Exercise 1. Suppose you have observed \( N \) samples \( k_1, \ldots, k_N \) drawn from a Poisson distribution. Compute the posterior density for the unobserved rate \( \lambda \), i.e.

\[
p(\lambda | k_1, \ldots, k_N) \propto \prod_{n=1}^{N} p(k_n; \lambda) p(\lambda)
\]

using the conjugate Gamma prior \( p(\lambda | \alpha, \beta) \).

2 points

Exercise 2. Consider the linear regression model, i.e.

\[
p(t|X, \theta) = \prod_{n=1}^{N} \mathcal{N}(t_n | w^T x_n, \sigma^2)
\]

with design matrix \( X \) and target outputs \( t \).

Now, compute the Bayesian posterior \( p(w|X, t, \sigma^2) \), assuming that the noise variance \( \sigma^2 \) is known and using a (multi-variate) Gaussian prior

\[
p(w) = \mathcal{N}(w | 0, \sigma_0^2 I)\]

3 points

Exercise 3. Show that the evidence of the Bayesian linear regression model is given by

\[
\ln p(t|X) = \frac{D}{2} \ln \tau_0 + \frac{N}{2} \ln \tau_e - \frac{N}{2} \ln 2\pi - \frac{\tau_e}{2} (X\mu_N - t)^T (X\mu_N - t) - \frac{\tau_0}{2} \mu_N^T \mu_N + \frac{1}{2} \ln |\Sigma_N|
\]

where \( N \) denotes the number of data points, \( D \) the number of inputs and the posterior parameters are given by

\[
\begin{align*}
\mu_N &= \tau_e \Sigma_N X^T t \\
\Sigma_N &= (\tau_0 I + \tau_e X^T X)^{-1}
\end{align*}
\]

2 points
Exercise 4. Fit a linear model to the data in data_simpson.csv that can be downloaded from the course website.

- What if I told you that $x$ denotes annual school teacher’s salary (in thousands of US dollars) and $y$ is the average total SAT score?
- Interpret your fit in the light of this background information.

2 points

Exercise 5. Replicate figure 3.9 of PRML illustrating Bayesian linear regression:

- Data sets in this example were generated by drawing $N = 1, 2, 4$ and 25 points $(x_n, t_n)$ with

$$x_n \sim \text{Uniform}(0, 1)$$

$$t_n = \sin(2\pi x) + 0.3\epsilon$$

where $\epsilon \sim N(0,1)$.

- A linear model with 9 radial basis functions equally spaced between 0 and 1 was then fitted to the data:
  - Assume that the noise precision is known, i.e. $\tau_\epsilon = \frac{1}{0.3^2}$, and use a prior precision of $\tau_0 = 1$.
  - Find a suitable width for the basis functions!

3 points