Supplementary exercises for sheet 4

You don't have to hand out those supplementary exercises; if you struggle with one of them, Alix Deleporte (who wrote them) will be happy to help!

Useful things in probabilities

Exercise 1. True or false: independence.

1. If an event E is independent from an event F and if E is independent from an event G, then E is independent from $F \cup G$.
2. If three events A, B, C are two by two independent, then they are independent. Namely, if $P(A \cap B) = P(A)P(B), P(A \cap C) = P(A)P(C), P(B \cap C) = P(B)P(C)$, then $P(A \cap B \cap C) = P(A)P(B)P(C)$.

Exercise 2.

1. Let $0 \leq p_1 \leq p_2 \leq 1$. Give a probability space and two random variables X_1, X_2 on this space, such that X_1 and X_2 respectively follow a Bernoulli law of parameter p_1 and p_2, and such that $X_1 \leq X_2$ almost surely.
2. Generalise to a finite family of Bernoulli laws.
3. Once you learn about probabilities on general measure spaces, come back again to this exercise and generalise to an increasing family of random variables $(X_p)_{0 \leq p \leq 1}$, each of them following a Bernoulli law of parameter p.

Exercise 3. Let N be a random variable with values in $\mathbb{N}_0$. Show that

$$\mathbb{E}[N] = \sum_{n \geq 1} \mathbb{P}(N \geq n).$$

Exercise 4. For $0 \leq r < 1$, the *geometric law with parameter r* is the probability measure on $\mathbb{N}$ given by $p_k = (1-r)r^k$.

1. Prove that, indeed, this is a probability measure.
2. We let X_1 and X_2 be independent random variables on $\mathbb{N}$, respectively following geometric laws with parameters r_1, r_2. Show that $\min(X_1, X_2)$ follows a geometric law with parameter $r_1 r_2$ (hint: consider $\mathbb{P}(X_j \leq k)$).

Exercise 5. Let E be an interval of $\mathbb{R}$ and let $f : E \to \mathbb{R}$ be a convex function. Let X be a random variable, on a finite probability space, with values into E. We suppose that X and $f(X)$ have finite expectation. Show the Jensen inequality:

$$f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)].$$

One can proceed by induction, or draw a picture.

Exercise 6. Let Ω be a finite set. The *entropy* of a probability $\mathbb{P}$ on Ω is

$$H(\mathbb{P}) = -\sum_{\omega\in\Omega} \mathbb{P}(\{\omega\})\log(\mathbb{P}(\{\omega\})).$$

Here, we take the convention $0\log(0) = 0$.

If $\mathbb{P}_1$ and $\mathbb{P}_2$ are two probabilities on Ω and if $\mathbb{P}_2(\{\omega\}) > 0$ for all $\omega \in \Omega$, then the *relative entropy* of $\mathbb{P}_1$ with respect to $\mathbb{P}_2$ is

$$D(P\|Q) = \sum_{\omega\in\Omega} \mathbb{P}_1(\{\omega\})\log\left(\frac{\mathbb{P}_1(\{\omega\})}{\mathbb{P}_2(\{\omega\})}\right).$$

1. Show, using the Jensen inequality, that D is non-negative and that $D(P\|Q) = 0$ implies $P = Q$.
2. Deduce that, among probabilities on Ω, the entropy is maximal for the uniform law on Ω, and only for this one.
3. The definition of H extends to probabilities on an enumerable set. Find a probability on $\mathbb{N}$ whose entropy is infinite.
4. Given two finite sets Ω_1, Ω_2 and two probability measures P_1, P_2 respectively on Ω_1 and Ω_2, the product measure $P_1 \otimes P_2$ on $\Omega_1 \times \Omega_2$ is defined by the following: for any $(\omega_1, \omega_2) \in P_1 \times P_2$,

$$P_1 \otimes P_2(\{(\omega_1, \omega_2)\}) = P_1(\{\omega_1\})P_2(\{\omega_2\}).$$

   Show that $H(P_1 \otimes P_2) = H(P_1) + H(P_2)$.
5. Let $f : \Omega_1 \to \Omega_2$ be a map between finite sets and let P be a probability measure on Ω_1. The probability measure $f(P)$ on Ω_2 is defined as follows: for any $\omega_2 \in \Omega_2$,

$$f(P)(\{\omega_2\}) = P(f(\omega_1) = \omega_2).$$

   Show that $H(f(P)) \leq H(P)$; discuss equality.

One can interpret entropy as the creation of information: the information created by two independent probabilistic events is the sum of the informations created by those events, and any transformation decreases the amount of information available.

Exercise 7.

1. Show (again) that, given for every $m \in \mathbb{N}$ a non-decreasing family $(x_{m,n})_{n\geq 0}$ of positive real numbers, with limit $\overline{x_m} \in [0, +\infty]$, one has, as $n \to +\infty$,

$$\sum_{m\in\mathbb{N}} x_{m,n} \to \sum_{m\in\mathbb{N}} \overline{x_m}.$$

2. Deduce that, if $(X_n)_{n\geq 0}$ is a non-decreasing family of positive random variables (i.e. $X_n(\omega) \leq X_{n+1}(\omega)$ for all ω and n) converging (pointwise) to a random variable $\overline{X}$, then as $n \to +\infty$,
$$\mathbb{E}[X_n] \to \mathbb{E}[\overline{X}].$$
3. Deduce that, if $(X_n)_{n\in\mathbb{N}_0}$ is a family of positive random variables, then
$$\mathbb{E}\left[\sum_{n=0}^{+\infty} X_n\right] = \sum_{n=0}^{+\infty} \mathbb{E}[X_n].$$

Applications

Exercise 8. The keyring of a watchman contains 10 keys; only one of them opens the door of the guardroom. To enter it, he follows one of the two scenarios:

- Case A: he picks a key at random, tries it, puts it aside if it does not work, and so on.
- Case B: he picks a key at random, tries it, but leaves it on the keyring if it does not work, and so on.

One writes respectively X_A and X_B the random variables equal to the number of tries (including the successful one) before he can access the guardroom, in the first and second scenarios.

1. Determine the distribution of X_A and X_B.
2. The watchman uses method B once every three days on average. One specific days, after trying 8 keys he still hasn't opened the door. What is the probability that he used method B?

Exercise 9. Alice suggests to Bob to play dice with her. She shows him three fair dice with the following numbers on them:

- Dice 1: 1,2,5,6,7,9
- Dice 2: 1,3,4,5,8,9
- Dice 3: 2,3,4,6,7,8.

They will play the following game: Bob chooses a dice, then Alice chooses another, and they toss them; each of them wins if the number on their dice is strictly greater than the other.

1. Prove that the expected value on each dice is 5.
2. Prove nevertheless that Alice can always choose a dice to have a better chance to win than Bob.

Exercise 10. Alice wants Bob to play "heads or tails" with her but Bob is tired of Alice's incessant cheating. He suspects she rigged the coin and it will fall on head with a probability $p \in [0,1]$. He suggests the following game: they will repeatedly toss the coin (producing a sequence of iid random variables $(X_n)_{n\in\mathbb{N}}$, each of them following a Bernoulli law with parameter p), and group the tossings two by two. If $(X_{2n-1}, X_{2n}) = (\text{head}, \text{tails})$ then Bob wins, if $(X_{2n-1}, X_{2n}) = (\text{tails}, \text{head})$ then Alice wins; the two other cases are a draw.

1. Show that, indeed, this game is equivalent to repeatedly tossing, in an independent manner, a fair coin, but that there is a probability q (to be computed) of having a draw.
2. They toss the original coin once per second. Within a (large) time interval T, how many "fair tossings" do they produce?
3. Meditate using the exercise on entropy. Without knowing p, can you invent a better way to produce fair tossings? What if you know the value of p?

Exercise 11.

1. Prove that, for any X, Y independent random variables on a discrete probability set Ω, with values in $\mathbb{N}$, the generating function of $X+Y$ is $G_{X+Y} = G_X G_Y$.
2. What is the generating function for the uniform law on $\{2,\dots,12\}$?
3. Let X_1 and X_2 be iid random variables taking values in $\{1,\dots,6\}$. By studying the roots of the polynomial $G_{X_1} G_{X_2}$, prove that the law of $X_1 + X_2$ cannot be the uniform law on $\{2,\dots,12\}$.
4. How would you use two dice to generate the uniform law on $\{2,\dots,12\}$?

Exercise 12. n students take a tramway with n cars. They recognise and greet each other if they find themselves in the same car.

1. Compute the probability that none of the students will meet if each of them boards a random car, independently with uniform probability.
2. Show that this probability decreases if they board a random car, independently, but the law is not uniform.