02. Playing with Probabilities

To be completed Friday 11th of November, by 12 noon sharp. You are encouraged to submit in pairs, but are also allowed to submit alone. You must submit your solution as **a single pdf file**, to the folder on Ilias, by naming *exercise_01_name1_matriculationnumber1_name2_matriculationnumber2.pdf*, with the obvious replacements in the strings. **Not adhering to the formatting requirements may result in the submission not being graded.**
Exercises are only graded in a binary fashion as sufficient or insufficient. To be graded as sufficient, you do not necessarily have to have correct solutions to every sub-question, but you must have made a clear and earnest effort to solve the entire exercise. Ultimately, what constitutes sufficient is at the discretion of the tutors. To be admitted to the exam, you must have submitted sufficient answers to at least 5 of the 6 (maybe 7) exercise sheets.

1. **EXAMple Question —**
   A tournament has N levels and 2^N players with skills $S_1 > S_2 > ... > S_{2^N}$. At each level, random pairs are formed and the winner proceeds to the next level. The player with higher skills always wins.

   (a) What is the probability that the two most skilled players 1 and 2 do *not* meet at the first level?

   (b) What is the probability that the two most skilled players 1 and 2 will meet in the final level?
   (*Bonus*: can you compute this probability without using the result from part (a)?)

2. **Theory Question —**
   Information and *entropy* are fundamental concepts in probability, and are ubiquitous in statistical analysis. In this question we aim to provide you with an intuitive understanding of these quantities.

   Consider a fair coin C_f and a bent coin C_b. C_f has an equal probability of heads(H) and tails(T) when tossed, $p(\mathrm{C}_f = H) = 0.5$. C_b is biased, with $p(\mathrm{C}_b = H) = 0.99$.

   The *information* present in a random event x can be defined as a function of the probability of occurence of that event:
   $$I(x) = \log_2\left(\frac{1}{p(x)}\right) = -\log_2(p(x)).$$
   (Intuitively, information is the ability to distinguish possibilities. To distinguish 2^k possibilities you need k bits of information, hence the log form.)

   (a) Show that for a set of independently occurring events $\mathbf{x} = (x_1, x_2)$, $I(\mathbf{x}) = I(x_1) + I(x_2)$. For each of the coins C_f and C_b, compute the information contained in all possible events when the coin is tossed twice. Interpret these values with the intuition provided above.

   The *entropy* of a random variable X with probability distribution p_X is the *expected information* over all possible realizations of X:
   $$H(X) = \mathbb{E}_{p_X}[I(x)].$$
   (In the same way as information, entropy can be interpreted as the *average* number of bits needed to communicate each event of a random variable.)

   (b) For each of the two coins C_f and C_b, compute the entropy for a single toss of the coin. What is the trend of entropy vs information for different probabilities of heads for a coin? Use the above interpretation to justify your observation.

3. **Practical Question —**
   (*WordleBot*) In this exercise, you learn to apply the concepts of *information* and *entropy* to create a bot that can play the popular game Wordle. See `Exercise_02.ipynb`.