In this MidTerm there are three questions:

- Q1: 20 Marks
- Q2: 20 Marks
- Q3: 10 Marks

Q1. Consider realization $Y_1, \ldots, Y_n$ from an infinitely exchangeable binary sequence where, in de Finetti's representation, prior distribution $\pi_0(\theta) \equiv Beta(\alpha_0, \alpha_0)$ on the parameter θ is specified. Let $y_1, \ldots, y_n$ be the corresponding observed values. The posterior distribution will be denoted π_n.

(a) Find $\pi_n(\theta)$. 6 Marks

(b) Suppose that a limiting version of the prior in (a) is used, with $\alpha_0 \longrightarrow 0$. Find the posterior mode for the resulting posterior distribution. 4 Marks

(c) If $\alpha_0 = 2$, find the posterior predictive distribution, $p_n(y)$, for a single future data point. 6 Marks

(d) Compute the variance of the individual Y_i in this exchangeable model. 4 Marks

Solution:

(a) We have that the likelihood is

$$\mathcal{L}_n(\theta) = \prod_{i=1}^{n} \theta^{y_i}(1-\theta)^{1-y_i} = \theta^{s_n}(1-\theta)^{n-s_n}$$

and with a $Beta(\alpha_0, \alpha_0)$ prior, this means that 6 Marks

$$\pi_n(\theta) \propto \theta^{s_n+\alpha_0-1}(1-\theta)^{n-s_n+\alpha_0+1} \equiv Beta(s_n+\alpha_0, n-s_n+\alpha_0)$$

(b) If $\alpha_0 \longrightarrow 0$, we have that the posterior is proper provided $s_n > 1$ and $n > s_n$, so $n > 2$. On the log scale, π_n becomes

$$(s_n - 1)\log\theta + (n - s_n - 1)\log(1-\theta) = a\log\theta + b\log(1-\theta)$$

say. The derivative of the $\log \pi_n(\theta)$ wrt θ is $a/\theta - b/(1-\theta)$, and equating this to zero to get the mode yields

$$a - a\theta = b\theta \quad \therefore \quad \theta = \frac{a}{a+b} = \frac{s_n - 1}{n-2}$$

provided $n > 2$ and $s_n > 1$. 4 Marks

(c) We have that for $y = 0, 1$

$$\begin{aligned}
p_n(y) &= \int_0^1 f_Y(y;\theta)\pi_n(\theta)\, d\theta \\
&= \int_0^1 \theta^y(1-\theta)^{1-y} \frac{\Gamma(n+4)}{\Gamma(s_n+2)\Gamma(n-s_n+2)} \theta^{(s_n+2)-1}(1-\theta)^{(n-s_n+2)-1}\, d\theta \\
&= \frac{\Gamma(n+4)}{\Gamma(s_n+2)\Gamma(n-s_n+2)} \frac{\Gamma(s_n+y+2)\Gamma(n-s_n+1-y+2)}{\Gamma(n+5)} \\
&= \begin{cases} \dfrac{n-s_n+2}{n+4} & y = 0 \\ \dfrac{s_n+2}{n+4} & y = 1 \end{cases}
\end{aligned}$$

and zero otherwise. 6 Marks

(d) We have for any α_0

$$\mathbb{E}[Y_i] = \mathbb{E}_{\pi_0}[\theta] = \frac{\alpha_0}{2\alpha_0} = \frac{1}{2} \qquad \mathbb{E}[Y_i^2] \equiv \mathbb{E}[Y_i] = \frac{1}{2}$$

so therefore $\text{Var}[Y_i] = \frac{1}{4}$. 4 Marks

Q2. A Bayesian analysis based on the conditional model

$$f_Y(y;\theta) = \theta e^{-\theta y} \qquad y > 0$$

and zero otherwise, for $\theta > 0$, is to be carried out based on a random sample of size n.

(a) Find a conjugate prior for this problem. 4 Marks

(b) Find the Jeffreys prior for this problem. 6 Marks

(c) If

$$\pi_0(\theta) = 2e^{-2\theta} \qquad \theta > 0$$

find the posterior distribution for θ if a single data point y_1 is observed. 6 Marks

(d) Suppose that ϕ denotes the median of the distribution $f_Y(y;\theta)$. Explain how to compute the posterior distribution for ϕ. You do not have to perform the computation. 4 Marks

Solution:

(a) Here the likelihood is based on the $Exponential(\theta)$ model

$$\mathcal{L}_n(\theta) = \theta^n e^{-s_n\theta}$$

so a conjugate prior must take the form

$$\pi_0(\theta) \propto \theta^a e^{-b\theta} \equiv Gamma(\alpha_0, \beta_0)$$

say. 4 Marks

(b) Here we have, in the usual notation,

$$\ell(y;\theta) = \log\theta - y\theta$$
$$\dot{\ell}(y;\theta) = 1/\theta - y$$
$$\ddot{\ell}(y;\theta) = -1/\theta^2$$

So therefore the Jeffreys prior is

$$\pi_0(\theta) \propto \{1/\theta^2\}^{1/2} = \frac{1}{\theta}$$

as in the scale family case. 6 Marks

(c) With this prior, we have

$$\pi_n(\theta) \propto \theta e^{-\theta y_1} \times e^{-2\theta} = \theta e^{-(y_1+2)\theta} \equiv Gamma(2, y_1+2)$$

6 Marks

(d) This is a simple transformation exercise; $\pi_n(\theta)$ yields the posterior for ϕ including the Jacobian. In this Exponential case, the median ϕ solves

$$1 - e^{-\theta\phi} = \frac{1}{2} \qquad \Longrightarrow \qquad \phi = \frac{\log 2}{\theta}$$

so that

$$\pi_n(\log 2/\phi)\frac{\log 2}{\phi^2} \qquad \phi \in \mathbb{R}^+$$

4 Marks

Q3. A Bayesian analysis based on the $Normal(0, \theta)$ conditional model

$$f_Y(y;\theta) = \left(\frac{1}{2\pi\theta}\right)^{1/2} \exp\left\{-\frac{1}{2\theta}y^2\right\} \quad y \in \mathbb{R}$$

is to be carried out based on a random sample of size n.

(a) Find a sufficient statistic for this problem. 4 Marks

(b) Find the posterior distribution under a conjugate prior specification. 6 Marks

Solution:

(a) Here the likelihood is

$$\mathcal{L}_n(\theta) = c\left(\frac{1}{\theta}\right)^{n/2} \exp\left\{-\frac{1}{\theta}\frac{\sum_{i=1}^{n} y_i^2}{2}\right\}$$

so therefore a sufficient statistic is

$$T_n = \sum_{i=1}^{n} Y_i^2.$$

4 Marks

(b) Inspection of $\mathcal{L}_n$ leads us to conclude that we must have

$$\pi_0(\theta) \propto \left(\frac{1}{\theta}\right)^{a} \exp\left\{-\frac{b}{\theta}\right\} \equiv InvGamma(\alpha_0/2, \beta_0/2)$$

say, so that

$$\pi_n(\theta) \equiv InvGamma(\alpha_n/2, \beta_n/2)$$

with

$$\alpha_n = \alpha_0 + n \qquad \beta_n = \beta_0 + t_n.$$

6 Marks