T-61.5040 Oppivat mallit ja menetelmät
T-61.5040 Learning Models and Methods
Pajunen, Viitaniemi

Exercises 4, 9.2.2007

Problem 1.

Consider an unknown parameter θ which can have only two values, θ_1 and θ_2. You can observe a binary variable $x \in \{0, 1\}$. If $\theta = \theta_1$ then the probabilities of data are $p(x = 0|\theta_1) = 0.8$ and $p(x = 1|\theta_1) = 0.2$. If $\theta = \theta_2$ then the probabilities of data are $p(x = 0|\theta_2) = 0.4$ and $p(x = 1|\theta_2) = 0.6$.

i) Given only the above information, what would you use as a prior distribution $p(\theta)$?

ii) Compute the posterior $p(\theta|x)$

iii) Assume you have independent observations $x_1, \ldots, x_n$. What is the posterior $p(\theta|x_1, \ldots, x_n)$? Can you find a single statistic of the data that determines the posterior?

Problem 2.

Assume that an unknown parameter θ is in the interval $0 \leq \theta \leq 1$. You can observe a variable x with probability $p(x|\theta) = \frac{2x}{\theta^2}$, $0 < x < \theta$.

i) Compute the posterior $p(\theta|x)$ when $p(\theta) = 1$

ii) Compute the posterior when $p(\theta) = 3\theta^2$

iii) Compute the posterior mean $\mathrm{E}(\theta|x)$ for both posteriors

Problem 3.

You have completed a Bayesian inference on θ, and the resulting posterior distribution is $p(\theta|D) = 0.9N(\theta|0, 1) + 0.1N(\theta|4, 0.1^2)$. You are going to use the result to predict the value of y, which is known to be $y = \exp(\theta)$.

i) What is approximately the most probable value of θ, i.e. one that maximizes the posterior? What value y is predicted by it?
Hint: sketch the posterior density and check which of the two obvious choices maximize the posterior.

ii) What is the expected value of y?
Hint: consider the two Normal distributions separately again, and use the mean of a lognormal distribution. *If θ is Normal, $N(\theta|\mu, \sigma^2)$, then $y = \exp(\theta)$ is lognormal. The mean of y is $\exp(\mu + \frac{1}{2}\sigma^2)$.*

Problem 4.

Assume $\theta \in \{\theta_1, \theta_2, \theta_3\}$ and you have obtained a posterior $p(\theta_1|D) = 1/2$, $p(\theta_2|D) = 1/4$, $p(\theta_3|D) = 1/4$. The value $\tilde{y} \in \{1, 2, 3\}$ can be predicted using θ.

i) If $p(\tilde{y} = 1|\theta = \theta_1) = 1$ and if $p(\tilde{y} = 3|\theta = \theta_2) = 1$ and if $p(\tilde{y} = 3|\theta = \theta_3) = 1$, what is the predictive distribution $p(\tilde{y}|D)$?

ii) Let's define the probabilities $p(\tilde{y}|\theta)$ by the following table:

	θ_1	θ_2	θ_3
1	0.35	0.5	0.1
2	0.3	0.4	0.4
3	0.35	0.1	0.5

What is the predictive distribution $p(\tilde{y}|D)$? Compare the $\tilde{y}$ that maximizes the predictive distribution to the values maximizing $p(\tilde{y}|\theta)$ for each θ.