Deep|Bayes summer school

Practical Session: Bayesian reasoning

August 20, 2019

1. [Basic Bayesian reasoning] Albus Dumbledore believes that the Dark Lord somehow survived after his death spell bounced off on the historical night of October 31st, 1981. While prophecies about the Dark Lord surviving that night are ambiguous, the Dark Mark on the hand of Severus Snape is a real piece of evidence. One would expect it to fade after the fall of the Dark Lord, but is this a sufficient evidence? Is it that unlikely for a magical mark to stay around after the maker dies?

   Suppose that, if the Dark Lord dies, the Dark Mark continues to exist with a twenty percent probability. On the other hand, if the Dark Lord's sentience lives on, the Dark Mark will stay with one hundred percent chance. Additionally, let's say the prior odds were a hundred-to-one against the Dark Lord surviving. Given that the Severus Snape's Dark Mark has not faded, what is the probability of Dark Lord being alive?

2. [Frequentist framework] Let $X = \{x_1, \dots, x_N\}$ be N independent dice rolls. For brevity, we denote the number of times a dice comes up as face $k \in \{1, \dots, K\}$ 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. Compute the maximum likelihood estimate for $\theta = (\theta_1, \dots, \theta_K)$. Do not forget that $\theta \in S_K$, i.e. $\sum_{k=1}^{K} \theta_k = 1$ and $\theta_k \geq 0$ for $k = 1, \dots, K$.

3. [Bayesian framework] The conjugate prior distribution for multinomial likelihood defined in Eq. 1 is the Dirichlet distribution:

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

   where $\alpha_k > 0$ and $B(\alpha_1, \dots, \alpha_K)$ 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 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

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

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