Information Theory & Statistical Physics

Lecture: Prof. Dr. Johannes Berg
Exercises: Stephan Kleinbölting **Sheet 1**

Due date: 04.05.17 12:00 *To be discussed on: 10.05.17*

5 Entropy I 30 pts.

(a) (10 pts) Show that the entropy $H(p_1, \dots, p_n)$ assumes its maximum value for a uniform distribution.
Note: You need to account for the constraint of total probability $\sum p_n = 1$ via a Lagrange multiplier.

(b) (10 pts) Now consider a random variable A and maximize H under the additional constraint $\langle A \rangle = a$. What is the maximizing distribution? Does it look familiar?

In the lecture the entropy was introduced axiomatically. One of the axioms ensures that it can be decomposed into entropies of smaller ensembles.

$$H(p_1, p_2, ..., p_n) = H(p_1, 1 - p_1) + (1 - p_1)H(\frac{p_2}{1 - p_1}, \frac{p_3}{1 - p_1}, ..., \frac{p_n}{1 - p_1}). \tag{1}$$

(c) (10 pts) A biased coin is tossed until a “tail” is thrown. The probability for the coin to come as “head” in a single toss is f. Calculate the entropy of the (random) number of necessary tosses in two ways: using the definition of entropy, and using the decomposition of entropy (1).

6 Sum of random variables 20 pts.

Let X, Y be two discrete, independent random variables that take values in $\mathcal{A}$.

(a) (2 pts) Show that independence implies $p_X(x|y) = p_X(x)$.

Define a new random variable as their sum $Z = X + Y$.

(b) (10 pts) Show that Z is distributed according to the *convolution* of p_X and p_Y

$$p_Z(z) = \sum_{x \in \mathcal{A}} p_X(x) p_Y(z - x) \tag{2}$$

Hint: Start by considering $p_Z(z) := P(Z = z) = P(X + Y = z)$ and rewriting this as an expectation value over a conditional distribution.

(c) (8 pts) Consider two cubic dice that you may label from the set $\{0, \dots, 6\}$ as you like. Is it possible to label them in such a way, that the sum of a roll of both dice has a uniform distribution over $\{1, \dots 12\}$? If yes, find such a configuration.

7 Weak law of large numbers 10 pts.

Let X_n be a sequence of iid. random variables with mean μ and variance σ^2. Let $\bar{X} = \frac{1}{N} \sum_{i=1}^{N} X_i$ denote their arithmatic mean.

Show that in the limit of large N, the arithmatic mean converges (in mean square) to μ

$$\lim_{N \to \infty} \text{Var}(\bar{X} - \mu) = 0 \tag{3}$$