02477 Bayesian Machine Learning 2025: Assignment 1

This is the first assignment out of three in the Bayesian machine learning course 2025. The assignment is a group work of 3-5 students (make your own groups) and hand in via DTU Learn. The assignment is **mandatory**. The deadline is **2nd of March 23:59**.

Being able to manipulate probability distributions is crucial for probabilistic machine learning. The purpose of this exercise is to become more familiar with probabilistic reasoning and Bayesian computations, e.g.

- basic concepts in Bayesian machine learning (prior, likelihood, posterior)
- the sum rule for marginalization and the product rule constructing joint distributions
- computing the analytical expression for conditional distributions of simple Bayesian models
- manipulating Gaussian distributions

Part 1: The beta-binomial model

Your friend has set up a website for her new business. So far $N = 115$ potential customers has visited her site, but only $y = 4$ customers have completed a purchase. To plan her future investments, she asks you for help to compute the probability that at least one of the next $N^* = 20$ customers will make a purchase. You decide to model the problem using the beta-binomial model with a uniform prior distribution on the probability of making a purchase $\theta \in [0, 1]$:

$$\theta \sim \text{Beta}(a_0, b_0), \tag{1}$$
$$y|\theta \sim \text{Binomial}(N, \theta) \tag{2}$$

where $a_0 = b_0 = 1$.

Task 1.1: Compute the prior mean of θ and provide a 95%-credibility interval for the prior.
Hint: See Section 4.6.6 in Murphy1 for details on posterior/credibility intervals. The book discusses intervals for posterior distribution, but we can also use credibility intervals to summarize prior distributions.

Task 1.2: Compute the posterior mean of θ and provide a 95%-credibility interval for the posterior.

Let y^* denote the number of purchases during the next $N^* = 20$ visits.

Task 1.3: Compute and plot the posterior predictive distribution for y^* given $y = 4$.

Task 1.4: What is the posterior predictive probability that at least one of the next $N^* = 20$ customers will make a purchase?

Task 1.5: Compute mean and variance of the posterior predictive distribution for y^*.

Part 2: Linear Gaussian systems

Let $\boldsymbol{z}_1, \boldsymbol{z}_2 \in \mathbb{R}^2$, and $y \in \mathbb{R}$ be random variables and consider the following linear Gaussian system

$$\boldsymbol{z}_1 \sim \mathcal{N}(\mathbf{0}, v\boldsymbol{I}) \tag{3}$$
$$\boldsymbol{z}_2|\boldsymbol{z}_1 \sim \mathcal{N}(\boldsymbol{z}_1, v\boldsymbol{I}) \tag{4}$$
$$y|\boldsymbol{z}_2 \sim \mathcal{N}(\boldsymbol{a}^T\boldsymbol{z}_2, \sigma^2) \tag{5}$$

where $\boldsymbol{a} \in \mathbb{R}^2$ is constant. The joint distribution of $(\boldsymbol{z}_1, \boldsymbol{z}_2, y)$ is given by

$$p(y, \boldsymbol{z}_1, \boldsymbol{z}_2) = p(y|\boldsymbol{z}_2)p(\boldsymbol{z}_2|\boldsymbol{z}_1)p(\boldsymbol{z}_1). \tag{6}$$

To solve this part, you will need the equations for linear Gaussian systems in section 3.3 in Murphy1 as well as the basic rules of probability theory (sum rule, product rule, conditioning).

Task 2.1: Determine the distribution $p(y)$.

Hints: Compute $p(\boldsymbol{z}_2)$ first. The equations for linear Gaussian systems in Section 3.3 in Murphy1 will be handy.

Task 2.2: Determine the distribution $p(y, \boldsymbol{z}_2|\boldsymbol{z}_1)$.

Task 2.3: Determine the distribution $p(y|\boldsymbol{z}_1)$.

Task 2.4: Determine the distribution $p(\boldsymbol{z}_1|y)$.

Hint: Start by using Bayes' rule.

Part 3: A conjugate model for count data

In week 1, we studied the Beta-binomial model, which is an example of a so-called **conjugate model**. In this exercise, you will work with another conjugate model, namely the Poisson-Gamma model, where the likelihood takes the form of a Poisson distribution, $y_i|\lambda \sim \text{Poisson}(\lambda)$ with rate $\lambda > 0$, and the prior on λ is a Gamma distribution, i.e. $\lambda \sim \text{Gamma}(a_0, b_0)$, where $a_0 > 0$ and b_0 are hyperparameters known as the *shape* and *rate*, respectively.

The Poisson distribution is discrete distribution, which is often applied to model **count data**, where $y_i \in \{0, 1, 2, \dots\}$ are a non-negative integers. The probability mass function (PMF) for the Poisson distribution is:

$$p(y_i|\lambda) = \frac{\lambda^{y_i} e^{-\lambda}}{y_i!} \tag{7}$$

for $y_i \in \{0, 1, 2, \dots\}$.

The Gamma distribution, e.g. $\lambda \sim \text{Gamma}(a_0, b_0)$, is distribution over the non-negative real line, i.e. $\lambda > 0$ with the following probability density function (PDF):

$$p(\lambda|a_0, b_0) = \frac{b_0^{a_0}}{\Gamma(a_0)} \lambda^{a_0 - 1} e^{-b_0 \lambda}, \tag{8}$$

where $\Gamma(x)$ is the gamma function.

Consider now the following model

$$y_i|\lambda \sim \text{Poisson}(\lambda) \tag{9}$$

$$\lambda \sim \text{Gamma}(a_0, b_0), \tag{10}$$

where $\{y_i\}$ are assumed to *conditionally independent* given λ. Suppose we collect N observations such that $\boldsymbol{y} = \{y_i\}_{i=1}^N$

Task 3.1: Determine the joint distribution of $(\boldsymbol{y}, \lambda)$

Task 3.2: Show that the functional form of a Gamma distribution is given by $\log p(\lambda|a, b) = (a-1)\log(\lambda) - b\lambda +$ constant.

Task 3.3: Derive the analytical expression for the posterior distribution $p(\lambda|\boldsymbol{y})$ and show that it is a Gamma-distribution

Suppose we observe $y_1 = 7$, $y_2 = 4$, $y_3 = 8$, $y_4 = 11$, and $y_5 = 12$ such that $N = 5$. Assume $a_0 = 1$ and $b_0 = \frac{1}{10}$.

Task 3.4: Determine the posterior distribution for λ given the data above and report the mean.

Hints: If $\lambda \sim Gamma(a, b)$, then $\mathbb{E}[\lambda] = \frac{a}{b}$.

Task 3.5: Plot $p(\lambda)$ and $p(\lambda|\boldsymbol{y})$ for $\lambda \in [0, 30]$.

Hints: Implement the log density first for numeric stability. The function scipy.stats.gammaln implements the logarithm of the $\Gamma(x)$ function, i.e. $\log \Gamma(x)$.