Diskrete Stochastik und Informationstheorie – 7 May 2014

Exercise 19. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space with $A, B \in \mathcal{A}$, $A \cap B = \emptyset$ and $\mathbb{P}(A) = \mathbb{P}(B) = \frac{1}{4}$. Let $X, Y : \Omega \to \{0, 1, -1\}$ be random variables defined by

$$X(\omega) = \begin{cases} 1 & \omega \in A \\ -1 & \omega \in B \\ 0 & \text{else} \end{cases} \qquad Y(\omega) = \begin{cases} -1 & \omega \in A \\ 1 & \omega \in B \\ 0 & \text{else} \end{cases}$$

(a) Argue that X and Y are not independent and that

$$H(X) = H(Y) = H(X, Y) = I(X; Y) = \frac{3}{2} \qquad \text{and} \qquad H(X|Y) = H(Y|X) = 0.$$

Whenever possible, avoid numeric calculations such as Shannon's formula and use general, more intuitive properties of entropy and mutual information instead.

(b) Let $Z = X \cdot Y$. Again avoiding calculations as far as possible, argue that

$$H(Z) < H(X, Y) = H(X, Y, Z) \qquad \text{and} \qquad H(Z|X) = 0 \text{ but } H(X|Z) > 0.$$

(c) Calculate $H(X|Z)$ and show that $H(X|Z) = I(X; Y \mid Z)$.

Exercise 20. Let X be a random variable with values in $\{x_1, \ldots, x_n\} \subseteq \mathbb{R}$ (each with non-zero probability), and let $Y = f(X)$ be another random variable for some function $f : \{x_1, \ldots, x_n\} \to \mathbb{R}$.

(a) Show that $H(Y) < H(X)$ if f is not injective (i.e., if there are elements $x_i \neq x_j$ with $f(x_i) = f(x_j)$).

(b) Which of the (in)-equalities $H(X) \leq H(Y)$, $H(X) = H(Y)$, $H(X) \geq H(Y)$ hold in general for $f(x) = e^x$, which for $f(x) = |x|$? What can you say about $I(X; Y)$?

Exercise 21. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space with random variables X, Y, Z defined on it, and let f be a function.

(a) Show that $I(X; X) = H(X)$ and that $I(f(X); X) = H(f(X)) \leq H(X)$.

(b) Give examples of X, Y and Z such that $I(X; Y \mid Z) < I(X; Y)$ and examples where $I(X; Y \mid Z) > I(X; Y)$.

Exercise 22. (Jensen's inequality) Let $f : I \to \mathbb{R}$ be a convex function defined on an intervall $I \subseteq \mathbb{R}$. Let $x_1, \ldots, x_n \in I$ be points in the interval, and let $q_i \in [0, 1]$, $i = 1, \ldots, n$ be coefficients such that $\sum_{i=1}^{n} q_i = 1$. Show that

$$f\left(\sum_{i=1}^{n} q_i x_i\right) \leq \sum_{i=1}^{n} q_i \cdot f(x_i).$$

Hint: Use induction on n.