Information Theory Exercise Sheet #1

University of Amsterdam, Master of Logic, Fall 2014
Lecturer: Christian Schaffner
TA: Philip Schulz

Out: Wednesday, 29 October 2014
(due: Wednesday, 5 November 2014, 13:00)

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. Prove 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. **Expected values** Let X and Y be two real random variables with joint distribution P_{XY}.

   (a) Show that expected values are linear: For arbitrary real numbers $a, b \in \mathbb{R}$, it holds that

   $$\mathbb{E}_{XY}[aX + bY] = a\, \mathbb{E}_X[X] + b\, \mathbb{E}_Y[Y] \; .$$

   (b) Show that if X and Y are independent, it holds that

   $$\mathbb{E}_{XY}[X \cdot Y] = \mathbb{E}_X[X] \cdot \mathbb{E}_Y[Y] \; .$$

   Give a and example of a joint distribution P_{XY} for which $\mathbb{E}_{XY}[X \cdot Y] \neq \mathbb{E}_X[X] \cdot \mathbb{E}_Y[Y]$.

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 $\mathcal{A} = \{\texttt{0}, \texttt{1}, \texttt{2}, \ldots, \texttt{9}, \texttt{a}, \texttt{b}, \texttt{c}, \ldots, \texttt{z}\}$. With probability 1/3, x is a uniformly random numeral $\texttt{0}, \texttt{1}, \texttt{2}, \ldots, \texttt{9}$, with probability 1/3, x is a random vowel $\{\texttt{a}, \texttt{e}, \texttt{i}, \texttt{o}, \texttt{u}\}$ and with probability 1/3, x is one of the 21 consonants. Estimate the entropy of X.

Homework

1. **Email** Please send an email to Philip (P.Schulz@uva.nl) and Chris (c.schaffner@uva.nl) stating your name, the program and year you are following (e.g. 2nd year Master of Logic), and (at least) one sentence about your motivation to follow this course. `2 p.`

2. (a) Compute the entropy of a perfectly shuffled (i.e. uniformly distributed over all possible orders) deck of 52 playing cards (assuming that you intend to draw one card after the other without replacement). `2 p.`

   (b) 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 p.`

3. Prove the following inequality for real numbers $p_1, p_2, \ldots, p_n \in [0,1]$: `3 p.`

$$(1-p_1)(1-p_2)\cdots(1-p_n) \geq 1 - p_1 - p_2 - \ldots - p_n\,.$$

   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.

4. 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: `3 p.`

$$\begin{aligned} H(X) &= H(X) + H(g(X)|X) && (1)\\ &= H(X, g(X)) && (2)\\ &= H(g(X)) + H(X|g(X)) && (3)\\ &\geq H(g(X)) && (4) \end{aligned}$$

5. Consider the following random experiment with two fair (regular six-sided) 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 and Z are defined as follows:

$$X = (A+B) \mod 2, \qquad Y = (A \cdot B) \mod 2, \qquad Z = |A - B|.$$

   (a) Find the joint distribution P_{AB}. `1 p.`

   (b) Determine $H(X)$, $H(Y)$ and $H(Z)$. `3 p.`

   (c) Compute $H(Z|A=1)$. `1 p.`

   (d) Compute $H(AB)$, i.e. the joint entropy of A and B. `1 p.`

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

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

$$D(P||Q) := \sum_{\substack{x \in \mathcal{X} \\ P(x)>0}} P(x) \log \frac{P(x)}{Q(x)}\,.$$

   Note that if $Q(x) = 0$ for some x, then $D(P||Q) = \infty$.

   (a) Prove that $D(P||Q) \geq 0$, and that equality holds if and only if $P = Q$. `5 p.`
   Hint: Use Jensen's inequality.

   (b) The mutual information between two random variables X and Y is defined as $I(X;Y) := H(X) - H(X|Y)$. Show that the mutual information can be expressed in terms of the relative entropy, i.e. that $I(X;Y) = D(P_{XY}||P_X P_Y)$ `3 p.`

   (c) Use (a) and (b) to prove that $H(X|Y) \leq H(X)$. `2 p.`