

where $\pi(\theta)\propto 1$ and that we have only one observable. Consider the LINEX loss function

$$\ell(\theta,\delta) = \exp(c(\theta-\delta)) - c(\theta-\delta) - 1$$

1. Show that $\ell(\theta,\delta)\geq 0$
2. Find the Bayes estimator $\hat{\delta}$ under LINEX loss function and under the given Bayesian model.

Hint-1: Random variable B follows a log-normal distribution $B\sim \text{LN}(\mu_A,\sigma_A^2)$ with parameters μ_A,σ_A^2 if $B=\exp(A)$ where $A\sim \text{N}(\mu_A,\sigma_A^2)$.

Hint-2: If $B\sim \text{LN}(\mu_A,\sigma_A^2)$ then $\text{E}_{\text{LN}(\mu_A,\sigma_A^2)}(B)=\exp(\mu_A+\frac{\sigma_A^2}{2})$.

Hint-3: It is

$$-\frac{1}{2}\frac{(\mu-\mu_1)^2}{v_1^2} - \frac{1}{2}\frac{(\mu-\mu_2)^2}{v_2^2} ... - \frac{1}{2}\frac{(\mu-\mu_n)^2}{v_n^2} = -\frac{1}{2}\frac{(\mu-\hat{\mu})^2}{\hat{v}^2} + C$$

where

$$\hat{v}^2 = \left(\sum_{i=1}^{n}\frac{1}{v_i^2}\right)^{-1}; \quad \hat{\mu} = \hat{v}^2\left(\sum_{i=1}^{n}\frac{\mu_i}{v_i^2}\right); \quad C = \frac{1}{2}\frac{\hat{\mu}^2}{\hat{v}^2} - \frac{1}{2}\sum_{i=1}^{n}\frac{\mu_i^2}{v_i^2}$$

Exercise 3. ($\star\star$)Suppose we wish to estimate the values of a collection of discrete random variables $\vec{X}=X_1,\dots,X_n$. We have a posterior joint probability mass function for these variables, $p(\vec{x}|y)=p(x_1,\dots,x_n|y)$ based on some data y. We decide to use the following loss function:

$$\ell(\hat{\vec{x}}, \vec{x}) = \sum_{i=1}^{n} (1 - \delta(\hat{x}_i, x_i)) \tag{1}$$

where $\delta(a, b) = 1$ if $a = b$ and zero otherwise.

1. Derive an expression for the estimated values, found by minimizing the expectation of the loss function. [Hint: use linearity of expectation.]

2. When the probability distribution is a posterior distribution in some problem, this type of estimate is sometimes called 'maximum posterior marginal' (MPM) estimate. Explain why this name is appropriate.

3. Explain in words what the loss function is measuring. Compare with the loss function for MAP estimation.

Exercise 4. (Example from the Lecture's handout) Consider a Bayesian model

$$\begin{cases} y_i|\mu & \overset{\text{iid}}{\sim} \mathrm{N}_d(\mu, \Sigma), \qquad i = 1, ..., n \\ \mu & \sim \mathrm{N}_d(\mu_0, \Sigma_0) \end{cases}$$

where uncertain $\mu \in \mathbb{R}^d$, $d \geq 1$, and known Σ, μ_0, Σ_0. Find the C_a parametric HPD credible set for μ.

Hint-1: If $z = (z_1, ..., z_d)^\top$ such as $z_j \overset{\text{iid}}{\sim} \mathrm{N}(0, 1)$ for $j = 1, ..., d$, and $\xi = z^\top z = \sum_{j=1}^{d} z_j^2$, then $\xi \sim \chi^2_d$

Hint-2: It is

$$-\frac{1}{2}\sum_{i=1}^{n}(x-\mu_i)^\top \Sigma_i^{-1}(x-\mu_i)) = -\frac{1}{2}(x-\hat{\mu})^\top \hat{\Sigma}^{-1}(x-\hat{\mu})) + C(\hat{\mu}, \hat{\Sigma}) \quad ;$$

$$\hat{\Sigma} = (\sum_{i=1}^{n} \Sigma_i^{-1})^{-1}; \quad \hat{\mu} = \hat{\Sigma}(\sum_{i=1}^{n} \Sigma_i^{-1}\mu_i);$$

$$C(\hat{\mu}, \hat{\Sigma}) = \underbrace{\frac{1}{2}(\sum_{i=1}^{n} \Sigma_i^{-1}\mu_i)^\top (\sum_{i=1}^{n} \Sigma_i^{-1})^{-1} (\sum_{i=1}^{n} \Sigma_i^{-1}\mu_i) - \frac{1}{2}\sum_{i=1}^{n} \mu_i^\top \Sigma_i^{-1}\mu_i}_{=\text{independent of } x}$$