Exercise session 1

1. Let Ω be a finite set. For probability distributions $p, q \in \mathcal{P}(\Omega)$ define the *Kullback-Leibler* divergence as

$$D(q||p) = \sum_{\omega \in \Omega} q(\omega) \ln \frac{q(\omega)}{p(\omega)},$$

with the convention that $0 \ln \frac{0}{0} = 0$, $0 \ln 0 = 0$ and $\ln 0 = \infty$. Prove that

(a) the Kullback-Leibler divergence is convex (by computing the second derivatives).

(b) $D(p||q) \geq 0$ for all p, q and $D(p||q) > 0$ unless $p = q$.

(c) if $q = q_1 \otimes q_2$ and $p = p_1 \otimes p_2$ are product distributions, then

$$D(q||p) = D(q_1||p_1) + D(q_2||p_2).$$

2. The *entropy* of a probability distribution p on a finite set Ω is defined as

$$H(p) = -\sum_{\omega \in \Omega} p(\omega) \ln p(\omega).$$

Prove that

(a) $H(p) \geq 0$ and $H(p) = 0$ iff there is $\omega \in \Omega$ such that $p(\omega) = 1$.

(b) the uniform distribution is the unique global maximiser of the entropy.

(c) if q is a distribution on $\Omega \times \Omega$ and q_1, q_2 are its marginals, then $H(q) \leq H(q_1) + H(q_2)$.

3. The *entropy* of a random variable X with values in a finite set Ω is defined as $H(X) = H(p_X)$, where p_X is the probability distribution defined by

$$p_X(\omega) = P[X = \omega].$$

If Y is a second random variable, then $H(X, Y)$ is the entropy of the Ω^2-valued random variable (X, Y). Furthermore, the *conditional entropy* is defined as

$$H(X|Y) = \sum_{x,y \in \Omega} P[X = x, Y = y] \ln \frac{P[X = x, Y = y]}{P[Y = y]}.$$

Finally, the *mutual information* of X,Y is defined as

$$I(X, Y) = \sum_{x,y \in \Omega} P[X = x, Y = y] \ln \frac{P[X = x, Y = y]}{P[X = x] P[Y = y]}.$$

Show that

(a) $H(X, Y) = H(X) + H(Y|X)$.

(b) $I(X, Y) = H(X) - H(X|Y)$.

(c) $I(X,Y) \geq 0$ and $I(X,Y) = 0$ iff X,Y are independent.

4. Suppose that $X_1,\ldots,X_n$ are independent random variables with value sin a finite set Ω that are identically distributed according to $p \in \mathcal{P}(\Omega)$. Let

$$\lambda(\omega) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{X_i = \omega\}$$

and suppose that $q \in \mathcal{P}(\Omega)$ is such that $nq(\omega)$ is an integer for every $\omega \in \Omega$. Show that

$$P[\lambda = q] = \exp(-nD(q||p) + o(n)).$$

Exercise session 2

5. Assuming that $\mu(\sigma) = \exp(-E(\sigma)/T)/Z$, show that

$$U = T^2 \frac{\partial \ln Z}{\partial T}.$$

6. In the Ising model, the variable nodes V are points on the d-dimensional integer grid $\{1, \ldots, n\}^d$. There is a contraint node a_x connected to each single variable node x. In addition, there is a constraint node $a_{x,y}$ between two variable nodes x, y iff $\|x - y\|_1 = 1$. The configuration space is $\{\pm 1\}^V$ and the weight functions are

$$\psi_{a_{x,y}}(\sigma_x, \sigma_y) = \exp(\beta \sigma_x \sigma_y), \qquad \psi_{a_x}(\sigma_x) = \exp(\beta B \sigma_x),$$

where $B \in \mathbb{R}$ is a parameter called the *external field* and $\beta > 0$ is called the *inverse temperature*. Write the Belief Propagation equations for the Ising model with $d = 1$.

7. Let

$$\eta_x = \ln \frac{\mu_{x \to a_x}(1)}{\mu_{x \to a_x}(-1)}.$$

Derive the recurrence

$$\eta_{x+1} = 2\beta(B-1) + \ln \frac{1 + \exp(\eta_x + 2\beta)}{1 + \exp(\eta_x - 2\beta)}$$

and show that the limit $\eta^* = \lim_{x \to \infty} \eta_x$ exists.

8. Derive an expression for $\lim_{n \to \infty} \frac{1}{n} \ln Z$ in terms of η^*.