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

Exercises 3, 2.2.2007

Problem 1.

Assume that you live in New York, a city with approximately 10 million inhabitants. Also assume that the fingerprint recognition system of the N.Y.P.D. is such that about 1 million different sets of fingerprints can be distinguished. Assume further that the system partitions the fingerprints into equivalence classes that are roughly of the same size.

Now a burglary is committed. Assume that exactly one citizen of New York is guilty of the burglary and leaves his/her fingerprints at the place of crime. By a chance the police notices that your fingerprints match the ones found at the place of crime. You are accused of the burglary based on the matching of the fingerprints only. What is the probability that you are innocent? Assume no other information to be known.
(S. Eddy, D. MacKay, Nature: Is the Pope the Pope?, June 27, 1996.)

Problem 2.

Assume that you are presented with three closed doors. Behind one of them there is a prize waiting for you. You can choose one of the doors blindly, but not open it.
After this, another of the remaining doors is opened. (The door to be opened is chosen so that there is no prize behind that door.)
You are now free to change your choice of the door if you wish. Should you change or not?

Problem 3.

(In this exercise, we will study some basic properties of probability and show that the lack of any of these basic properties results in the "Dutch Book".) Consider a finite set of events $A_1, \ldots, A_n$. Assume that you have invented a new way of representing uncertainty. You meet a Bayesian friend and decide to trade bets with him. For simplicity, you decide to trade tickets T_i to bet on event A_i so that the seller of the ticket pays the buyer 1 EUR if the event A_i happens, nothing otherwise. The trading procedure is such that first you trade the tickets, after which you observe outcomes of all the events A_i practically simultaneously.
We know that the Bayesian will buy T_i from you for less than $p(A_i)$, and sell T_i to you for more than $p(A_i)$. You can also bundle together several individual transactions. The Bayesian will accept the trade if 1) the net price he pays is less than the sum of the winning probabilities of the tickets he receives minus the sum of winning probabilities of the tickets he turns over to you, or 2) the net price he receives is more than the sum of the winning probabilities of the tickets he turns over minus the sum of winning probabilities of the tickets he receives.

Whatever your ingenious method, you must have a way of deciding what is the limit price at which you buy or sell. Denote this function by $q(A_i)$. A *Dutch Book* means a set of bets which guarantee that you lose money, whichever the outcome.

i) Show that violating the axiom that equivalent events have equal probabilities leads to a Dutch Book.

ii) Show that if the event A_1 will happen with probability 1, then $q(A_1) \neq 1$ leads to a Dutch Book.

iii) Show that violating the sum rule leads to a Dutch Book. (The sum rule says that for two mutually exclusive but exhaustive events A_1 and A_2, $p(A_1) + p(A_2) = 1$.)

Problem 4.

Bayesian Inference involves conditional probabilities and their manipulation. Two useful *iteration formulas* are often helpful:

$$\mathrm{E}(x) = \mathrm{E}(\mathrm{E}(x|y)) \quad (1)$$

$$\mathrm{Var}(x) = \mathrm{E}(\mathrm{Var}(x|y)) + \mathrm{Var}(\mathrm{E}(x|y)) \quad (2)$$

The "outer" integrals are taken over the distribution $p(y)$ and the "inner" integrals are taken over the distribution $p(x|y)$.

i) Bayesian Inference proceeds from a prior distribution $p(\theta)$ to a posterior distribution $p(\theta|D)$. Use Formula (2) to interpret how the variance of the distribution of θ changes on average when learning from data D.

ii) Consider a random variable $s = \sum_{i=1}^{N} s_i$, where s_i are iid variables with $\mathrm{Exp}(\mu)$-distribution, and N has a $\mathrm{Poisson}(\lambda)$-distribution. Calculate the mean and the variance of s.
Recall that if a random variable $X \sim \mathrm{Poisson}(\lambda)$, we have $E(X) = \lambda$ and $\mathrm{Var}(X) = \lambda$.
If a random variable $Y \sim \mathrm{Exp}(\mu)$, we have $\mathrm{E}(Y) = 1/\mu$ and $\mathrm{Var}(Y) = 1/\mu^2$.