Foundations of Statistical Inference, BS2a, Exercises 3

1. The number of phone calls a man receives in a week follows a Poisson distribution with mean θ. At the start of week 1, the man's opinion about the value of θ corresponds to the gamma distribution

$$\pi(\theta) = \frac{1}{54}\theta^2 e^{-\theta/3},\ \theta > 0.$$

In the 4 weeks following the start of week 1, the man received 3, 7, 6, and 10 phone calls, respectively. Determine the posterior distribution of θ and the predictive distribution for the number of calls that he will receive in week 5.

2 Consider a vector of observations X with sampling model $X|\theta \sim f(\cdot|\theta)$ with $\theta \in \mathbb{R}$ and a prior distribution with density π (wrt Lebesgue measure). Consider the loss function for estimating θ

$$L(\theta, \delta) = h(\theta - \delta), \quad h(u) = e^u - u - 1$$

1. Show that for all $u \in \mathbb{R}$ $h(u) \geq 0$ and show that h is a convex function.
2. Determine the form of the Bayesian estimator and prove that it is almost surely unique as soon as

$$\int_{\mathbb{R}} [e^\theta + |\theta|]\pi(\theta|x)d\theta < +\infty$$

3. Assume that $X = (X_1, \cdots, X_n)$ with $X_i \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$ with prior $\pi(\mu, \sigma) \propto 1/\sigma$, $\theta = (\mu, \sigma^2)$

   Compute the Bayesian estimator under the loss function

$$L(\theta, \delta) = e^{\mu-\delta} - \mu + \delta - 1 \quad \delta \in \mathbb{R}$$

   Assume that σ is known and do the same thing.

3. In order to measure the intensity, θ, of a source of radiation in a noisy environment a measurement X_1 is taken without the source present and a second, independent measurement X_2 is taken with it present. It is known that X_1 is $N(\mu, 1)$ and X_2 is $N(\mu + \theta, 1)$, where μ is the mean noise level.
(a) If the prior distribution for μ is $N(\mu_0, 1)$ while the prior for θ is constant.

- (i) Show that the posterior is almost surely defined and write down the joint posterior distribution of μ and θ up to a constant of proportionality.
- (ii) Hence obtain the posterior marginal distribution of θ.
- (iii) The usual estimate of θ is $x_2 - x_1$ explain why $\frac{1}{2}(2x_2 - x_1 - \mu_0)$ might be better.

(b) Compute Jeffrey's prior on (μ, θ) and study the propriety of the posterior distribution.

(c) Derive the Bayesian estimator associated to the quadratic loss under Jeffrey's and compute the frequentist risk, the posterior risk and the integrated risk associated to the given loss, estimator and prior.

4. Let $X|\theta \sim f(\cdot|\theta)$ with $\theta \in \Theta \subset \mathbb{R}^d$. Let π be an improper prior.

1. Show that
$$m(x) = \int_\Theta f(x|\theta)\pi(\theta)d\theta$$
is improper as a measure on $\mathcal{X}$.
2. Assume that $X \in \mathcal{X}$ where $\text{card}(\mathcal{X}) < +\infty$ i.e. $\mathcal{X}$ is finite. Show that there exists $x \in \mathcal{X}$ such that
$$\int_\Theta P[X = x|\theta]\pi(\theta)d\theta = +\infty$$

5 Let $(X_1, \cdots, X_n) \overset{iid}{\sim} f(\cdot|\theta)$, $\theta \in \Theta$ where $f(\cdot|\theta)$ is a canonical exponential family
$$f(x|\theta) = w(x)e^{\theta^T B(x) - D(\theta)}, \quad \Theta \subset \mathbb{R}^d$$

Let π be a prior density on Θ with respect to Lebesgue measure.
(a) Show that the posterior distribution of θ depends only on $T_n = \sum_{i=1}^n B(X_i)$. Show that this result holds true outside exponential family if T_n is any sufficient statistics for θ.

(b) Let $E(a,b)$ be the distribution of the shifted exponential with density

$$\frac{1}{b}e^{-(x-a)/b}, \qquad x > a$$

where $a \in \mathbb{R}, b > 0$ are parameters. Let $X_1, \ldots, X_n$ be a random sample from the distribution $E(a,b)$.

- (i) If a is known, derive Jeffrey's prior on b and compute the posterior mean $\hat{b}$. Show that $\tilde{b} = (n-1)\hat{b}/n$ is MVUE and attains the Cramer Rao lower bound.
- (ii) If a is known and a $Gamma(\alpha, \beta)$ prior is chosen on $1/b$, find the posterior mean. Is it an MVUE? Compute the predictive distribution of X_{n+1} .
- (iii) If b is known and a prior density π_a on a is considered. Show that the posterior distribution on b depends only on $X_{(1)} = \min X_i$, compute the posterior mean associated with the prior $\pi(a) \propto 1$. Show that the posterior is defined and compute it.
- (iv) Show that for all $X_{(1)}$ and all $b > 0$ $\Pi(n|a - X_{(1)}| > z|X_1, \cdots, X_n)$ goes to zero as z goes to infinity uniformly in n. Do we have a Bernstein von Mises property as described in the lectures? Compute frequentist risk of a and $\hat{a}$ (the posterior mean) under the quadratic loss.
- (v) Suppose now that both a and b are unknown. Define $T_1(X) = X_{(1)}$ and $T_2(X) = \sum_{i=1}^n X_i - X_{(1)}$ and show that the posterior distribution depends only on (T_1, T_2). Show that the family of priors

  $$\pi(a|b) \propto e^{\alpha(a-\beta)/b}\mathbf{1_{a<z_1}}, \quad \pi(\mathbf{b}) \propto \mathbf{b^{-m-1}e^{-c/b}1_{b>0}}$$

  is conjuguate.