Problem Sheet 2: Prior choice

1. **Bernoulli Model**

   Consider the model $\mathcal{P} = \{\mathcal{B}(\theta), \theta \in (0,1)\}$. Compute the Fischer information.

2. **Conjugate Distributions**

   Show that the following families of prior distributions are conjugate for $n \geq 1$. In each case, give the expression for the posterior mean.

   1. The family of Gaussian distributions $\mathcal{N}(\mu, \sigma^2)$ for $\mathcal{P} = \{P_\theta^{(n)} = \mathcal{N}(\theta, 1)^{\otimes n}, \theta \in \mathbb{R}\}$.
   2. The family of Gamma distributions $\mathcal{G}(a, b)$ for $\mathcal{P} = \{P_\lambda^{(n)} = \mathcal{E}(\lambda)^{\otimes n}, \lambda > 0\}$.
   3. The family of Beta distributions $\mathcal{B}(a, b)$ for $\mathcal{P} = \{P_p^{(n)} = \mathcal{B}(n, p), p \in [0,1]\}$.

3. **Sequential Posterior and Information**

   We place ourselves in the following Bayesian framework for observations $X_1, \ldots, X_n$ and for θ:

   $$\theta \sim \Pi, \quad d\Pi = \pi d\nu,$$

   $$X_1, \ldots, X_n \mid \theta \sim P_\theta^{\otimes n}, \quad dP_\theta = p_\theta d\mu.$$

   1. Characterize the posterior distributions of $\theta \mid X_1$ and $\theta \mid X_1, X_2$ by giving the corresponding posterior densities.
   2. Show that the posterior distribution of θ given X_1, X_2 in the original model (where θ follows the prior Π) coincides with the posterior distribution of θ given X_2 in a model where one takes $\Pi[\cdot \mid X_1]$ as the prior distribution.
   3. Generalize the previous result to $\theta \mid X_1, \ldots, X_n$. Is there a difference taking the posterior directly with respect to all observations versus sequentially, observation by observation?
   4. Does the order of X_i in the conditioning $\theta \mid X_1, \ldots, X_n$ matter? What would happen if $X_1, \ldots, X_n \mid \theta$ did not follow a product distribution (i.e., not i.i.d.)?
   5. We now ask if conditioning on an observation might have no effect, i.e., if we can have $\mathcal{L}(\theta \mid X) = \mathcal{L}(\theta)$. Let $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$ and

      $$X \mid \theta \sim \mathcal{N}\left(\frac{\theta_1 + \theta_2}{2}, 1\right), \quad \theta \sim \Pi,$$

      with Π of the form $d\Pi(\theta) = 2h(\theta_1 + \theta_2)h(\theta_1 - \theta_2)d\theta_1 d\theta_2$, for h measurable positive and $\int h(u)du = 1$. Furthermore, let:

      $$z_1 = \frac{\theta_1 + \theta_2}{2} \quad \text{and} \quad z_2 = \frac{\theta_1 - \theta_2}{2}.$$

(a) Determine the prior distribution of $z = (z_1, z_2)$. Deduce the marginal distribution of z_2.

(b) Calculate the posterior density $\theta \mid X$ and then that of $z \mid X$.

(c) Deduce the law of $z_2 \mid X$ and conclude. Comment (do we have identifiability?).

4. **Improper prior?**

   Let $X_1, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ and $\pi(\mu, \sigma) = 1/\sigma$ with $\Theta = \mathbb{R} \times \mathbb{R}_*^+$. What is the value of the marginal likelihood/density $p(x_1, \ldots, x_n)$ in this model? Is the measure $\pi(\mu, \sigma)$ usable?

5. **Jeffreys Prior for Exponential**

   Let $X \mid \theta$ follow an exponential distribution $\mathcal{E}(\theta)$. What is the Jeffreys prior for this model?

6. **Jeffreys Prior for Binomial**

   Calculate the Jeffreys prior on θ when $X \mid \theta \sim \mathcal{B}(n, \theta)$.

7. **Conjugation - Multinomial Case**

   Consider $X \mid \theta$ following a multinomial distribution with $X = (X_1, \ldots, X_d)$ and $\theta = (\theta_1, \ldots, \theta_d)$ such that $0 \leq \theta_i \leq 1$ and $\sum_{i=1}^d \theta_i = 1$:

   $$P(X_1 = k_1, \ldots, X_d = k_d \mid \theta) = \frac{n!}{k_1! \ldots k_d!} \theta_1^{k_1} \ldots \theta_d^{k_d}.$$

   Show that the Dirichlet distribution is conjugate for this likelihood.