Problem Sheet 6

The *Inverse-Gamma* distribution $\mathrm{IG}(a, b)$ is the distribution on $\mathbb{R}^+$ with density, for $a, b > 0$,

$$x \mapsto x^{-(a+1)} e^{-b/x} \frac{b^a}{\Gamma(a)}.$$

1. **Conjugate Gaussian family in $\mathbb{R}^d$**

   Let $\mathcal{P} = \{P_\theta = \mathcal{N}(\theta, \Sigma), \theta \in \mathbb{R}^d\}$, where Σ is a known invertible variance-covariance matrix. Let Π be the prior distribution $\mathcal{N}(0, \Lambda)$ on θ, with Λ invertible.

   (a) Show that if $(X_1, \ldots, X_n)|\theta \sim P_\theta^{\otimes n}$, the posterior distribution is written as $\theta|(X_1, \ldots, X_n) \sim \mathcal{N}(\theta_X, \Sigma_X)$, with

   $$\begin{aligned} \Sigma_X &= (n\Sigma^{-1} + \Lambda^{-1})^{-1} \\ \theta_X &= n\Sigma_X \Sigma^{-1} \overline{X}. \end{aligned}$$

   (b) Is the class of prior distributions $\{\mathcal{N}(0, \Lambda), \Lambda \text{ invertible}\}$ conjugate?

2. **Gaussian family with unknown mean and variance**

   Consider the model $\mathcal{P} = \{P_\theta = P_{\mu,\sigma^2} = \mathcal{N}(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma^2 > 0\}$ where we set $\theta = (\mu, \sigma^2)$. We have observations $X = (X_1, \ldots, X_n)$, with distribution $P_\theta^{\otimes n}$ given θ. We define the measure Π:

   $$d\Pi(\mu, \sigma^2) = \frac{1}{\sigma^2} d\mu d\sigma^2,$$

   where $d\mu d\sigma^2$ is interpreted as $d\mathrm{Leb}_{\mathbb{R}}(\mu) d\mathrm{Leb}_{\mathbb{R}^+}(\sigma^2)$.

   (a) Verify that Π is an improper prior "distribution".

   (b) Show that $\mathcal{L}(\sigma^2|X)$ is an $\mathrm{IG}(\frac{n-1}{2}, \frac{s}{2})$ distribution, where $s = \sum_{i=1}^n (X_i - \overline{X})^2$. To do this:

      i. Show that the integral $\int p_{\mu,\sigma^2}(X)\sigma^{-2} d\mu$ is finite. It thus allows us to define, up to a proportionality constant, a "joint density" of (σ^2, X).

      ii. Deduce $\mathcal{L}(\sigma^2|X)$.

      iii. Verify also in passing that the posterior distribution $\mathcal{L}((\mu, \sigma^2)|X)$ is well-defined.

   (c) Characterize the posterior distribution $\mathcal{L}(\mu, \sigma^2|X)$ using $\mathcal{L}(\mu|\sigma^2, X)$ and $\mathcal{L}(\sigma^2|X)$.

   (d) Construct a credible region for μ at level $1 - \alpha$.

3. **Empirical Bayes and normal distributions**

   We are in the framework of the fundamental model $\mathcal{P} = \{P_\theta = \mathcal{N}(\theta, 1), \theta \in \mathbb{R}\}$. We have n i.i.d. observations $X_1, \ldots, X_n$ given θ from the distribution P_θ. Let $\Pi = \Pi_\mu = \mathcal{N}(\mu, 1)$ be a prior distribution on θ. We propose to determine μ using an empirical Bayes method.

   (a) What does this method consist of?

   (b) We construct $\hat{\mu}$ using the marginal maximum likelihood method. Recall the principle of this method in two lines maximum.

   (c) Show that the marginal distribution of $(X_1, \ldots, X_n)$ is that of a Gaussian vector. You may draw inspiration from exercise 5 of PS4.

   (d) Deduce that $\hat{\mu} = \overline{X}$. What is the final posterior distribution suggested by the empirical Bayes method?

4. **Empirical Bayes and Poisson distributions**

   Let $\mathcal{P} = \{P_\theta = \mathcal{P}(\theta), \theta > 0\}$. We have observations $X_1, \ldots, X_n$ i.i.d. with distribution P_θ given θ. Let $\Pi = \Pi_\lambda = \mathcal{E}(\lambda)$ be a prior distribution on θ. We propose to determine λ using an empirical Bayes method.

   (a) Show that the marginal distribution of X_1 is a geometric distribution with parameter $\lambda/(\lambda + 1)$.

   (b) Calculate the marginal density of $(X_1, \ldots, X_n)$ as a function of λ.

   (c) Deduce that $\hat{\lambda}^{EB} = 1/\overline{X}$, then the final posterior distribution suggested by the empirical Bayes method.