**Exercise 1 (1.5 points).** Show that $H(X,Y,Z)-H(X,Y) \leq H(X,Z)-H(X)$, and find conditions for equality. Show that $I(X;Y|Z) \geq I(Y;Z|X) - I(Y;Z) + I(X;Y)$, and find conditions for equality.

**Exercise 2 (2 points).** Let $X_0, X_1, \ldots$ be random variables with values in $\mathbb{Z}$, that describe a random walk with inertia: $X_0 = 0$ with probability 1, $X_1 = 1$ or $-1$ with equal probability $1/2$, and for each $n \in \mathbb{N}$,

$$\begin{cases} \mathbb{P}(X_{n+2} - X_{n+1} = X_{n+1} - X_n) = 9/10 \\ \mathbb{P}(X_{n+2} - X_{n+1} = -(X_{n+1} - X_n)) = 1/10 \end{cases}$$

Compute $H(X_0, \ldots, X_n)$.

**Exercise 3 (1 point).** Compute binary Huffman code and Shannon-Fano code of the sum $S = D_1 + D_2$ of two independent dices with 6 faces. Compare their efficiencies with the entropy of $S$. Suppose now that there are $N$ dices $D_1, \ldots, D_N$, with 6 faces each. When $N$ tends to $\infty$, what can you tell about the efficiencies of binary Huffman codes and Shannon-Fano codes for the sum of those $N$ dices, and the entropy of the sum $D_1 + \ldots + D_N$?
**Problem (2 points):** Suppose now there are only two dices, but with an even number $M$ of faces each, the same question as above.

**Exercise 4 (Axiomatization of entropy, 3 points).** For each $m$, let us denote

$$D_m = \{(x_1, \ldots, x_m) \in [0,1]^m \mid \sum_{i=1}^{m} x_i = 1\}$$

the set of probability distributions over $m$ events, and for each $m$, let $H_m : D_m \to [0, +\infty)$ be a positive function over the set of probability distributions. Suppose that, for any $m$ and $n$,

1. **Robustness.** $H_m$ is symmetric and continuous,
2. **Product of independent random events.**
   $H_{mn}(\frac{1}{mn}, \ldots, \frac{1}{mn}) = H_m(\frac{1}{m}, \ldots, \frac{1}{m}) + H_n(\frac{1}{n}, \ldots, \frac{1}{n})$
3. **Grouping two events.**
   $H_m(p_1, p_2, \ldots, p_m) = H_{m-1}(p_1+p_2, p_3, \ldots, p_m) + (p_1+p_2)H_2(\frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2})$.
4. **Normalization.**
   $H_2(\frac{1}{2}, \frac{1}{2}) = 1$.

Let $G : \bigcup_{m \in \mathbb{N}} D_m \to [0, +\infty)$ defined by $G(p_1, \ldots, p_m) = H_m(p_1, \ldots, p_m)$. The goal of this problem is proving that $G$ is equal to $H$, the measure of (binary) entropy.

**Hints:** First prove that you can add as many zeros as you want to a probability distribution without changing its value. Secondly, extrapolate the grouping property to handle an arbitrary number of events. Then prove by induction on $m$ that $G$ and $H$ coincide on binary distributions, i.e. distributions $(p_1, \ldots, p_{2^m})$ such that for each $1 \leq i \leq m$, there exists $n_i$ such that $p_i = n_i/2^m$. Conclude with a density argument.