UNIVERSITY OF WARSAW

Faculty of Mathematics, Informatics and Mechanics
University of Warsaw

Statistical Data Analysis 2

Exercises # 1

Krzysztof Gogolewski,
based on the notes of Ewa Szczurek

Problem 1: Conditional independence

Let X, Y, Z be random variables. X and Y are said to be *conditionally independent* given Z (in symbols $X \perp Y \mid Z$) if

$$P(X, Y \mid Z) = P(X \mid Z)P(Y \mid Z).$$

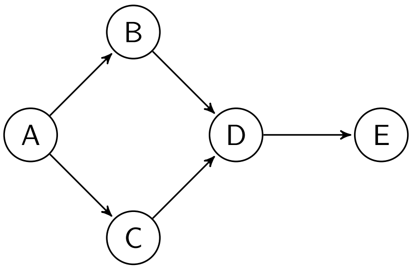
This condition is equivalent to

$$P(X \mid Y, Z) = P(X \mid Z).$$

Using the laws of probability, show that this equivalence holds (show both directions of the proof).

Problem 2: Markov blanket

Consider the following graphical structure of a Bayesian network:



Determine the Markov blanket $\mathrm{MB}(C)$ of the node C and show that the conditional probability $P(C \mid A, B, D, E)$ can be expressed as

$$P(C \mid A, B, D, E) = P(C \mid \mathrm{MB}(C)).$$

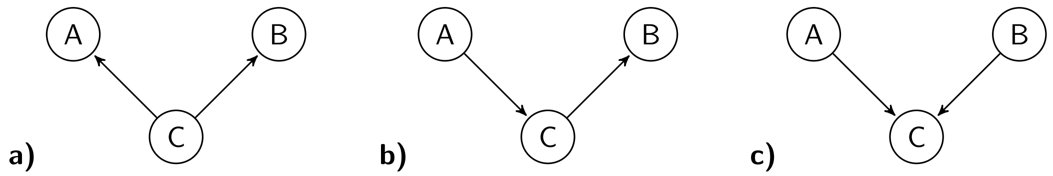
Problem 3: Coin tossing

a) Little Joe had too much time during his home-schooling and he conducted the following experiment. He threw a coin $n = 200$ times and counted the number of times when the heads appeared. After all the tries there were $k = 134$ heads. Joe wants to confirm if his coin is fair, because ha has a feeling that his brother Tom might have swap it. How can we help Joe?

b) Simulate a vector describing the outcome of n coin tosses for which the probability of heads is equal to p. Consider two approaches. First, find the maximum likelihood estimator $\hat{p}$ for the probability of heads. Next, find the maximum a posteriori estimate of p_{MAP} given the prior $\mathcal{B}$ distribution.

Problem 4: Conditional independence and BNs

Consider the following graphical structures, corresponding to (different) Bayesian networks. For which network does the statement $A \perp B \mid C$ hold? For which does the statement $A \perp B$ hold? Prove your answers by the laws of probability.


A
B
C
a)
A
B
C
b)
A
B
C
c)