Bayesian Modelling – Problem Sheet 2

Please hand in your solutions for Problems 3-6 by 6pm on Wednesday 20/02/2019

Problem 1 (Jacobian formula in one dimension)

Let $X \sim f_X$ be a real-valued random variable having a continuous probability density function f_X, $g : \mathbb{R} \to \mathbb{R}$ be a one-to-one and continuously differentiable mapping and $Y = g^{-1}(X)$. Show that $Y \sim f_Y$ where $f_Y : \mathbb{R} \to \mathbb{R}_+$ is defined by

$$f_Y(y) = \Big|\frac{\mathrm{d}}{\mathrm{d}y} g(y)\Big| f_X(g(y)), \quad y \in \mathbb{R}.$$

Problem 2

Let $x_1, \dots, x_n$ be n observations that we model as independent random variables such that

$$X_k \sim \tilde{f}(x_1|\theta), \quad k = 1, \dots n$$

with $\theta \in \Theta \subset \mathbb{R}^d$ an unknown parameter.

Let $I_n(\theta)$ be the Fisher information matrix of this model (we assume that $\tilde{f}$ is such that $I_1(\theta)$ is well-defined). Show that $I_n(\theta) = nI_1(\theta)$ and conclude that the Jeffreys prior for this model does not depend on the sample size n.

Problem 3

Let $x \in \mathbb{R}^d$ be an observation that we model as a $\mathcal{N}_d(\theta, I_d)$ random variable, with $\theta \in \Theta := \mathbb{R}^d$ and I_d the d-dimensional identity matrix. Let $\pi(\theta)$ be the Laplace's prior for this model.

1. Show that the Laplace's prior is also the Jeffreys prior for this model.
2. Compute $\mathbb{E}_\pi[\theta|x]$, the posterior expectation of θ given x.
3. Assume now that we are interested in the parameter $\eta = \|\theta\|^2$. Show that $\mathbb{E}_\pi[\eta|x] = \|x\|^2 + d$.
4. For $c \in \mathbb{R}$ let $\delta_c : \mathbb{R}^d \to \mathbb{R}_+$ be the estimator of η defined by

$$\delta_c(z) = \|z\|^2 + c, \quad z \in \mathbb{R}^d.$$

   Show that, under the quadratic loss function and for any $\theta \in \mathbb{R}^d$, the frequentist risk $R(\|\theta\|^2, \delta_c)$ is minimized for $c = -d$. Deduce that

$$R(\|\theta\|^2, \delta_{-d}) < R(\|\theta\|^2, \delta_d), \quad \forall \theta \in \Theta$$

   and thus that the estimator $\mathbb{E}_\pi[\eta|\cdot]$ is not admissible.
5. Show that, for any $c \in \mathbb{R}$, we have $\sup_{\theta \in \Theta} R(\|\theta\|^2, \delta_c) = +\infty$.

Problem 4

Let $x_1, \dots, x_n$ be n observations that we model as independent Poisson random variables,

$$X_k|\theta \sim \text{Poisson}(\theta), \quad k = 1, \dots n,$$

where $\theta \in \Theta := \mathbb{R}_{>0}$ is an unknown parameter. The likelihood function of observation $x_1 \in \{0\} \cup \mathbb{N}$ is therefore given by

$$\tilde{f}(x_1|\theta) = \frac{\theta^{x_1}}{x_1!} e^{-\theta}, \quad \theta \in \Theta.$$

We assign to θ a $\text{Gamma}(\alpha_0, \beta_0)$ prior distribution, i.e.

$$\pi(\theta) = \frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)} \theta^{\alpha_0 - 1} e^{-\beta_0 \theta}, \quad \theta \in \Theta,$$

where $\alpha_0, \beta_0 \in (0, +\infty)$ are positive hyper-parameters. Below we denote by $x = (x_1, \dots, x_n)$ the vector of observations.

1. Derive the posterior distribution of θ. Is the prior conjugate? Explain your answer.
2. Compute the posterior mean and posterior variance of θ.
3. Using the observation x and a quadratic loss function, compute the estimate of the loss $L(\theta, \mathbb{E}_\pi[\theta|x])$ obtained by minimizing the posterior expected loss.
4. Consider now the Laplace's prior defined by

$$\pi(\theta) \propto 1, \quad \theta \in \Theta.$$

   a) Derive the posterior mean $\mathbb{E}_\pi[\theta|x]$ and the posterior variance. How are these quantities related to those obtained in part 2?

   b) Show that, for the quadratic loss function, the integrated risk of the estimator $\delta^\pi := \mathbb{E}_\pi[\theta|\cdot]$ is infinite; that is, $r(\pi, \delta^\pi) = +\infty$.

   Hint: The following result may be useful: If $Y_1, \dots, Y_n$ are independent $\text{Poisson}(\lambda)$ *random variables then $\sum_{k=1}^n Y_i \sim \text{Poisson}(n\lambda)$ (with $\lambda > 0$).*

5. We now consider the Jeffreys' prior for this model.

   a) Compute the Jeffreys' prior for this model. Is this prior proper or improper?

   b) Derive the posterior mean and posterior variance under this prior.

Problem 5

The probability of success θ in a Bernoulli trial lies in the interval $\Theta := [0, 1]$, so if we are completely ignorant of its true value the uniform distribution on the unit interval seems to be the natural prior to use:

$$\pi(\theta) = \mathbb{1}_{[0,1]}(\theta), \quad \theta \in \Theta. \tag{1}$$

But if we are completely ignorant of θ, we are also completely ignorant of

$$\eta := \frac{\theta}{1-\theta}$$

which takes values in the positive real line.

1. Compute $\pi^*(\eta)$, the prior distribution for $\eta \in \mathbb{R}_+$ which is implied by the prior distribution on θ specified in (1). Check that $\pi^*(\eta)$ is a proper density function on $\mathbb{R}_+$.

2. For $a \geq 0$ compute $\pi^*(\{\eta : \eta \leq a\})$; that is, the probability that $\eta \leq a$ under the prior $\pi^*(\eta)$. Compute this probability for $a = 3$.

3. Using the results obtained in parts 1 and 2, comment on whether the prior density $\pi^*(\eta)$ can be interpreted as a non-informative prior density.

Problem 6

Observations $x_1, \dots, x_n$ are modelled as independent Geometric random variables,

$$X_k|\theta \sim \text{Geom}(\theta) \quad \text{for } k = 1, \dots, n$$

where $\theta \in \Theta := [0, 1]$ is an unknown parameter. Note that $x_1 \in \{1, 2, \dots\}$ and that the likelihood function of observation x_1 is given by

$$\tilde{f}(x_1|\theta) = (1-\theta)^{x_1 - 1}\theta, \quad \theta \in \Theta.$$

1. Derive the Jeffreys' prior for this model. Is the prior proper or improper?

2. Show that the Jeffreys' prior yields a valid posterior distribution. Compute this latter as well as the posterior mean.