LUND UNIVERSITY
Electrical and Information Technology

Home Exam in Information Theory, EITN45

May 30 – June 5, 2015

Name: ____________________

Id Number: ____________________

Programme: __________

Nbr of sheets: __________

Mark with a cross the problems you solved.				
1	2	3	4	5

Signature: ____________________

Control protocol						
1	2	3	4	5	$\sum$	Grade

LUND UNIVERSITY
Electrical and Information Technology

Home Exam in Information Theory, EITN45

May 30 – June 5, 2015

- Problems shall be solved and handed in individually.
- Write your name on each paper.
- Start a new solution on a new sheet of paper. Use only one side of the paper.
- Solutions should clearly show the line of reasoning.
- Program code, e.g. MATLAB, python or C/C++, used for derivations should be handed in as a complement to the solution. However, the solutions should be fully understandable without the code. Normal command line calculations, e.g. entropy, is not required to hand in.

Good luck!

Problem 1

Compress the following text string using LZ77 with $S = 16$ and $B = 8$,

I scream, you scream, we all scream for icecream.

Initialise the encoder with the beginning of the string. Give the codewords and the compression ratio.

(10p)

Problem 2

Consider a 2-dimensional uniform distribution over the shaded area shown in Figure 2.1. Derive

(a) $H(X, Y)$
(b) $H(X)$ and $H(Y)$
(c) $I(X; Y)$
(d) $H(X|Y)$ and $H(Y|X)$

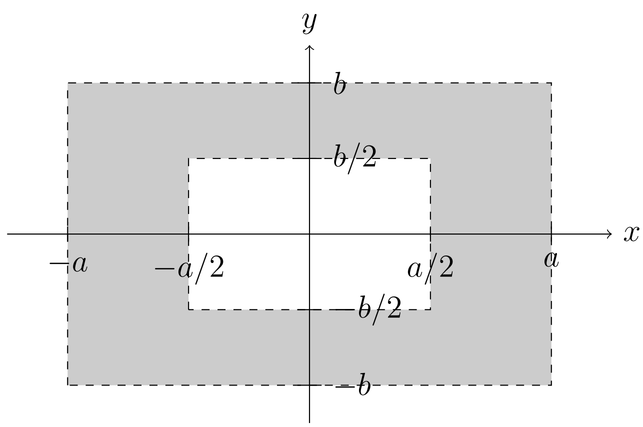

Figure 2.1: A 2-dimension uniform distribution.

(2+3+2+3=10p)

Problem 3

Consider a random variable with k outcomes, $X \in \{x_i\}_{i=1}^k$, and probabilities $p_i = P(X = x_i)$. To construct a source code for the variable, order the outcomes according to the probabilities, with the highest to the left, i.e. the vector becomes $x_1, x_2, \dots, x_k$ where $p_1 \geqslant p_2 \geqslant \cdots \geqslant p_k$. Split the vector in two parts such that their sums are as close as possible. That is, find an index q such that

$$\left|\sum_{i=1}^{q} p_i - \sum_{j=q+1}^{k} p_j\right|$$

is as small as possible. Label the left part with 0 and the right part with 1. Continue the procedure iteratively in the same way with each of the parts until there are only two outcomes in the vector. The labeling constitute the codeword for each outcome.

The described method is often referred to as Fano coding. For each of the probability vectors below find the Fano code and state if it is also a Huffman code and/or if it is optimal.

$$P_{\text{a}} = \left(\frac{4}{10}, \frac{3}{10}, \frac{2}{10}, \frac{1}{10}\right)$$

$$P_{\text{b}} = \left(\frac{6}{21}, \frac{5}{21}, \frac{4}{21}, \frac{3}{21}, \frac{2}{21}, \frac{1}{21}\right)$$

$$P_{\text{c}} = \left(\frac{15}{43}, \frac{7}{43}, \frac{7}{43}, \frac{7}{43}, \frac{7}{43}\right)$$

(4+3+3=10p)

Problem 4

In Figure 4.1(a) the number of day (i.e. light) hours in Lund over the months in a year is shown. In Figure 4.1(b) the number of sun hours in Lund over a year is shown. Define two random variables,

$$T = \text{Time of day, either Day or Night}$$
$$W = \text{Weather, either Sunny or Not sunny}$$

Define another variable, M, which tells which month of the year it is.

(a) If you look out the window one day, how much information do you gain about what month it is by observing if it is day or night, i.e. T?

(b) If you know that it is day, i.e. $T = \text{Day}$, how much information do you get about the weather by observing the month?

Assume it is not a leap year.

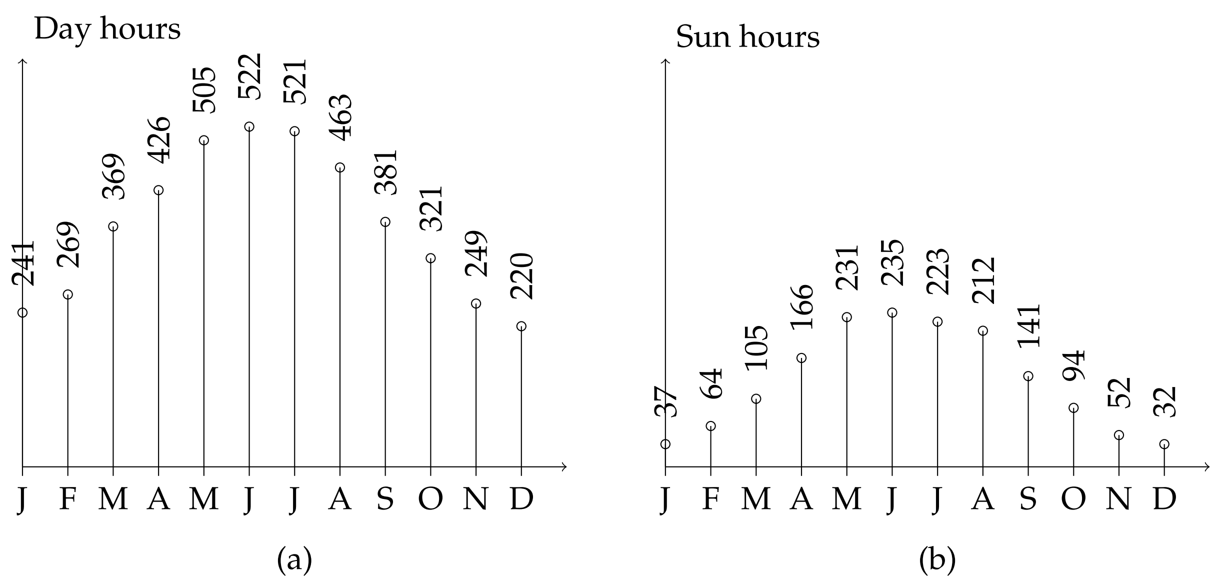

Figure 4.1: Number of day hours (a) and sun hours (b) per month in Lund over the year.

(5+5=10p)

Problem 5

A communication scheme uses 4-PAM system, meaning there are four different signal alternatives differentiated by their amplitudes, see Figure 5.1. During the transmission there is a noise, Z, added to the signal so the received signal is $Y = X + Z$. The noise has the distribution as viewed in Figure 5.2.

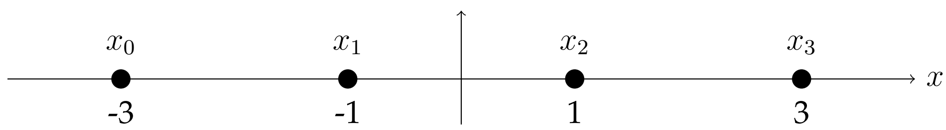

Figure 5.1: 4-PAM signalling.

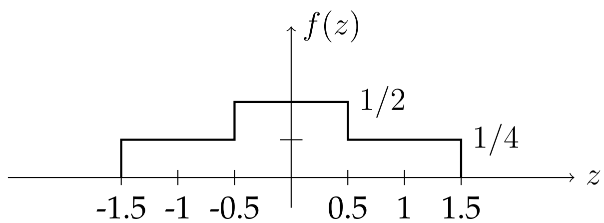

Figure 5.2: Density function of the additive noise.

(a) Assuming the signal alternatives are equally likely, how much information can be transmitted in each transmission (channel use)?

(b) What is the capacity for the transmission, i.e. what is the maximum mutual information using the given signal alternativess and the noise. For what distribution on X can it be obtained? How is the average power of the transmitted signal affected by the optimisation?

(5+5=10p)