CS475 Spring 2006
Homework 1

Due Date: Feb 15 2006

Problem 1

In class, we used a small example with three random variables (Toothache, Catch and Cavity), each taking on two possible values. The joint probabilities that each random variable takes on the respective values is given bellow (they are different than values from the slides):

	toothache		¬toothache	
	catch	¬catch	catch	¬catch
cavity	0.048	0.032	0.054	0.036
¬cavity	0.002	0.018	0.081	0.729

As in the class, we use upper-case letters for *random variables* (e.g. "Cavity"), and lower-case letters for *events* that a random variable takes on a particular value (e.g. "¬cavity" is an event where the random variable Cavity takes on the value of ¬cavity).

Use the values in the table to answer the following questions. Please explain your answers and include all necessary calculations.

1. Is the event "catch" (unconditionally) independent of "toothache"? Is it conditionally independent of "toothache" given "cavity"?

2. Is the random variable "Catch" (unconditionally) independent of the random variable "Toothache"? Recall that to have independence of random variables, all events corresponding to all possible outcomes of both random variables must be independent.

3. Is the random variable "Catch" *conditionally* independent of "Toothache", given "Cavity"?

Problem 2

Prove the Law of Total Probability:

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$$

where $A \subseteq \Omega$ is an arbitrary event (e.g. set of possible outcomes of an experiment), and $\{B_i\}_{i=1}^n$ is a partition of the sample space Ω into measurable sets, i.e.:

$\forall i, j \in \{1, \ldots, n\}, i \neq j : B_i \cap B_j = \emptyset$	(B_is are disjoint)
$\bigcup_{i=1}^n B_i = \Omega$	(they cover the whole sample space)
$\forall i \in \{1, \ldots, n\} : P(B_i) > 0$	(and they are measurable)

The sample space Ω is a set of all possible outcomes in the world, so $P(\Omega) = 1$. The measurability of B_is is a technical requirement, to allow to condition on B_i in expression $P(A|B_i)$. The Law of Total probability allows one to decompose a complex event (A) into smaller parts of which the probabilities are known (see Problem 3).

Hint: Prove $P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$ first, and then generalize the proof (B^c is a complement of B, i.e. $\Omega \backslash B$). Use the fact that $P(B_1) + P(B_2) = P(B_1 \cup B_2)$ if $B_1 \cap B_2 = \emptyset$.

Problem 3

Assume there are two different tests for presence of HIV virus in blood. You are given information about what is the probability of the tests to show positive ($T+$) or negative ($T-$) given that the person actually is HIV positive ($HIV+$) or not ($HIV-$). The numbers for both tests are given bellow:

Test A: $P(T+|HIV+) = 0.999$ and $P(T-|HIV-) = 0.99$

Test B: $P(T+|HIV+) = 0.99$ and $P(T-|HIV-) = 0.999$

1. Compute $P(T-|HIV+)$ and $P(T+|HIV-)$ for each test. Note: we assume the tests either show $T+$ or $T-$, and a person is either $HIV+$ or $HIV-$, so there are no other possible outcomes.

2. Assume you know the value of $P(HIV+)$ (probability that a randomly selected person from a population is HIV positive). How would you now calculate $P(T+)$? (*Hint:* use the Law of Total Probability)

3. Again, assuming that $P(HIV+)$ is known, use the Bayes Rule to express $P(HIV+|T+)$ (which is what is actually interesting: probability that one is HIV positive if the test shows positive)

4. Find an estimate of $P(HIV+)$ for the world, USA or Africa, and compute the actual $P(HIV+|T+)$ for both Test A and Test B. Which one is more accurate (has fewer false positives)? Does the ranking depend on the value of $P(HIV+)$? Explain your answer. Note: you will *not* be graded on accuracy of the $P(HIV+)$ you use, but you may find the results interesting if you use realistic values. If you cannot find anything, assume that $P(HIV+) = 0.006$.

Problem 4

Show that the following three statements are equivalent:

$$\begin{aligned} i) &\quad P(X|Y,Z) = P(X|Z) \\ ii) &\quad P(Y|X,Z) = P(Y|Z) \\ iii) &\quad P(X,Y|Z) = P(X|Z)P(Y|Z) \end{aligned}$$

where X,Y and Z are random variables.

Equivalence of the first two statements show that conditional independence is symmetric (X and Y are conditionally independent given Z, and the order of X and Y doesn't matter). The third statement is analogous to the definition of unconditional independence: $P(X,Y) = P(X)P(Y)$.