STOCKHOLMS UNIVERSITET
MATEMATISKA INSTITUTIONEN
Joanna Tyrcha

February 9, 2021

**Take Home Exam, Part 1:**
**Statistical Information Theory (MT7037)**

***Please read the following instructions carefully before starting:***

- Deadline February 23, 2021, at 8 pm.
- Credit points will be given based on clear logical explanation and steps leading to the final solution, so do not simply state the final answer.
- Credit for the course is earned by writing reports for **two take-home exams**. Obtainable grades together for two take-home exams:
  $A(\geq 90\ points), B(\geq 80 points), C(\geq 70\ points), D(\geq 60\ points), \geq (50\ points), Fx\ and\ F.$
  Note that we reserve the right to rescale the points for grades by a factor *c* (0.9< *c* < 1) stated depending on the outcome of the exam.
- The report must be completed independently! Plagiarism or other forms of cheating is a serious act - to underline this your report must as cover page contain the signed "confirmation sheet" that your work is made in accordance with the rules for written exams at Stockholm University.

**<u>Exercise 1</u>** (16 points)

a) (6 points) Let $X$ and $Y$ be discrete random variables. Show that $H\big((X+Y)|X\big) = H(Y|X)$. Argue that if $X$ and $Y$ are independent then $H(X+Y) \geq H(Y)$.

b) (10 points) Show that among all ***N***-valued random variables (i.e. with values $k = 1, 2, \ldots$) with expected value $\mu$, the geometric distributed random variable with expected value $\mu$ has the maximum value of Shannon entropy.

*Reminder:* The probability function in a geometric distribution is $p(k) = p(1-p)^{k-1}, k = 1, 2, \ldots$

**<u>Exercise 2</u>** (9 points)

a) (3 points) Show that the difference entropy is translation invariant i.e. that for a constant *c*, $h(X+c) = h(X)$, where $X$ is a continuous random variable.

b) (6 points) Suppose that $(X, Y, Z)$ are jointly normal distributed and that $X \to Y \to Z$ forms a Markov chain. Let $X$ and $Y$ have the correlation coefficient equal to $\rho_{XY}$ , and let $Y$ and $Z$ have the correlation coefficient equal to $\rho_{YZ}$.
Show that mutual information $I(X;Z) = -\frac{1}{2}\log\left(1-\rho_{XY}^2\rho_{YZ}^2\right)$.

**<u>Exercise 3</u>** (25 points)

a) (5 points) Define the transfer entropy, starting from Schreiber's definition in *Phys. Rev. Lett., Vol. 85, Number 2, 461-464, 2000*, in terms of mutual information and Shannon entropy.

b) (20 points) To complete this part, you need to read the paper by Seghouane and Amari in *Neural Computation 24, 1722-1739, 2012 "Identification of Directed Influence: Granger Causality, Kullback-Leibler Divergence, and Complexity"* and answer the following questions:

   (i) (5 points) Look at the formula (2.1) in the paper. Is it equivalent to any of the concepts you have learned during the course? If yes, then to which ones?

(ii) (5 points) What is the difference between the directed influence measure defined in the formula (2.2) and transfer entropy definition by Schreiber?

(iii) (5 points) Comment on the formula (4.5) in terms of what you have learned during the course.

(iv) (5 points) Please provide your opinion about the directed influence measures presented in the paper (the potential advantages and/or disadvantages).

**Good Luck!**