

# Information Theory Exercise Sheet #1

## (Probability and Entropy)

University of Amsterdam, Master of Logic, Fall 2015

Lecturer: Christian Schaffner

TA: Mathias Madsen

Out: Wednesday, 28 October 2015  
(due: Wednesday, 4 November 2015, 9:00)

1. **Inverse Probabilities** Suppose a certain disease occurs in 1% of the population. Suppose further that there is a test which can detect the disease somewhat reliably, producing false positives in only 5% of healthy individuals, and false negatives in 5% of sick individuals. Given that a person tests positive, what is the probability that the person has the disease?
2. **Bayes Factors** The Bayes factor is the ratio

$$\frac{P(\mathcal{A}|\mathcal{B})}{P(\mathcal{A})}.$$

This is the factor by which the probability of event  $\mathcal{A}$  changes if event  $\mathcal{B}$  is observed. Prove that

$$\frac{P(\mathcal{A}|\mathcal{B})}{P(\mathcal{A})} = \frac{P(\mathcal{B}|\mathcal{A})}{P(\mathcal{B})}.$$

3. **Dependent Sets** Give an example of a joint distribution over three binary random variables  $X$ ,  $Y$ , and  $Z$  such that
  - (a)  $P_{X|Y}(x|y) = P_X(x)$  for all  $x$  and  $y$ ;
  - (b)  $P_{X|Z}(x|z) = P_X(x)$  for all  $x$  and  $z$ ;
  - (c)  $P_{X|Y,Z}(x|y,z) \neq P_X(x)$  for at least one  $x$ ,  $y$ , and  $z$ .
4. **Expectation of a Product** Prove that if  $X$  and  $Y$  are independent, then

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

Given an example that shows that this may fail if  $X$  and  $Y$  are dependent.

5. **Binary Entropy** Compute the entropy of a random variable  $X$  with

$$P_X(1) = 1 - P_X(0) = \frac{1}{3}.$$

6. **Sum Distribution** Let  $X$  and  $Y$  be two independent binary random variables with

$$P_X(1) = P_Y(1) = \frac{1}{2}.$$

Compute  $H(X + Y)$ .

7. **Shuffling Cards**

- (a) Compute the entropy of the probability distribution which models the shuffling of a deck of cards.  
 (b) Do the same for a double deck (with twice as many cards).

8. **Geometric Entropy** Suppose you flip a fair coin until it comes up heads the first time, and record the number of coin flips as the random variable  $X$ . Compute  $H(X)$ .

9. **Squares and Expectations** Use Jensen's inequality to derive an inequality between  $E[X^2]$  and  $E[X]^2$ . Use this inequality as an alternative proof that  $\text{Var}[X] \geq 0$ .

10. **Random Dice** Suppose that

$$\begin{aligned} P_X(x) &= 1/6, & x = 1, 2, \dots, 6; \\ P_{Y|X}(y|x) &= 1/x, & y = 1, 2, \dots, x. \end{aligned}$$

Compute  $H(X)$ ,  $H(Y|X)$ , and  $H(X, Y)$ .

11. **Mutual Information** Let the binary random variables  $X$  and  $Y$  be defined by the joint probability distribution in the following table:

	$X = 0$	$X = 1$
$Y = 0$	0.3	0.2
$Y = 1$	0.1	0.4

Compute  $I(X; Y)$ .

12. **Stirling's Approximation** Prove that

$$\ln(n!) \leq n \ln(n).$$

Then use the approximation  $\ln(n!) \approx n \ln(n)$  to prove that

$$\frac{1}{n} \ln \binom{n}{np} \approx -p \ln(p) - (1-p) \ln(1-p) = h(p),$$

where we assume that  $n$  is an integer,  $p \in (0, 1)$ , and  $np$  is an integer.

13. **The Weak Law of Large Numbers** In this exercise, you will prove that averages converge to expectations in a certain precise sense. The proof is using a number of steps, each of which is interesting in its own right.

- (a) **Mean and Variance of Averages** Suppose that  $X_1, X_2, \dots, X_n$  are independent and identically distributed variables sampled from a distribution with mean  $E[X]$  and variance  $\text{Var}[X]$ . Prove that

$$E \left[ \frac{1}{n} \sum_{i=1}^n X_i \right] = E[X] \quad (1)$$

$$\text{Var} \left[ \frac{1}{n} \sum_{i=1}^n X_i \right] = \frac{\text{Var}[X]}{n} \quad (2)$$

- (b) **The Markov Bound** Suppose that  $S$  is a random variable which only takes on non-negative values (that is,  $P(S \geq 0) = 1$ ). Prove that

$$P(S \geq s) \leq \frac{E[S]}{s}$$

(For instance, less than 1/5 of the population earns more than 5 times the average income.)

- (c) **Chebyshev's Inequality** Suppose  $X$  is a random variable with mean  $E[X]$  and variance  $\text{Var}[X]$ . Prove that

$$P(|X - E[X]| \geq \varepsilon) \leq \frac{\text{Var}[X]}{\varepsilon^2}.$$

- (d) **The Weak Law of Large Numbers** Suppose that  $X_1, X_2, \dots, X_n$  are i.i.d. random variables with a shared mean  $E[X]$  and variance  $\text{Var}[X]$ . Prove that

$$P \left( \left| \frac{1}{n} \sum_{i=1}^n X_i - E[X] \right| \geq \varepsilon \right) \leq \frac{\text{Var}[X]}{n\varepsilon^2}.$$

14. **Tail-heavy Distributions\*** Give an example of a random variable  $S$  for which the Markov bound holds with equality for every  $s$ .
15. **Convergence of Averages\*** Prove a stronger version of the weak law of large numbers which holds for any sequence of independent random variables with a shared mean  $E[X]$  and bounded variances  $\text{Var}[X_i] \leq B$ .

Homework for this week are exercises 7, 10, 11, and 13. The starred exercises are optional. To confirm you are following the course, please send an email with your student number, name, and study to ([mathias.winther@gmail.com](mailto:mathias.winther@gmail.com)).