Homework 8

Exercise 1. (extended law of large numbers)
Let $(\mu_n,\, n \geq 1)$ be a sequence of real numbers such that

$$\lim_{n\to\infty} \frac{\mu_1 + \ldots + \mu_n}{n} = \mu \in \mathbb{R}$$

Let $(X_n,\, n \geq 1)$ be a sequence of square-integrable random variables such that

$$\mathbb{E}(X_n) = \mu_n, \quad \forall n \geq 1 \quad \text{and} \quad \mathrm{Cov}(X_n, X_m) \leq C_1 \exp(-C_2\,|m-n|), \quad \forall m, n \geq 1.$$

for some constants $C_1, C_2 > 0$ (the random variables X_n are said to be *weakly* correlated). Let finally $S_n = X_1 + \ldots + X_n$.

a) Show that

$$\frac{S_n}{n} \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} \mu$$

b) Application to auto-regressive processes: Let $(Z_n,\, n \geq 1)$ be a sequence of i.i.d. $\sim \mathcal{N}(0,1)$ random variables, $x, a \in \mathbb{R}$ and $(X_n,\, n \geq 1)$ be the sequence of random variables defined recursively as

$$X_1 = x, \quad X_{n+1} = aX_n + Z_{n+1}, \quad n \geq 1$$

For what values of $x, a \in \mathbb{R}$ does the sequence $(X_n,\, n \geq 1)$ satisfy the assumptions made in a)? Compute μ in this case.

Exercise 2. (another extension of the weak law of large numbers)
Let $(X_n,\, n \geq 1)$ be a sequence of i.i.d. square-integrable random variables such that $\mathbb{E}(X_1) = \mu \in \mathbb{R}$ and $\mathrm{Var}(X_1) = \sigma^2 > 0$.

Let $(T_n,\, n \geq 1)$ be another sequence of random variables, independent of the sequence $(X_n,\, n \geq 1)$, with all T_n taking values in the set of natural numbers $\mathbb{N}^* = \{1, 2, 3, \ldots\}$. Define

$$p_k^{(n)} = \mathbb{P}(\{T_n = k\}) \quad \text{for } n, k \geq 1 \qquad \left(\text{so } \sum_{k\geq 1} p_k^{(n)} = 1 \quad \forall n \geq 1\right)$$

a) Find a sufficient condition on the numbers $p_k^{(n)}$ guaranteeing that

$$\frac{X_1 + \ldots + X_{T_n}}{T_n} \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} \mu \tag{1}$$

Hint: You should use the law of total probability here: if A is an event and the events $(B_k,\, k \geq 1)$ form a partition of Ω, then:

$$\mathbb{P}(A) = \sum_{k\geq 1} \mathbb{P}(A \,|\, B_k)\, \mathbb{P}(B_k)$$

b) Apply the above criterion to the following case: each T_n is the sum of two independent geometric random variables $G_{n1} + G_{n2}$, where both G_n are distributed as

$$\mathbb{P}(\{G_n = k\}) = q_n^{k-1}\,(1 - q_n) \quad k \geq 1$$

where $0 < q_n < 1$.

b1) Compute first the distribution of T_n, as well as $\mathbb{E}(T_n)$, for each $n \geq 1$.

b2) What condition on the sequence $(q_n,\ n \geq 1)$ ensures that conclusion (1) holds?

Hint: Solving question b1) above may help you guessing what the answer to b2) should be.

Exercise 3*.

Let X and Y be random variables defined on common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Define

$$d(X,Y) = \mathbb{E}\left(\log_2\left(1 + \frac{|X-Y|}{1+|X-Y|}\right)\right).$$

a) First, we would like to confirm that $d(X,Y)$ is a distance metric. Show that $d(X,Y)$ satisfies the triangle inequality. That is, $d(X,Z) \leq d(X,Y) + d(Y,Z)$ for any X, Y, and Z.

Hint: the function $f(x) = \log_2(1+x)$ is sub-additive, e.g. $f(x+y) \leq f(x) + f(y)$.

Next, we would like to check if convergence with respect to $d(X,Y)$ is equivalent to convergence in probability (a distance metric with this property is sometimes called a Ky-Fan metric).

b) Let $(X_n, n \geq 1)$ be sequence of random variables and X be another random variable, all defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Show that if $X_n \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} X$ then $\lim_{n\to\infty} d(X_n, X) = 0$.

c) Is the converse true? That is, if $\lim_{n\to\infty} d(X_n, X) = 0$ then $X_n \underset{n\to\infty}{\overset{\mathbb{P}}{\to}} X$. If yes, prove the statement. If no, provide a counter example.