# Information Theory SV3

**Question 1:** (a)   Consider two discrete probability distributions $p(x)$ and $q(x)$ over the same set of four values $\{x\}$ of a random variable:

| $x$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $p(x)$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{4}$ | $\frac{1}{2}$ |
| $q(x)$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ |

    (i)   Calculate the cross-entropy $H(p,q)$ (unfortunate notation!) between $p(x)$ and $q(x)$.

    (ii)   Calculate their Kullback–Leibler divergence $D_{KL}(p\|q)$.

    (iii)   Comment on the use of the metrics $H(p,q)$ and $D_{KL}(p\|q)$ in machine learning and for calculating the efficiency of codes.

(b)   Explain the concept of entropy in:

    (i)   Classical thermodynamics ("entropy", developed by Clausius and others).

    (ii)   Statistical mechanics ("statistical entropy", developed by Boltzmann and others – presented in Lecture 12).

    (iii)   Information theory ("information entropy", developed by Shannon).

(c)   A continuous communication channel adds Gaussian white noise to signals transmitted through it. The ratio of signal power to noise power is 30 decibels, and the frequency bandwidth of this channel is 10 MHz. Roughly what is the information capacity $C$ of this channel, in bits/second?

(d)   Given two (binary) noisy channels with capacities $C_1$ and $C_2$, determine the capacity of the concatenation of the two channels (i.e. the output of the first channel is the input of the second).

(e)   Explain what is meant by the Kullback–Leibler divergence, and why it is a divergence and not a distance.

(f)   $Y$ and $Z$ are two continuous random variables. $Y$ has an exponential probability density distribution $p(x)$ over $x \in [0,\infty)$: $p(x) = e^{-x}$. $Z$ has a uniform probability density distribution: $p(x) = 1/\alpha$ for $x \in [0,\alpha]$, else $p(x) = 0$.

Calculate the differential entropies $h(Y)$ and $h(Z)$ for these two continuous random variables, and find the value of $\alpha$ for which these differential entropies are the same. Sketch these distributions.