To be solved in Class

1. **Probability theory** Prove Bayes’ theorem.

   **Theorem 1 (Bayes’ theorem)** Let $E_1$ and $E_2$ be probability events with $P[E_2] \neq 0$. Then,
   \[ P[E_1|E_2] = \frac{P[E_1] \cdot P[E_2|E_1]}{P[E_2]} . \]

2. Proof the **union bound** which states that for arbitrary events $E_1, E_2$, we have
   \[ P[E_1 \cup E_2] \leq P[E_1] + P[E_2] . \]

3. Show that expected values are linear. For two real random variables $X$ and $Y$ with joint distribution $P_{XY}$ and arbitrary real numbers $a, b \in \mathbb{R}$, it holds that
   \[ E_{XY}[aX + bY] = a \ E_X[X] + b \ E_Y[Y] . \]

4. What is the probability that two (or more) students in our information-theory class have the same birthday? Let us assume that everybody was born in the same year.

5. ([MacKay] Example 2.3:) Jo has a test for a nasty disease. We denote Jo’s state of health by the random variable $A$ ($A = 1$ is Jo has the disease and $A = 0$ if not) and the test result by $B$ ($B = 1$ if the test is positive and $B = 0$ if the test is negative).

   The test is 95% reliable: in 95% of cases of people who really have the disease, a positive result is returned, and in 95% of cases of people who do not have the disease, a negative result is obtained. Finally, 1% of people of Jo’s age and background have the disease.

   If Jo has the test and it is positive, what is the probability that Jo has the disease?

6. ([MacKay], Example 2.13:) A source produces a character $x$ from alphabet $A = \{0, 1, 2, \ldots, 9, a, b, c, \ldots, z\}$. With probability $1/3$, $x$ is a uniformly random numeral $0, 1, 2, \ldots, 9$, with probability $1/3$, $x$ is a random vowel $\{a, e, i, o, u\}$ and with probability $1/3$, $x$ is one of the 21 consonants. Estimate the entropy of $X$. 

**Homework**

1. (a) [1 point] Compute the entropy of a perfectly shuffled (i.e. uniformly distributed over all possible orders) deck of 52 playing cards.

   (b) [2 points] Now suppose we have a perfectly shuffled big deck, consisting of two identical decks of 52 cards (so 104 cards in total). Compute the entropy of the shuffled big deck.

2. [2 points] Prove the following inequality for real numbers $p_1, p_2, \ldots, p_n \in [0, 1]$:

   \[(1 - p_1)(1 - p_2) \cdots (1 - p_n) \geq 1 - p_1 - p_2 - \ldots - p_n.\]

   *Hint:* For an event $E$, the event $\bar{E}$ is the event that $E$ does not occur, hence $\Pr[\bar{E}] = 1 - \Pr[E]$. Consider independent events $E_i$ with probabilities $p_i = P[E_i]$ and use the union bound.

3. [2 points] Entropy of functions of a random variable. Let $X$ be a discrete random variable. Show that the entropy of a function $g$ of $X$ is less than or equal to the entropy of $X$ by justifying the following steps:

   \[
   H(X) = H(X) + H(g(X)|X) \quad (1) \\
   = H(X, g(X)) \quad (2) \\
   = H(g(X)) + H(X|g(X)) \quad (3) \\
   \geq H(g(X)) \quad (4)
   
   4. Consider the following random experiment with two fair dice. First, the first die is thrown, and let the outcome be $A$. Then, the second die is thrown until the outcome has the same parity (even, odd) as $A$. Let this final outcome of the second die be $B$. The random variables $X, Y, Z$ are defined as follows:

   \[X = (A + B) \mod 2, \quad Y = (A \cdot B) \mod 2, \quad Z = |A - B|.\]

   (a) [1 point] Find the joint distribution $P_{AB}$.

   (b) [2 points] Determine $H(X), H(Y)$ and $H(Z)$.

   (c) [1 point] Compute $H(Z|A = 1)$.

   (d) [2 points] Compute $H(AB)$, i.e. the joint entropy of $A$ and $B$.

   (e) [2 points] A random variable $M$ describes whether the sum $A + B$ is larger than seven, between five and seven, or smaller than five. How much entropy is present in this random variable $M$?

5. For two distributions $P$ and $Q$ over $\mathcal{X}$, the relative entropy or Kullback-Leibler divergence is defined as

   \[
   D(P||Q) := \sum_{x \in \mathcal{X}} P(x) \log \frac{P(x)}{Q(x)}. 
   
   Note that if $Q(x) = 0$ for some $x$, then $D(P||Q) = \infty$.

   (a) [5 points] Prove that $D(P||Q) \geq 0$, and that equality holds if and only if $P = Q$.

   *Hint:* Use Jensen’s inequality.

   (b) [3 points] Show that the mutual information can be defined in terms of the relative entropy, i.e. that $I(X;Y) = D(P_{X,Y}||P_X P_Y)$

   (c) [1 point] Use (a) and (b) to prove that $H(X|Y) \leq H(X)$.