Solutions to Homework 8

Exercise 1*. a) Because the random variables X_n are identically distributed and bounded, it holds that there exists $M > 0$ such that $|X_n(\omega)| \leq M$ for all $n \geq 1$ and $\omega \in \Omega$. (*Note:* all this could hold with probability 1 instead of $\forall \omega \in \Omega$). So it holds that

$$\frac{S_n}{n} - \frac{1}{n} \sum_{j=2}^{n} X_j = \frac{X_1}{n} \underset{n\to\infty}{\to} 0 \quad \text{almost surely}$$

Likewise, it holds for any $k \geq 1$ that

$$\frac{S_n}{n} - \frac{1}{n} \sum_{j=k}^{n} X_j \underset{n\to\infty}{\to} 0 \quad \text{almost surely}$$

meaning that

$$A = \left\{ \frac{S_n}{n} \text{ converges} \right\} = \left\{ \frac{1}{n} \sum_{j=k}^{n} X_j \text{ converges} \right\} \in \sigma(X_k, X_{k+1}, \ldots)$$

As this holds for every $k \geq 1$, this proves that $A \in \mathcal{T}$.

b) From the conclusion of Exercice 1.b), we know that $\mathbb{P}(A) = 1$ even when the X_n are just uncorrelated random variables, not necessarily independent. In order to find a counter-example, there remains therefore to find another event B.

To this end, let us consider the sequence $(Y_n,\, n \geq 1)$ of i.i.d. random variables such that $\mathbb{P}(\{Y_1 = +a\}) = 2/3$ and $\mathbb{P}(\{Y_1 = -2a\}) = 1/3$, and let Z be a random variable independent of the sequence $(Y_n,\, n \geq 1)$ such that $\mathbb{P}(\{Z = +1\}) = \mathbb{P}(\{Z = -1\}) = 1/2$. Let us finally define $X_n = Y_n \cdot Z$ for $n \geq 1$.

Choosing $a = 1/\sqrt{2}$, the random variables X_n have zero mean, unit variance and are uncorrelated, as

$$\mathbb{E}(X_n) = \mathbb{E}(Y_n) \cdot \mathbb{E}(Z) = 0, \quad \mathrm{Var}(X_n) = \mathbb{E}(X_n^2) = \mathbb{E}(Y_n^2) \cdot \mathbb{E}(Z^2) = \left(a^2 \, \frac{2}{3} + 4a^2 \, \frac{1}{3} \right) \cdot 1 = 2a^2 = 1$$

and for $n \neq m$:

$$\mathbb{E}(X_n \cdot X_m) = \mathbb{E}(Y_n \cdot Y_m) \cdot \mathbb{E}(Z^2) = \mathbb{E}(Y_n) \cdot \mathbb{E}(Y_m) \cdot 1 = (2a/3 - 2a/3)^2 = 0$$

Now, let $B = \{Z = +1\}$. The event B belongs to the tail σ-field $\mathcal{T}$ for the following reason: for any value of $n \geq 1$, the value of $X_n = Y_n \cdot Z$ determines the value of Z, as

$$Z = +1 \quad \text{if and only if} \quad X_n = +a \quad \text{or} \quad X_n = -2a$$

and likewise,

$$Z = -1 \quad \text{if and only if} \quad X_n = -a \quad \text{or} \quad X_n = +2a$$

So Z is measurable with respect to $\sigma(X_n) \subset \sigma(X_n, X_{n+1}, \ldots)$ for any $n \geq 1$, so Z is measurable with respect to $\mathcal{T} = \cap_{n\geq 1}\sigma(X_n, X_{n+1}, \ldots)$, i.e., $B = \{Z = +1\} \in \mathcal{T}$, but $\mathbb{P}(B) = 1/2 \notin \{0, 1\}$.

Exercise 2. a) Let us compute first

$$\mathbb{E}(S_1) = \frac{1}{2}\left(\frac{3S_0}{2} + \frac{S_0}{2}\right) = S_0$$

Assuming now that $\mathbb{E}(S_n) = S_0$ (more precisely, that the expectation stays constant over n coin tosses), let us compute $\mathbb{E}(S_{n+1})$:

$$\begin{aligned}\mathbb{E}(S_{n+1}) &= \mathbb{E}(S_{n+1}|\{X_1 = +1\})\,\mathbb{P}(\{X_1 = +1\}) + \mathbb{E}(S_{n+1}|\{X_1 = -1\})\,\mathbb{P}(\{X_1 = -1\}) \\ &= \frac{1}{2}\left(\mathbb{E}(S_{n+1}|\{S_1 = \frac{3S_0}{2}\}) + \mathbb{E}(S_{n+1}|\{S_1 = \frac{S_0}{2}\})\right) = \frac{1}{2}\left(\frac{3S_0}{2} + \frac{S_0}{2}\right) = S_0\end{aligned}$$

Note: The computation is slightly unorthodox here, but we will see a cleaner way to prove this later in the course.

b) Y_n is the sum of n i.i.d. random variables, as the following computation shows:

$$Y_n = \log\left(\frac{S_n}{S_0}\right) = \log\left(\prod_{j=1}^{n}\left(1 + \frac{X_j}{2}\right)\right) = \sum_{j=1}^{n}\log\left(1 + \frac{X_j}{2}\right)$$

and these random variables are bounded, so by the central limit theorem,

$$\frac{Y_n - n\,\mu}{\sqrt{n}\,\sigma} \underset{n\to\infty}{\overset{d}{\to}} Z \sim \mathcal{N}(0,1)$$

where $\mu = \mathbb{E}(\log(1 + X_1/2)) = \frac{1}{2}\left(\log(3/2) + \log(1/2)\right) \simeq -0.144$ and

$$\sigma^2 = \mathrm{Var}(\log(1 + X_1/2)) = \frac{1}{2}\left(\log(3/2)^2 + \log(1/2)^2\right) - \mu^2 \simeq 0.3$$

This is saying that for large n, we have

$$Y_n \simeq -0.144n + \sqrt{0.26n}\,Z \quad \text{in particular:} \quad Y_{100} \simeq -14.4 + 5.4\,Z$$

Therefore

$$\begin{aligned}\mathbb{P}(\{S_{100} > S_0/10\}) &= \mathbb{P}(\{S_{100}/S_0 > 1/10\}) = \mathbb{P}(\{Y_{100} > -\log(10)) \\ &\simeq \mathbb{P}\left(\left\{Z > \frac{-2.3 + 14.4}{5.4}\right\}\right) = \mathbb{P}(\{Z > 2.24\})\end{aligned}$$

which is roughly 1% (so you can imagine what $\mathbb{P}(\{S_{100} > S_0\})$ looks like ...).

Therefore, the process $(S_n,\, n \geq 1)$, unexpectedly perhaps, "crashes" to zero with high probability as n gets large, even though it seemed a priori a "fair game" with constant expectation. This is an important example among a large class of processes called "martingales"; we will come back to this!

Note: The random process $(S_n,\, n \geq 1)$ is not unrelated to the following *deterministic* process defined recursively as

$$x_0 \in \mathbb{N}^*, \quad x_{n+1} = \begin{cases} x_n/2 & \text{if } x_n \text{ is even} \\ 3x_n + 1 & \text{if } x_n \text{ is odd}\end{cases}$$

in which an even number gets multiplied by 1/2 and an odd number gets approximately multiplied by 3/2 (because it first gets multiplied by 3 and then necessarily divided by 2, as $3x_n + 1$ is even). So if you consider that even and odd numbers appear naturally with probability 1/2, then the two processes have something in common. But in the deterministic case, one has no proof that the process ultimately reaches the value 1 as n gets large: this is the famous Collatz conjecture, which remains unsolved until now.

Exercise 3. a) let us compute $\mathbb{E}(S_n) = \sum_{j=1}^n \mathbb{E}(X_j^{(n)}) = n\,\frac{\lambda}{n} = \lambda$ and

$$\mathrm{Var}(S_n) = \sum_{j=1}^n \mathrm{Var}(X_j^{(n)}) = n\,\frac{\lambda}{n}\,\left(1 - \frac{\lambda}{n}\right) = \lambda - \frac{\lambda^2}{n}$$

b) So $\mu = \lim_{n\to\infty} \mathbb{E}(S_n) = \lambda$ and $\sigma^2 = \lim_{n\to\infty} \mathrm{Var}(S_n) = \lambda$.

c) Let us compute the characteristic function of S_n:

$$\begin{aligned}\phi_{S_n}(t) &= \mathbb{E}(\exp(itS_n)) = \mathbb{E}(\exp(it\,(X_1^{(n)} + \ldots + X_n^{(n)}))) = \mathbb{E}(\exp(itX_1^{(n)}))\cdots\mathbb{E}(\exp(itX_n^{(n)})) \\ &= \left(\mathbb{E}(\exp(itX_1^{(n)}))\right)^n = \left(e^{it}\,\frac{\lambda}{n} + 1 - \frac{\lambda}{n}\right)^n = \left(1 + \frac{\lambda\,(e^{it}-1)}{n}\right)^n \underset{n\to\infty}{\to} \exp\left(\lambda\,(e^{it}-1)\right)\end{aligned}$$

This limiting function is the characteristic function of $Z \sim \mathcal{P}(\lambda)$. Indeed, one can check that

$$\phi_Z(t) = \mathbb{E}(\exp(itZ)) = \sum_{k\geq 0} e^{itk}\,\frac{\lambda^k\,e^{-\lambda}}{k!} = e^{-\lambda}\sum_{k\geq 0}\frac{(\lambda\,e^{it})^k}{k!} = \exp(\lambda\,(e^{it}-1))$$

which allows us to conclude that $S_n \underset{n\to\infty}{\overset{d}{\to}} Z$.

d) The computation of the characteristic function is similar here:

$$\mathbb{E}\left(e^{itT_n}\right) = \left(\frac{1}{n}\,e^{it} + \left(1 - \frac{1}{n}\right)\right)^{\lceil\lambda n\rceil} = \left(1 + \frac{1}{n}\,(e^{it}-1)\right)^{\lceil\lambda n\rceil} \underset{n\to\infty}{\to} \exp(\lambda(e^{it}-1))$$

and leads actually exactly to the same result: T_n converges in distribution towards a Poisson random variable Z of parameter λ.

e) No, as each random variable S_n is constructed from a different set of random variables $X_1^{(n)}, \ldots, X_n^{(n)}$, which depends on n. The same holds for the random variables T_n.