Homework accompanying the lecture „Basics in Applied Mathematics“

Homework 4

Hand in: Tuesday, 12.11.2024, after the lecture in the mailbox at the Math Institut
(Don't forget to put your name on your homework.
Please hand in your solutions in groups of two.)

Exercise 1 (4 points)

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a discrete probability space and X_1, X_2 two independent random variables on Ω. Prove:

a) If X_i are Poisson distributed with parameters $\lambda_1, \lambda_2 > 0$, then their sum $Y = X_1 + X_2$ is Poisson distributed with parameter $\lambda_1 + \lambda_2$.

b) If X_i are Poisson distributed with parameters $\lambda_1, \lambda_2 > 0$, then X_1 given $X_1 + X_2 = l$ is binomial distributed with parameter $n = l$ and $p = \frac{\lambda_1}{\lambda_1+\lambda_2}$, meaning

$$P(X_1 = k | X_1 + X_2 = l) = \binom{l}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^k \left(1 - \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^{l-k}$$

for $k = 0, 1, ..., l$.

Exercise 2 (4 points)

Let X be a discrete random variable on a discrete probability space $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$. Prove lemma 3.2 and remark 3.3 found in the lecture notes, more precise:

a) Show that $\mathbb{P}^X$ as defined in lemma 3.2 is a discrete a probability measure on $(\Omega^X, \mathcal{P}(\Omega^X))$.

b) Show that F is monotonically increasing, right-continuous and prove

$$\lim_{x \to -\infty} F(x) = 0 \quad \text{as well as} \quad \lim_{x \to \infty} F(x) = 1 \ .$$

Hint: For right-continuity, i.e. $\lim_{x \searrow x_0} F(x) = F(x_0)$ you can take any strictly falling sequence $x_1 > x_2 > ...$ converging to x_0 and split $\mathbb{R}$ into disjoint sets between the points $x_\bullet$. With Kolmogorov (iii) and some standard arguments on infinite series, one can then show the convergence $\lim_{N \to \infty} F(x_N) = F(x_0)$. (The other two convergence results work similarly.)

Exercise 3 (8 points)

Let $Y_1, Y_2, ...$ and $C_1, C_2, ...$ be independent random variables with the distribution

$$\mathbb{P}(Y_n = 1) = \frac{1}{2} = \mathbb{P}(Y_n = -1)$$
$$\mathbb{P}(C_n = 1) = \frac{1}{2} = \mathbb{P}(C_n = 0) \ .$$

Define the random variables $X_1, X_2, ...$ recursively by

$$X_1 = Y_1 \quad \text{and} \quad X_{n+1} = \mathbb{1}_{C_n=1} Y_{n+1} + \mathbb{1}_{C_n=0} X_n = \begin{cases} Y_{n+1} & \text{, if } C_n = 1 \\ X_n & \text{else} \end{cases} .$$

The sequence $X_1, X_2, ...$ could heuristically be generated by the following process: Randomly generate $X_1 = Y_1$ and flip a coin C_1. If $C_1 = 0$ (Tails), then X_2 is set to be X_1. Otherwise X_2 is generated independently. Continue like this such that every X_{n+1} is either independently generated or copied from the value of X_n depending on the coin flip C_n.

a) By induction over $i \in \mathbb{N}$ prove that $\mathrm{E}[X_i^2] = 1$ and $\mathrm{E}[X_i] = 0$.

b) By induction over $j \in \mathbb{N}_0$, prove that the mean $\mathrm{E}[X_i X_{i+j}]$ has the form $\frac{1}{2^j}$.

c) By induction over n prove that the variance $\mathbb{V}\Big[\sum_{i=1}^{n} X_i\Big]$ is no grater than $3n$.

d) Prove that $\mathbb{P}\Big(\Big|\frac{1}{n}\sum_{i=1}^{n} X_i\Big| > \varepsilon\Big) \xrightarrow{n\to\infty} 0$ for all $\varepsilon > 0$.

HINT: You may use without proof that X_i, C_n, Y_{n+1} are jointly independent for all $i \leq n \in \mathbb{N}$ and also the fact that $\mathrm{E}[f(X)g(Y)h(Z)] = \mathrm{E}[f(X)]\,\mathrm{E}[g(Y)]\,\mathrm{E}[h(Z)]$ holds for all independent random variables X, Y, Z and functions f, g, h.
Keep in mind that X_i and X_{i+j} are not independent, so $\mathrm{E}[X_i X_{i+j}] \neq \mathrm{E}[X_i]\,\mathrm{E}[X_{i+j}]$, however the product $X_i X_{i+j}$ is independent of $\mathbb{1}_{C_{i+j}=0} = f(C_{i+j})$, so $\mathrm{E}[\mathbb{1}_{C_{i+j}=0} X_i X_{i+j}] = \mathrm{E}[\mathbb{1}_{C_{i+j}=0}]\,\mathrm{E}[X_i X_{i+j}]$.

Programming exercise 4 (2 points)

In front of you is a plate of $n \geq 1$ cooked spaghetti. You successively knot two randomly selected spaghetti ends together until all the ends are knotted.
Please implement a function that simulates this procedure in the jupyter notebook *homework04.ipynb* found in the git repository. The task is mainly split into two parts:

a) Write a function that simulates the number of rings that yield the procedure at a given n and empirically approximate the expectation.

b) Write a function that returns the lengths of each ring in a vector and plot the empirical cumulative distribution of the largest at the

Note: The true expectation can be evaluated explicitly using mathematical induction. The distribution of the maximal length of a ring could be much harder to calculate.