PROBABILITY

Exercise sheet 6

Exercise 1 (C.d.f vs r.v.). *Let X be a random variable and F_X its cumulative distribution function. Prove that for any $x < y \in \mathbb{R}$*

(1) $\mathbb{P}_X(X < x) = F(x-)$
(2) $\mathbb{P}_X(X > x) = 1 - F(x)$
(3) $\mathbb{P}(X \in (x, y)) = F(y-) - F(x)$.
(4) $\mathbb{P}_X(X = x) = F(x) - F(x-)$.

Exercise 2. *Prove that a random variable X is discrete if and only if there is a countable set $S \subseteq \mathbb{R}$ such that for all $s \in S$ we have that $\mathbb{P}_X(X = s) > 0$ and $\mathbb{P}_X(X \in S) = 1$. We call S the support of the discrete random variable X.*

Exercise 3 (Binomial r.v. is the number of occurring events). *Prove that for $n \geq 1$, $p \in (0, 1)$ we have that*

$$\sum_{k=0}^{n} p^k (1-p)^{n-k} \binom{n}{k} = 1$$

and deduce that the Binomial law as defined in class really gives rise to a probability law of $(\mathbb{R}, \mathcal{F}_E)$.

Now, suppose we have n mutually independent events $E_1, \dots, E_k$ of probability p on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Consider the random number of events that occurs: $X = \sum_{i=1}^n 1_{E_i}$. Prove that X is a random variable and has the law $Bin(n, p)$.

Exercise 4 (Uniform distribution as a maximal entropy distribution). *Consider a discrete probability space $(S, \mathcal{P}(S), \mathbb{P})$ where S is finite. The entropy of the distribution is given by $-\sum_{s \in S} \mathbb{P}(\{s\}) \log \mathbb{P}(\{s\})$. You can think of entropy as of measuring the amount of surprise a random outcome of this distribution brings us. When having to choose a probability distribution on the set S with no a priori knowledge, it is reasonable to choose the distribution that assumes the least extra information, that maximises the surprise, i.e. that maximises the entropy. Prove that among all probability distributions on $\{1, \dots, n\}$ the uniform distribution has maximal entropy.*

Exercise 5 (Random variables in the random walk). *For $p \in (0, 1)$ consider the probability space $(\{-1, 1\}, \mathcal{P}(\{-1, 1\}), \mathbb{P}_p)$, where $\mathbb{P}_p(\{1\}) = p$. Consider the product probability space of n copies of this probability space, i.e. $(\{-1, 1\}^n, \mathcal{P}(\{-1, 1\}^n), \mathbb{P}_\Pi)$. Writing $(X_1, \dots, X_n)$ for any element of $\{-1, 1\}^n$, explain why each X_1 is a random variable. Explain why $S_n = X_1 + \cdots + X_n$ is a random variable. Why can you think of the model as of a (biased) random walk?*

- *Show that each step X_1 has the law of a Bernoulli random variable on $\{-1, 1\}$ of parameter p.*
- *Prove that $\frac{S_n + n}{2}$ has the law of a Binomial random variable with parameter p. How does $\frac{S_n + n}{2}$ behave as $n \to \infty$ for $p = 1/2$ and for $p \neq 1/2$?*

Exercise 6 (Random variables in the random graph)**.** *For $p \in (0,1)$ consider the Erdos-Renyi random graph model with vertex set $V = \{1, 2, \dots n\}$ and edge set $E \subseteq \{\{i,j\} : 1 \le i \neq j \le n\}$. In other words, consider the product probability space $|E|$ copies of $(\{0,1\}, \mathcal{P}(\{0,1\}), \mathbb{P}_p)$ where $\mathbb{P}_p(\{1\}) = p = 1 - \mathbb{P}_p(\{0\})$. Further, enumerate the edges as $e_1, e_2, \dots, e_N$ and assign the edge e_i to be present in the graph G_p if and only if $\omega_i = 1$.*

- *Show that the total number of edges of this graph is a random variable. What is its distribution?*
- *Consider the following exploration algorithm of the graph: to start you list all possible edges $\{i,j\}$ in some arbitrary order $\hat{E} = (\hat{e}_1, \hat{e}_2, \dots, \hat{e}_N)$ (that might be a relabelling of the initial order) and you pick a starting vertex i. Now*
  (1) You start from the vertex i, then pick the first edge $\hat{e}_j = \{i,k\}$ in the list that is incident to i and if this edge is present in the random graph you move to vertex k and remove this edge $\hat{e}_j$ from the list $\hat{E}$. Otherwise you stop.
  (2) At each step, you repeat this process form vertex k, picking the first edge among remaining edges in $\hat{E}$ that is incident to k. If no such edge exists, you also stop.

  Let L denote the number of edges traversed in this process. Prove that L is a random variable. For a fixed value of p, what does the probability $\mathbb{P}_\Pi(L = k)$ converge to as $n \to \infty$? Does anything change if you don't fix the order of the edges in the beginning, but rather on each step choose arbitrarily among edges that are still in the list $\hat{E}$?

0.1. $\star$ **For fun (non-examinable)** $\star$.

Exercise 7 ($\star$ Number of triangles)**.** *Consider the Erdös-Rényi graph described above. We call $\{i,j,k\}$ a triangle if all the edges $\{i,j\}, \{j,k\}, \{i,k\}$ are present in the random graph G_p. We call a triangle isolated if none of i, j, k has any other edges incident to it. Now consider $p = \lambda/n$. Prove that as $n \to \infty$, the number of isolated triangles T_n converges to the Poisson random variable in the sense that $\mathbb{P}_\Pi(T_n = k)$ converges as $n \to \infty$ to $\frac{\lambda^k e^{-\lambda}}{k!}$ for some parameter $\lambda > 0$.*