Practical Session: Bayesian reasoning

1. [Bayesian reasoning] During medical checkup, one of the tests indicates a serious disease. The test has high accuracy 99% (the probability of true positive is 99%, the probability of true negative is 99%). However, the disease is quite rare, and only one person of 10000 is affected. Calculate the probability that the examined person has the disease.

2. [Data modeling: Bayesian vs frequentist] Let $X = \{x_1, \ldots, x_N\}$ be N independent dice rolls. For brevity, we denote the number of times a dice comes up as face $k \in \{1, \ldots, K\}$ as $N_k = \sum_{n=1}^{N} \mathbb{I}(x_n = k)$. With this notation the likelihood has the form

$$p(X \mid \theta) = \prod_{k=1}^{K} \theta_k^{N_k}, \tag{1}$$

where θ_k is the probability of outcome k.

The conjugate prior distribution for multinomial likelihood defined in Eq. 1 is the Dirichlet distribution:

$$\mathrm{Dir}(\theta|\alpha) = \frac{1}{B(\alpha_1, \ldots, \alpha_K)} \prod_{k=1}^{K} \theta_k^{\alpha_k - 1}, \quad \theta \in S_K$$

where $\alpha_k > 0$ and $B(\alpha_1, \ldots, \alpha_K)$ is the normalizing constant, also known as the multivariate Beta function.

(a) Check that the Dirichlet distribution is indeed the conjugate prior for multinomial likelihood.

(b) Train the model, i.e. derive the posterior distribution $p(\theta|X, \alpha)$.

(c) Compute the maximum likelihood estimate for $\theta = (\theta_1, \ldots, \theta_K)$. Do not forget that $\theta \in S_K$, i.e. $\sum_{k=1}^{K} \theta_k = 1$ and $\theta_k \geq 0$ for $k = 1, \ldots, K$.

(d) Compute the expectation of the posterior distribution and compare it with the maximum likelihood estimate. To compute the expectation of Dirichlet distribution, you may use the following formula:

$$\mathbb{E}\theta_k = \frac{\alpha_k}{\sum_{l=1}^{K} \alpha_l}$$

(*) Derive the posterior predictive distribution $p(x_{N+1} = k|X, \alpha) = \int_{S_K} p(x_{N+1} = k|\theta) p(\theta|X, \alpha) \mathrm{d}\theta$. To simplify the answer, you may use the following expression for the multivariate Beta function

$$B(\alpha_1, \ldots, \alpha_K) = \frac{\prod_{k=1}^{K} \Gamma(\alpha_k)}{\Gamma(\sum_{k=1}^{K} \alpha_k)}$$

and the multiplicative property of the Gamma function $\Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$.