Questions 2

Dr. Roman Belavkin

BIS4435

Question 1

Name and briefly describe three main sources of uncertainty.

Question 2

What is probability? What is the probability of 'raining or not raining' tomorrow?

Question 3

Suppose that a university database has a variable called 'Age' storing the age of a student, and it can have 100 values. What is your estimate of the prior probability of each value? What if the database contains records of a 20,000 of students, and all of them are older than 16?

Question 4

The probability of disjunction (logical OR) of several disjoint events is simply the sum of their probabilities. For example, for two independent events E_1 and E_2

$$P(E_1\, OR\, E_2) = P(E_1) + P(E_2)$$

If, however, the events are not disjoint (i.e. the events can happen together), then

$$P(E_1\, OR\, E_2) = P(E_1) + P(E_2) - P(E_1, E_2)$$

where $P(E_1, E_2)$ is the joint probability of E_1 and E_2 (logical AND).

a) What is the probability of disjunction of two events that are not disjoint but are independent of each other?

b) Consider two fair and independent coins tossed together. What is the probability that at least one of the coins will be 'Heads' (i.e. P(heads OR heads))?

Question 5

Suppose you know $P(A, B)$ - the joint probability distribution of events A and B. Let also $P(A)$ and $P(B)$ be the probabilities of each event individually (i.e. the marginal probabilities, which can be computed from the joint distribution).

a) What these probabilities can tell you about the relation between events A and B?

b) How could you use this information to reduce the uncertainty about one event based on information about another?

c) Why is it better to use the information about event A to assess the probability of B, then simply using probability $P(B)$?

Question 6

Consider two systems: A bicycle and an airplane. Why is the uncertainty associated with an airplane higher than the uncertainty of a bicycle?

Question 7

Suppose that a database has recorded a very unusual case - its values are very different from other, more typical cases. Why is this case more interesting from the information theory point of view?

Question 8

Let variable x can have values 1, 2 and 3 with probabilities $P(1) = 1/5$, $P(2) = 3/5$ and $P(3) = 1/5$. What is the expected value of x? Compare it with mean value of $(1, 2, 2, 2, 3)$?

Question 9

Consider the following sets of values for variables x and y:

x	y
-1	-2
0	0
1	2

Compute the expected values and variances of x and y. You can compute them as the means and the mean square deviations. Compare the results. Which variable is more uncertain (risky)?

Question 10

Compute the covariance and correlation between x and y from Question 9:

x	y
-1	-2
0	0
1	2

Are these variables correlated, uncorrelated or anticorrelated?

Question 11

Suppose the database contains data for m independent variables. What should the covariance and the correlation matrices look like?

Question 12

Suppose of you have a choice between two lotteries A and B:

- Lottery A: The utility can have values -1, 0 or 1.
- Lottery B: The utility can have values -2, 0 or 2.

Suppose that all values have equal probabilities.

a) What choice does the maximum expected utility principle suggest?

b) Have you lost some information by using only the expected values?