Book of exercises, MACS201 and MACS203

F. Roueff*

September 9, 2021

Contents

Disclaimer

Section 6 is out of the scope of MACS201.
MACS203 is only concerned with Section 6.

*Alain Durmus wrote most of the solutions when he was a teacher assistant of MACS201.

1 Hilbert spaces

Exercise 1.1. Let X and Y be two complex valued random variables in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, for some probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

1. Determine the constant $m = \operatorname{proj}(X | \operatorname{Span}(1))$.
2. Determine the random variable $Z = \operatorname{proj}(X | \operatorname{Span}(1, Y))$.

Exercise 1.2 (The Fourier basis is dense). The first questions of this exercise are dedicated to the proof of Theorem 1.3.1. Let f be a continuous 2π-periodic function and f_n be defined as in (1.10).

1. Determine the Fejér kernel J_n, which satisfies
$$\frac{1}{n}\sum_{k=0}^{n-1} f_k = \int_{\mathbb{T}} J_n(x-t)\, f(t)\, \mathrm{d}t \; .$$
2. Show that we can write, for all $t \in \mathbb{R}$,
$$J_n(t) = \frac{1}{2\pi}\sum_{k=-n+1}^{n-1}(1-|k|/n)\mathrm{e}^{\mathrm{i}kt} = \frac{1}{2\pi n}\left|\sum_{j=0}^{n-1}\mathrm{e}^{\mathrm{i}jt}\right|^2 \; .$$
3. Deduce that $J_n \geq 0$, $\int_{\mathbb{T}} J_n = 1$ and that for any $\epsilon \in (0, \pi]$,
$$\sup_{n\geq 1} n \sup_{\epsilon \leq |t| \leq \pi} J_n(t) < \infty \; .$$
4. Conclude the proof of Theorem 1.3.1.

Let now μ be a finite measure on $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$. Let F be a closed set in $\mathbb{T}$. Define $f_n(x) = (1 - n\, d(F, x))_+$, where $d(F, x) = \inf\{|y-x| \;:\; y \in F\}$.

5. Show that $f_n \to \mathbb{1}_F$ in $L^1(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mu)$.

By Proposition A.1.3, we know that μ is regular that is, for all $A \in \mathcal{B}(\mathbb{T})$,
$$\mu(A) = \inf\{\mu(U) \;:\; U \text{ open set } \supset A\} = \sup\{\mu(F) \;:\; F \text{ closed set} \subset A\} \; .$$

6. Deduce that for all $A \in \mathcal{B}(\mathbb{T})$ and all $\epsilon > 0$, there exists a continuous 2π-periodic function g_ϵ such that
$$\int |\mathbb{1}_A - g_\epsilon| \mathrm{d}\mu \leq \epsilon \; .$$
7. Deduce that the set of continuous 2π-periodic functions is dense in $L^1(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mu)$ endowed with the L^1 norm.
8. Deduce that the set of continuous 2π-periodic functions is also dense in $L^2(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mu)$ endowed with the L^2 norm.
9. Conclude the proof of Corollary 1.3.2.

2 Probability

Exercise 2.1 (Pseudo-inverse d'une fonction de répartition, ordre stochastique). Let F be the distribution function of a probability on $\mathbb{R}$. Define (as usual) its *pseudo-inverse* by

$$F^{-1}(t) = \inf\{x : \ F(x) \geq t\}, \qquad t \in (0,1) .$$

1. Show that for all $t \in (0,1)$, $F \circ F^{-1}(t) \geq t$, with equality if F is continuous.
2. Show that for all $x \in \mathbb{R}$, $F^{-1} \circ F(x) \leq x$, , with equality if F is (strictly) increasing.
3. Let X be a r.v. with distribution function F; show that $X = F^{-1}(F(X))$ a.s.
4. Show that if F is continuous, then $F(X)$ has a uniform distribution on $[0,1]$. Compute $\mathbb{E}\left[F^n(X)\right]$ for all $n \in \mathbb{N}$.
5. Show that if U is a uniform r.v. on $[0,1]$, then $F^{-1}(U)$ has distribution function F.

Let X and Y two r.v. with distribution functions F and G. Suppose that for all $s \in \mathbb{R}$, $F(s) \leq G(s)$. We say that F is stochastically larger than or equal to G and we denote $G \leq_{sto} F$.

6. Provide examples of F and G such that $G \leq_{sto} F$.
7. Show that $G \leq_{sto} F$ if and only if there exists a random bidimensional vector (X,Y) such that $X \sim F$, $Y \sim G$ (this is called a *coupling* with marginals (F,G)) and $Y \leq X$ a.s.
8. Provide an example of F and G such that $G \leq_{sto} F$ and a coupling (X,Y) with marginals (F,G) for which $\mathbb{P}(Y \leq X) < 1$.
9. Show that if $G \leq_{sto} F$, then for any non-decreasing $f : \mathbb{R} \to \mathbb{R}$, we have $\mathbb{E}[f(Y)] \leq \mathbb{E}[f(X)]$ for any coupling (X,Y) with marginals (F,G).

Exercise 2.2. Let Y be an L^2 real valued r.v. defined on $(\Omega, \mathcal{A}, \mathbb{P})$. Let $\mathcal{B}$ be a sub-σ-field of $\mathcal{A}$. Define the conditional variance of Y given $\mathcal{B}$ by

$$\sigma^2(Y|\mathcal{B}) = \mathbb{E}[Y^2 \mid \mathcal{B}] - (\mathbb{E}[Y \mid \mathcal{B}])^2.$$

Denoting by $\sigma^2(Z)$ the variance of Z, show that

$$\sigma^2(Y) = \sigma^2(\mathbb{E}[Y \mid \mathcal{B}]) + \mathbb{E}[\sigma^2(Y|\mathcal{B})].$$

What is the result when Y is independent of $\mathcal{B}$?

Exercise 2.3. Show all the properties of Proposition 2.1.5 and also the following ones.

(a) If $\sigma(X) \vee \mathcal{H} = \sigma(\sigma(X) \cup \mathcal{H})$ (the smallest σ-field containing $\sigma(X)$ and $\mathcal{H}$) is independent of $\mathcal{G}$, $\mathbb{E}\left[X | \mathcal{H} \vee \mathcal{G}\right] = \mathbb{E}\left[X | \mathcal{H}\right]$. [Hint : first take an element of $\mathcal{H} \vee \mathcal{G}$ of the form $A \cap B$. Use the $\pi - \lambda$-theorem to conclude.]

(b) Let $X = F(Y,Z)$ where Y and Z are two random vectors valued in $\mathbb{R}^p$ and $\mathbb{R}^q$, respectively, and F measurable from $(\mathbb{R}^{p+q}, \mathcal{B}(\mathbb{R}^{p+q}))$ to $(\mathsf{X}, \mathcal{X})$. Suppose moreover that Y is $\mathcal{G}$-measurable and Z is independent of $\mathcal{G}$, then the conditional distribution of X given $\mathcal{G}$ is given by

$$\mathbb{P}^{X|\mathcal{G}}(\omega, A) = \mathbb{P}(F(Y(\omega), Z) \in A) \quad \text{for all } \omega \in \Omega \text{ and } A \in \mathcal{X} .$$

[Hint : first compute $\mathbb{P}^{(Y,Z)|\mathcal{G}}(\cdot, B \times C)$ for $B \in \mathcal{B}(\mathbb{R}^p)$ and $C \in \mathcal{B}(\mathbb{R}^q)$ and deduce $\mathbb{P}^{(Y,Z)|\mathcal{G}}(\cdot, D)$ for $D \in \mathcal{B}(\mathbb{R}^{p+q})$.]

Exercise 2.4. Let $\mathbf{X}$ and $\mathbf{Y}$ be as in Proposition 2.1.11.

1. In order to prove Proposition 2.1.11(i) and (ii), use Property (b) above.

2. Use the characterization of the orthogonal projection to prove Proposition 2.1.11(iii).

Exercise 2.5. Let μ and λ be two σ-finite measures on $(\Omega, \mathcal{F})$ and let a Borel function $\phi : \Omega \to \bar{\mathbb{R}}_+$ satisfy, for all $A \in \mathcal{F}$,

$$\mu(A) = \int_A \phi \, \mathrm{d}\lambda \,. \tag{1}$$

1. Show that if a Borel function $\psi : \Omega \to \bar{\mathbb{R}}$ satisfies, for all $A \in \mathcal{F}$,

   $$\int_A \psi \, \mathrm{d}\lambda = 0 \,,$$

   then $\psi = 0$ λ-a.e.
2. Show that , for all $A \in \mathcal{F}$ such that $\lambda(A) = 0$, we have $\mu(A) = 0$.
3. Show that $\phi > 0$ μ-a.e.
4. Give an example where we do not have that $\phi > 0$ λ-a.e.
5. Show that $\phi < \infty$ λ-a.e. [Hint : first consider the case where μ is finite]
6. Show that (1) uniquely defines ϕ up to a λ-null set.

Exercise 2.6. Prove Theorem 2.2.2.

Exercise 2.7. Let $(\mathbf{X}, \mathbf{Y})$ be a random vector valued in $\mathbb{R}^{p+q}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Show that the conditional density of $\mathbf{X}$ given $\mathbf{Y}$ is always well defined when

1. $\mathbf{X}$ and $\mathbf{Y}$ are discrete random variables (i.e. take values in countable sets).
2. $(\mathbf{X}, \mathbf{Y})$ admit a density with respect to the Lebesgue measure.
3. $\mathbf{Y}$ is a discrete random variable [Hint : show that $(\mathbf{X}, \mathbf{Y})$ admits a density with respect to $\xi \otimes \xi'$ with ξ' a counting measure and $\xi = \mathbb{P}^{\mathbf{X}}$.].
4. Deduce a formula for $\mathbb{E}[\mathbf{X}|\, \mathbf{Y}]$ in all the previous cases.
5. Give a new proof of Proposition 2.1.11(i) and (ii) in the case where $(\mathbf{X}, \mathbf{Y})$ has an invertible covariance matrix.

Exercise 2.8. Let X be a real valued random variable admitting a symmetric density f (with respect to the Lebesgue measure), $f(-x) = f(x)$ for all $x \in \mathbb{R}$.

1. Compute the conditional distribution of X given $|X|$.
2. Let Y be any other random variable defined on the same probability space as X and let $\mathbb{P}^{X|Y}$ denote the conditional distribution of X given Y. Show that for all Borel set A with null Lebesgue measure, we have that, for $\mathbb{P}^Y$-a.e. y, $\mathbb{P}^{X|Y}(y, A) = 0$.
3. Do we have that $\mathbb{P}^{X|Y}(y, \cdot)$ admits a density (with respect to the Lebesgue measure) for $\mathbb{P}^Y$-a.e. y?

Exercise 2.9. Let (X, Y) be an $\mathbb{R}^2$-valued r.v. Determine the conditional expectation and distribution of X given Y in the following choices for the distribution of (X, Y) :

1. the uniform distribution on the triangle $(0, 0), (1, 0), (0, 1)$
2. the uniform distribution on the square $(0, 0), (1, 0), (1, 1), (0, 1)$

Exercise 2.10. Let X and Y be two r.v.s defined on the same probability space. Suppose that X has its values in $\mathbb{N}$ and Y follows an exponential distribution with unit mean. Suppose also that the conditional distribution of X given Y is Poisson with mean Y. Determine the distribution of (X, Y) and that of X. Compute the conditional distribution of Y given X.

Exercise 2.11. Let $X_1, \ldots, X_p$ be independent r.v.'s following Poisson distributions with parameters $\lambda_1, \ldots, \lambda_p$.

1. Determine the conditional distribution of $(X_1, \ldots, X_{p-1})$ given $X_1 + \ldots + X_p$.
2. Compute $\mathbb{E}[X_1 \mid X_1 + X_2]$.

Exercise 2.12. Let $X_1, \ldots, X_n$ be i.i.d. random variables with density f, assumed to be continuous on $\mathbb{R}$.

1. Recall that the order statistic $(X_{(1)}, \ldots, X_{(n)})$ obtained by ordering$X_1, \ldots, X_n$ in an increasing order admits a density.
2. Determine the conditional distribution of $\min_{1\le i\le n} X_i$ given $\max_{1\le i\le n} X_i$.
3. Assuming that $\mathbb{E}\left[|X_i|\right] < \infty$ for all $i = 1, \ldots, n$, deduce an expression of $\mathbb{E}[\min_{1\le i\le n} X_i \mid \max_{1\le i\le n} X_i]$.

Exercise 2.13 (Statistiques d'ordre). Let $X = (X_1, \ldots, X_n)$ be a random vector with density $f(x)$, $x \in \mathbb{R}^n$. Let $R = (R(1), \ldots, R(n))$ be *rank statistic* of X, that is, for all $i \in \{1, \ldots, n\}$,

$$R(i) = \sum_{j=1}^{n} \mathbb{1}_{\{X_i \ge X_j\}} .$$

1. Show that, with probability 1, there exists a permutation σ of $\{1, \ldots, n\}$ such that $X_{\sigma(1)} < \ldots < X_{\sigma(n)}$. What is the relationship between σ and R ?
2. Show that, for any permutation r of $\{1, \ldots, n\}$ and all borel function $g : \mathbb{R}^n \to \mathbb{R}_+$
$$\mathbb{E}\left[g(X_{\sigma(1)}, \ldots, X_{\sigma(n)})\mathbb{1}_{\{R=r\}}\right] = \int g(x_1, \ldots, x_n)\mathbb{1}_{\{x_1<\cdots<x_n\}} f(x_{r(1)}, \ldots, x_{r(n)})\, \mathrm{d}x_1 \ldots \mathrm{d}x_n$$
3. Deduce the conditional distribution of R given $(X_{\sigma(1)}, \ldots, X_{\sigma(n)})$.
4. What do you obtain when the X_i's are iid ?

Exercise 2.14. Let P and Q be two probabilities on $(\Omega, \mathcal{F})$.

1. Let X and Y be two measurable functions defined on $(\Omega, \mathcal{F})$ valued in $(\mathsf{X}, \mathcal{X})$ and $(\mathsf{Y}, \mathcal{Y})$, respectively, such that, under P and Q, X and Y are independent. Express $\mathrm{KL}\left(P^{(X,Y)} \parallel Q^{(X,Y)}\right)$ with $\mathrm{KL}\left(P^X \parallel Q^X\right)$ and $\mathrm{KL}\left(P^Y \parallel Q^Y\right)$.
2. Let $X_1, \ldots, X_n$ be measurable functions defined on $(\Omega, \mathcal{F})$ such that the X_i's are iid under P and Q. Deduce $\mathrm{KL}\left(P^{X_{1:n}} \parallel Q^{X_{1:n}}\right)$ from $\mathrm{KL}\left(P^{X_1} \parallel Q^{X_1}\right)$.

3 Mathematical statistics

Exercise 3.1. Let $X_1, \ldots, X_n$ be n i.i.d. r.v.s with common density

$$p_\theta(x) = \exp(\theta - x)\mathbb{1}_{\{x \geq \theta\}} \ , \theta \in \mathbb{R} \ .$$

We want to estimate the translation parameter θ.

1. Propose an unbiased estimator $\hat{\theta}_n$ based on the empirical mean.
2. Compute the quadratic risk of $\hat{\theta}_n$.

Define the estimator $\tilde{\theta}_n$ by $\tilde{\theta}_n = \inf_{1 \leq i \leq n} X_i$.

3. What is the distribution of $\tilde{\theta}_n$?
4. Is the estimator $\tilde{\theta}_n$ unbiased ?
5. Compute the quadratic risk of $\tilde{\theta}_n$.
6. What is the best estimator according to these results?

Exercise 3.2. Let $X_1, \ldots, X_n$ be n i.i.d. r.v.'s with distribution $\mathbf{Ber}(\theta)$ (Bernouilli with parameter θ).

1. Show that $S_n = X_1 + \cdots + X_n$ is a sufficient statistic of the model.
2. Show that one can choose $\alpha > 0$ such that $\alpha(X_1 - X_2)^2$ is an unbiased estimator of the variance.
3. Compute the Rao Blackwellized estimator

$$Z = \mathbb{E}_\theta[\alpha(X_1 - X_2)^2 | S_n].$$

Exercise 3.3. Let $X_1, \ldots, X_n$ n i.i.d. r.v.s with distribution $\mathbf{Pn}(\theta)$ (Poisson distribution with parameter $\theta > 0$) defined by

$$\boldsymbol{P}_\theta\{X_1 = k\} = \frac{\theta^k}{k!} e^{-\theta}, \quad k \in \mathbb{N}.$$

Define $S_n = \sum_{i=1}^n X_i$.

1. Show that if $Y_1 \sim \mathbf{Pn}(\theta_1)$, $Y_2 \sim \mathbf{Pn}(\theta_2)$ and Y_1, Y_2 are independent, then $Y_1 + Y_2 \sim \mathbf{Pn}(\theta_1 + \theta_2)$.
2. Show that S_n is a sufficient statistic.

Let $x_0 \in \mathbb{N}$ and $X \sim \mathbf{Pn}(\theta)$. We want to estimate $\boldsymbol{P}_\theta(X \geq x_0)$ from $X_1, \ldots, X_n$. We consider the Rao-Blackwellized version of $\mathbb{1}(X_1 \geq x_0)$, that is

$$\widehat{P}_{x_0} := \mathbb{E}_\theta[\mathbb{1}(X_1 \geq x_0) | S_n].$$

3. Compute $\widehat{P}_{x_0}$ explicitly by using 1.
4. What is the result when $x_0 = 0$ and when $x_0 = 1$? Comment the result.

Exercise 3.4 (Neyman-Pearson test: Gaussian variable with known mean). 1. Let Y be a centered Gaussian vector of length n. We want to test the hypothesis H_1: $Y \sim \mathcal{N}(0, \Sigma_1)$ versus H_0: $Y \sim \mathcal{N}(0, \Sigma_0)$ where Σ_0, Σ_1 are invertible covariance matrices. Show that the Neyman-Pearson test consists in comparing $Y^T(\Sigma_1^{-1} - \Sigma_0^{-1})Y$ to a threshold.

2. Let X and V be two centered Gaussian r.v.'s with respective variances σ_X^2 and σ_V^2. The variable X is a true signal of interest and V is a measure noise. We observe $Y = X + V$ in the form of n independent observations.

   Propose a statistical test at level α for detecting the presence of the signal X.

3. For the previous test, give the value of the threshold as a function of the quantiles of the χ^2 distributions.

Exercise 3.5 (Uniform distribution). Let $X_1, \ldots, X_n$ be n i.i.d. r.v.s with uniform distribution on $[0, \theta]$, $\theta > 0$.

1. Let $\theta_0 < \theta_1$. What is the form of the Neyman-Pearson test of H_0 : "$\theta = \theta_0$" against H_1 : "$\theta = \theta_1$".
2. Same question for $\theta_0 > \theta_1$.
3. Let $\theta_0 > 0$. Build a test at level α of H_0 : "$\theta = \theta_0$" against H_1 : "$\theta \neq \theta_0$" from the test statistic $\hat{\theta}_n = \sup_{1\le i\le n} X_i$.
4. Compute the power function of the previous test.

Exercise 3.6 (Mixture model and EM algorithm). Let $\{g_\theta,\, \theta \in \Theta\}$ be a parametric collection of density functions with respect to a measure ν on $(\mathsf{X}, \mathcal{X})$. Suppose that for all θ, θ',

$$\int |\ln g_{\theta'}|\ g_\theta\ \mathrm{d}\nu < \infty\ .$$

Let K be a positive integer. The parametric model

$$\left\{ \sum_{k=1}^{K} \alpha_k g_{\theta_k}(x)\, \nu(dx),\ (\theta_1, \ldots, \theta_K) \in \Theta^K,\ (\alpha_1, \ldots, \alpha_K) \in [0,1]^K,\ \sum_{k=1}^{K} \alpha_k = 1 \right\} \tag{2}$$

is called a *micture model.* We will denote $\eta = (\theta_1, \ldots, \theta_K, \alpha_1, \ldots, \alpha_K)$ the corresponding parameter and $\mathbb{P}_\eta$ the associated distribution. In the following we observe $X_1, \ldots, X_n$ i.i.d. with distribution $\mathbb{P}_\eta$.

1. Show that one can define i.i.d. *hidden variables* $Y_1, \ldots, Y_n$ valued in $\{1, \ldots, K\}$ such that for all $i = 1, \ldots, n$ and all $A \in \mathcal{B}(\mathcal{X})$,

   $$\mathbb{P}_\eta(X_i \in A \,|\, Y_i = k) = \int_A g_{\theta_k}(x)\, \nu(dx)\ .$$

2. Compute $\mathbb{P}_\eta(Y_1 = k)$ for all $k = 1, \ldots, K$.
3. Compute the joint density f_η of (X_1, Y_1) and the distribution of Y_1 given X_1 under $\mathbb{P}_\eta$.

The maximum likelihood estimator for the model (2) is denoted by $\hat{\eta} = (\hat{\theta_1}, \ldots, \hat{\theta_K}, \hat{\alpha_1}, \ldots, \hat{\alpha_K})$ and maximizes

$$\eta = (\theta_1, \ldots, \theta_K, \alpha_1, \ldots, \alpha_K) \mapsto \sum_{k=1}^{K} \alpha_k g_{\theta_k}(x).$$

Even in the cases where $\{g_\theta,\, \theta \in \Theta\}$ is a well known simple collection of density functions (e.g. Gaussian), the max is difficult to compute. We can rely on the hidden variables abiove in order to implement the EM algorithm. Define

$$\begin{aligned}
\mathcal{Q}(X_{1:n}; \eta'; \eta) &= \mathbb{E}_\eta \left[\ln f_{\eta'}(X_1, \ldots, X_n, Y_1, \ldots, Y_n) \,|\, X_{1:n} \right] \\
&= \sum_{i=1}^{n} \mathbb{E}_\eta \left[\ln f_{\eta'}(X_i, Y_i) \,|\, X_i \right].
\end{aligned}$$

On notera

$$
\begin{aligned}
P_{i,k}(\eta) &:= \frac{\alpha_k\, g_{\theta_k}(X_i)}{\sum_{j=1}^K \alpha_j\, g_{\theta_j}(X_i)}, \\
T_{n,k}(\eta) &:= \sum_{i=1}^n P_{i,k}(\eta), \\
S_{n,k}(\eta,\theta') &= \sum_{i=1}^n P_{i,k}(\eta)\, \ln(g_{\theta'_k}(X_i)).
\end{aligned}
$$

4. (**Etape E**) Compute $\mathcal{Q}(X_{1:n};\eta'\,;\eta)$ explicitly.
5. (**Etape M**)
   (a) Compute $(\alpha_1^\star,\dots,\alpha_K^\star)$ that maximizes
   $$(\alpha'_1,\dots,\alpha'_K) \mapsto \mathcal{Q}(X_{1:n};\eta' = (\theta'_1,\dots,\theta'_K,\alpha'_1,\dots,\alpha'_K)\,;\eta)$$
   for $\eta, (\theta'_1,\dots,\theta'_K)$ given.
   (b) Take $\Theta = (0,\infty)$ and g_θ as the Gaussian density with mean zero and variance θ. Compute $\eta^\star$ that maximizes
   $$\eta' \mapsto \mathcal{Q}(X_{1:n};\eta'\,;\eta)\ .$$

Exercise 3.7. Let $\mathcal{P} = \{f_\theta(x)\,\nu(\mathrm{d}x),\ \theta\in\Theta\}$ and $\mathcal{Q} = \{g_\theta(x)\,\mu(\mathrm{d}x),\ \theta\in\Theta\}$ be Hellinger differentiable at θ with Fisher information matrices $\mathcal{I}(\theta)$ et $\mathcal{J}(\theta)$.

1. Show that $\{f_\theta(x)\otimes g_\theta(x)\ \nu(\mathrm{d}x)\otimes\mu(\mathrm{d}x),\ \theta\in\Theta\}$ is Hellinger differentiable at θ with Fisher information matrix $\mathcal{I}(\theta)+\mathcal{I}'(\theta)$.
2. Show that, for all $n\geq 1$, $\{\prod_{i=1}^n p_\theta(x_i)\,\mu^{\otimes n}(\mathrm{d}x),\ \theta\in\Theta\}$ is Hellinger differentiable at θ with Fisher information matrix $n\times\mathcal{I}(\theta)$.
3. Compare with Exercise 2.14.

4 Random processes: basic definitions

Exercise 4.1. Let X be a Gaussian vector, A_1 and A_2 two linear applications. Let us set $X_1 = A_1 X$ and $X_2 = A_2 X$. Give the distribution of (X_1, X_2) and a necessary and sufficient condition for X_1 and X_2 to be independent.

Exercise 4.2. Let X be a Gaussian random variable, with zero mean and unit variance, $X \sim \mathcal{N}(0,1)$. Let $Y = X\mathbf{1}_{\{U=1\}} - X\mathbf{1}_{\{U=0\}}$ where U is a Bernoulli random variable with parameter $1/2$ independent of X. Show that $Y \sim \mathcal{N}(0,1)$ and $\mathrm{Cov}\,(X,Y) = 0$ but also that X and Y are not independent.

Exercise 4.3. Let $n \geq 1$ and Γ be a $n \times n$ nonnegative definite hermitian matrix.

1. Find a Gaussian vector X valued in $\mathbb{R}^n$ and a unitary matrix U such that UX has covariance matrix Γ. [Hint : take a look at the proof of Proposition 4.2.3].

2. Show that
$$\Sigma := \frac{1}{2}\begin{bmatrix} \mathrm{Re}(\Gamma) & -\mathrm{Im}(\Gamma) \\ \mathrm{Im}(\Gamma) & \mathrm{Re}(\Gamma) \end{bmatrix}$$
is a real valued $(2n) \times (2n)$ nonnegative definite symmetric matrix.

Let X and Y be two n-dimensional Gaussian vectors such that
$$\begin{bmatrix} X & Y \end{bmatrix}^T \sim \mathcal{N}(0, \Sigma) \; .$$

3. What is the covariance matrix of $Z = X + \mathrm{i}Y$?

4. Compute $\mathbb{E}\left[ZZ^T\right]$.

The random variable Z is called a centered circularly-symmetric normal vector.

Let now T be an arbitrary index set, $\mu : I \to \mathbb{C}$ and $\gamma : T^2 \to \mathbb{C}$ such that for all finite subset $I \subset T$, the matrix $\Gamma_I = [\gamma(s,t)]_{s,t \in I}$ is a nonnegative definite hermitian matrix.

5. Use the previous questions to show that there exists a random process $(X_t)_{t \in T}$ valued in $\mathbb{C}$ such that, for all $s, t \in T$,
$$\mathbb{E}\left[X_t\right] = \mu(t) \quad \text{and} \quad \mathrm{Cov}\,(X_s, X_t) = \gamma(s,t) \; .$$

Exercise 4.4. Let $(\varepsilon_t)_{t \in \mathbb{Z}}$ be a sequence of i.i.d. real valued random variables. Determine in each of the following cases, if the defined process is strongly stationary.

1. $Y_t = a + b\varepsilon_t + c\varepsilon_{t-1}$ (a, b, c real numbers).

2. $Y_t = a + b\varepsilon_t + c\varepsilon_{t+1}$.

3. $Y_t = \sum_{j=0}^{+\infty} \rho^j \varepsilon_{t-j}$ for $|\rho| < 1$, assuming that $\mathbb{E}\left[|\epsilon_0|\right] < \infty$.

4. $Y_t = \varepsilon_t \varepsilon_{t-1}$.

5. $Y_t = (-1)^t \varepsilon_t$, $Z_t = \varepsilon_t + Y_t$.

Exercise 4.5. Let τ be a stopping time on the filtered measurable space $(\Omega, \mathcal{F}, (\mathcal{F}_k)_{k \in \mathbb{N}})$. Show that $\mathcal{F}_\tau$ in Definition 4.4.1 is a sub-σ-field of $\mathcal{F}$.

Exercise 4.6. Let τ be a $(\mathcal{F}_n)$-stopping time on $(\Omega, \mathcal{F}, \mathbb{P})$.

1. Let Y be a real random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. Show that Y is $\mathcal{F}_\tau$-measurable if and only if $Y\mathbb{1}_{\{\tau \leq n\}}$ is $\mathcal{F}_n$-measurable for all $n \in \mathbb{N}$.

2. Show that $A \in \mathcal{F}$ is $\mathcal{F}_\tau$-measurable if and only if $A \cap \{\tau = n\} \in \mathcal{F}_n$ for all $n \in \mathbb{N}$.

3. Let $n \in \mathbb{N}$ and Y be a real random variable $\mathcal{F}_n$-measurable. Prove that $Y\mathbb{1}_{\{\tau = n\}}$ is $\mathcal{F}_\tau$-measurable.

4. Let X be an integrable real random variable. Show that for all $n \in \overline{\mathbb{N}}$,
$$\mathbb{E}\left[\left. X\mathbb{1}_{\{\tau = n\}} \right| \mathcal{F}_\tau\right] = \mathbb{1}_{\{\tau = n\}}\mathbb{E}\left[\left. X \right| \mathcal{F}_n\right] \; .$$

5 Weakly stationary processes

Exercise 5.1. Let $(X_t)_{t\in\mathbb{Z}}$ and $(Y_t)_{t\in\mathbb{Z}}$ be two second order stationary processes that are uncorrelated in the sense that X_t and Y_s are uncorrelated for all t, s. Show that $Z_t = X_t + Y_t$ is a second order stationary process. Compute its autocovariance function, given the autocovariance functions of X and Y. Do the same for the spectral measures.

Exercise 5.2. Consider the processes of Exercise 4.4, with the additional assumption that $(\varepsilon_t)_{t\in\mathbb{Z}} \sim \mathrm{WN}(0, \sigma^2)$. Determine in each case, if the defined process is weakly stationary. In the case of Question 4, consider also $Z_t = Y_t^2$ under the assumption $\mathbb{E}[\varepsilon_0^4] < \infty$.

Exercise 5.3. Define χ as in (5.7).

1. For which values of ρ is χ an autocovariance function ? [Hint : use the Hertglotz theorem].
2. Exhibit a Gaussian process with autocovariance function χ.

Exercise 5.4. For $t \geq 2$, define

$$\Sigma_2 = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \ldots, \Sigma_t = \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & 1 & \rho \\ \rho & \cdots & \rho & 1 \end{bmatrix}$$

1. For which values of ρ, is Σ_t guaranteed to be a covariance matrix for all values of t [Hint: write Σ_t as $\alpha I + A$ where A has a simple eigenvalue decomposition]?
2. Define a stationary process whose finite-dimensional covariance matrices coincide with Σ_t (for all $t \geq 1$).

Exercise 5.5. Let X and Y two L^2 centered random variables. Define

$$\rho = \frac{\mathrm{Cov}\,(X, Y)}{\mathrm{Var}\,(Y)} ,$$

with the convention $0/0 = 0$. Show that

$$\mathrm{proj}\,(\,X|\,\mathrm{Span}\,(Y)) = \rho\, Y \quad \text{and} \quad \mathbb{E}\left[(X - \mathrm{proj}\,(\,X|\,\mathrm{Span}\,(Y)))^2\right] = \mathrm{Var}\,(X) - |\rho|^2\,\mathrm{Var}\,(Y) .$$

Exercise 5.6. Let (Y_t) be a weakly stationary process with spectral density f such that $0 \leq m \leq f(\lambda) \leq M < \infty$ for all $\lambda \in \mathbb{R}$. For $n \geq 1$, denote by Γ_n the covariance matrix of $[Y_1, \ldots, Y_n]^T$. Show that the eigenvalues of Γ_n belong to the interval $[2\pi m, 2\pi M]$.

Exercise 5.7. Let $X = (X_t)_{t\in\mathbb{Z}}$ be a centered weakly stationary process with spectral density f and denote by $\hat{X}$ its spectral representation field, so that, for all $t \in \mathbb{Z}$,

$$X_t = \int \mathrm{e}^{\mathrm{i}t\lambda}\, \mathrm{d}\hat{X}(\lambda) .$$

Assume that f is two times continuously differentiable and that $f(0) = 0$. Define, for all $t \geq 0$,

$$Y_t = X_{-t} + X_{-t+1} + \cdots + X_0 .$$

1. Build an example of such a process X of the form $X_t = \epsilon_t + a\epsilon_{t-1}$ with $\epsilon \sim \mathrm{WN}(0, 1)$ and $a \in \mathbb{R}$.
2. Determine g_t such that $Y_t = \int g_t\, \mathrm{d}\hat{X}$.
3. Compute

$$\lim_{n\to\infty} \int_{\mathbb{T}} \left| \frac{1}{n} \sum_{k=1}^{n} \mathrm{e}^{-\mathrm{i}k\lambda} \right|^2 \mathrm{d}\lambda$$

4. Show that
$$Z = \int (1 - \mathrm{e}^{-\mathrm{i}\lambda})^{-1} \, \mathrm{d}\hat{X}(\lambda) \, .$$
is well defined in $\mathcal{H}^X_\infty$.

5. Deduce from the previous questions that
$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} Y_t = Z \quad \text{in } L^2.$$

6. Show this result directly in the particular case exhibited in Question 1.

Exercise 5.8. Let $(X_t)_{t\in\mathbb{Z}}$ be a centered weakly stationary process with covariance function γ. Denote
$$\Gamma_t = \mathrm{Cov}\left(\left[X_1, \ldots, X_t\right]^T, =\right) \left[\gamma(i-j)\right]_{1\le i,j\le t} \qquad \text{for all } t \geq 1.$$
We temporarily assume that there exists $k \geq 1$ such that Γ_k is invertible but Γ_{k+1} is not.

1. Show that we can write X_n as $\sum_{t=1}^{k} \alpha_t^{(n)} X_t$, where $\alpha^{(n)} \in \mathbb{R}^k$, for all $n \geq k+1$.

2. Show that the vectors $\alpha^{(n)}$ are bounded independently of n.

Suppose now that $\gamma(0) > 0$ and $\gamma(t) \to 0$ as $t \to \infty$.

3. Show that, for all $t \geq 1$, Γ_t is invertible.

4. Deduce that Proposition 5.3.3 holds.

Exercise 5.9. Define $Z = X + Y$ with $X \sim \mathrm{WN}(0, \sigma^2)$ and $Y_t = Y_0$ for all t, where Y_0 is centered with positive variance and uncorrelated with $(X_t)_{t\in\mathbb{Z}}$.

1. Show that $\mathcal{H}^Z_{-\infty} \subseteq \mathrm{Span}\,(Y_0)$. [Hint : see Example 5.5.5]

Define, for all $t \in \mathbb{Z}$ and $n \geq 1$,
$$T_{t,n} = \frac{1}{n} \sum_{k=1}^{n} Z_{t-k}$$

2. What is the L^2 limit of $T_{t,n}$ as $n \to \infty$?

3. Deduce that $\mathcal{H}^Z_{-\infty} = \mathrm{Span}\,(Y_0)$.

Exercise 5.10. Define $(X_t)_{t\in\mathbb{Z}}$, $(U_t)_{t\in\mathbb{Z}}$ and $(V_t)_{t\in\mathbb{Z}}$ as in Theorem 5.5.2.

1. Show that
$$\mathcal{H}^X_{-\infty} \overset{\perp}{\oplus} \mathcal{H}^\epsilon_t = \mathcal{H}^X_t \, .$$

2. Deduce that $U_t = \mathrm{proj}\left(X_t \middle| \mathcal{H}^\epsilon_t\right)$, $V_t = \mathrm{proj}\left(X_t \middle| \mathcal{H}^X_{-\infty}\right)$ and that U and V are uncorrelated.

3. Show that $\mathcal{H}^X_{-\infty} = \mathcal{H}^V_t$ and $\mathcal{H}^\epsilon_t = \mathcal{H}^U_t$ for all $t \in \mathbb{Z}$. [Hint : observe that $\mathcal{H}^X_t \subset \mathcal{H}^U_t \oplus \mathcal{H}^V_t$ and use the previous questions]

4. Conclude the proof of Theorem 5.5.2.

6 Martingales

Exercise 6.1. Let $(X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ be a supermartingale such that $\mathbb{E}[X_n]$ is constant. Show that $(X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is a martingale.

Exercise 6.2. Let $(\epsilon_n)_{n\in\mathbb{N}}$ be a sequence of independent random variables with mean zero and variances given by $\mathrm{Var}[\epsilon_n] = \sigma_n^2$. Let $S_0 = T_0 = 0$ and for all $n \geq 1$,

$$S_n = \sum_{i=1}^{n} \epsilon_i \quad \text{et} \quad T_n = \sum_{i=1}^{n} \sigma_i^2 \, .$$

Show that $S_n^2 - T_n$ is a martingale.

Exercise 6.3. Let $\{X_n\}_{n\geq 0}$ and $\{Y_n\}_{n\geq 0}$ be two martingales L^2 (squared integrable) defined on the same filtered space.

1. Show that for all $m \leq n$, we have $\mathbb{E}[X_m Y_n | \mathcal{F}_m] = X_m Y_m$ a.s.
2. Show that $\mathbb{E}[X_n Y_n] - \mathbb{E}[X_0 Y_0] = \sum_{k=1}^{n} \mathbb{E}[(X_k - X_{k-1})(Y_k - Y_{k-1})]$.

Exercise 6.4 (Inégalité d'Azuma). Consider a martingale $((X_n, \mathcal{F}_n))_{n\in\mathbb{N}}$ and assume that there exist non negative real numbers $c_1, c_2, \ldots$ such that

$$|X_k - X_{k-1}| \leq c_k.$$

Denote by $Y_k = X_k - X_{k-1}$.

1. Show that, for $t > 0$ and $-c \leq x \leq c$,
$$\mathrm{e}^{tx} \leq \frac{\mathrm{e}^{ct} + \mathrm{e}^{-ct}}{2} + \frac{\mathrm{e}^{ct} - \mathrm{e}^{-ct}}{2c}\, x.$$
2. Denote by $f_c(x)$ the term on the right-hand side on the previous ineguality. Let Z be a random variable such that $|Z| \leq c$. Show that for any σ-field $\mathcal{F}$,
$$\mathbb{E}\left[\mathrm{e}^{tZ} \middle| \mathcal{F}\right] \leq f_c(\mathbb{E}[Z|\mathcal{F}]).$$
3. Let $t > 0$, show that for all a,
$$\mathbb{P}(Y_1 + \ldots + Y_m \geq a) \leq \mathrm{e}^{-ta}\mathbb{E}\left[\mathrm{e}^{tY_1 + \ldots + tY_m}\right].$$
4. Show that
$$\mathbb{E}\left[\mathrm{e}^{tY_n} \middle| \mathcal{F}_{n-1}\right] \leq f_{c_n}(0) \leq \mathrm{e}^{(c_n t)^2/2}.$$
5. Show that
$$\mathbb{P}(X_m - X_0 \geq a) \leq \exp\left(-at + \frac{t^2}{2}\sum_{k=1}^{m} c_k^2\right). \tag{3}$$
6. Find the minimizer of the term on the right-hand side in (3) and conclude that
$$\mathbb{P}(X_m - X_0 \geq a) \leq \exp\left(-\frac{a^2}{2\sum_{k=1}^{m} c_k^2}\right).$$

Exercise 6.5 (Jensen's inequality). Let X be a L^1 random variable valued in $]a,\ b[$ with $-\infty \leq a < b \leq +\infty$ and let $\Phi : (a,\ b) \to \mathbb{R}$ be a convex function. Assume that $\Phi(X)$ is L^1.

1. Show that
$$\Phi(\mathbb{E}[X]) \leq \mathbb{E}[\Phi(X)].$$
Hint: Use that Φ admits a left (and a right) derivative $\Phi'_<$ at any point $x \in (a, b)$, and, for all $y \in (a, b)$, $\Phi(y) \geq \Phi'_<(x)\,(y - x) + \Phi(x)$.

2. Let $\mathcal{G}$ be a sub σ-field. Show that

$$\Phi(\mathbb{E}[X \mid \mathcal{G}]) \leq \mathbb{E}[\Phi(X) \mid \mathcal{G}] \quad \mathbb{P}\text{-a.s.}$$

Hint: Show first that the inequality holds on $\{|\Phi(\mathbb{E}[X \mid \mathcal{G}])| \vee |\Phi'(\mathbb{E}[X \mid \mathcal{G}])| \leq n\}$ for any given $n \geq 1$.

3. Prove Proposition 6.1.2.

Exercise 6.6. Let $(Y_n)_{n\in\mathbb{N}}$ be a sequence of *i.i.d.* random variables with $\mathbb{P}\{Y_k = 1\} = \mathbb{P}\{Y_k = -1\} = 1/2$. Denote by $\mathcal{F}_0 = \{\emptyset,\ \Omega\}$, $\mathcal{F}_n = \sigma(Y_1,\ \ldots,\ Y_n)$, $S_0 = 0$ and $S_n = Y_1 + \ldots + Y_n$ for $n \geq 1$. Define $\operatorname{sgn}(x) = \mathbb{1}_{\{x>0\}} - \mathbb{1}_{\{x<0\}}$ and consider the random process $M_0 = 0$ and for all $n \geq 1$,

$$M_n = \sum_{k=1}^{n} \operatorname{sgn}(S_{k-1})Y_k.$$

1. What is the predicatble part of the Doob decomposition of the sub-martingale (S_n^2)?

2. Show that (M_n) is a martingale and compute the quadratic variation $\langle M\rangle$ of M.

3. What is the Doob decomposition of $(|S_n|)$? Show that M_n is measurable with respect to the σ-field $\sigma(|S_1|,\ \ldots,\ |S_n|)$.

Exercise 6.7 (Wald Part I). Let $(X_n)_{n\geq 1}$ be a sequence of *i.i.d.* random variables, with mean $\mu \in \mathbb{R}$. Consider τ a L^1 stopping time. Show that

$$\mathbb{E}\left[\sum_{k=1}^{\tau} X_k\right] = \mathbb{E}\left[\tau\right]\mu\ .$$

We shall begin with the case where $X_k \geq 0$.

Exercise 6.8. Consider a supermartingale $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, \mathbb{P})$. Assume that there exists a constant M such that for all $n \geq 1$,

$$\mathbb{E}\left[|X_n - X_{n-1}||\mathcal{F}_{n-1}\right] \leq M \text{ p.s.}$$

1. Show that if $(V_n)_{n\geq 1}$ is a non-negative random process such that for all $n \geq 0$, V_n is $\mathcal{F}_{n-1}$ measurable, then we have:

$$\mathbb{E}\left(\sum_{n=1}^{\infty} V_n|X_n - X_{n-1}|\right) \leq M\mathbb{E}\left(\sum_{n=1}^{\infty} V_n\right)$$

2. Let ν be a L^1 stopping time. We use the following identity

$$\mathbb{E}(\nu) = \sum_{n\geq 1} \mathbb{P}(\nu \geq n)$$

   (a) Deduce from question 1 that $\mathbb{E}\left[\sum_{n\geq 1} \mathbb{1}_{\{\nu\geq n\}}|X_n - X_{n-1}|\right] < +\infty$.

   (b) What is the value of $\sum_{n\geq 1} \mathbb{1}_{\{\nu\geq n\}}(X_n - X_{n-1})$? Conclude that X_ν is L^1.

3. Show that $(X_{\nu\wedge p})_{p\geq 0}$ converges towards X_ν in L^1 as p goes to infinity.

4. Conclude finally that if $\nu_1 \leq \nu_2$ are two stopping times with ν_2 in L^1 then

$$\mathbb{E}\left(X_{\nu_2}|\mathcal{F}_{\nu_1}\right) \leq X_{\nu_1}$$

   We can use the fact, after proving it, that if $A \in \mathcal{F}_{\nu_1}$, then $A \cap \{\nu_1 \leq k\} \in \mathcal{F}_{\nu_1\wedge k}$.

Exercise 6.9. On a probability space $(\Omega,\ \mathcal{F},\ \mathcal{P})$, we consider *i.i.d.* Bernoulli random variables with mean $1/2$, X_n^i, with $n \geq 1$ and $i \in \{1,\ 2\}$. Define $S_n^i = \sum_{k=1}^n X_k^i$ and $\nu_i = \inf\{n \geq 1 :\ S_n^i = a\}$ where $a \geq 1$ is an integer. Denote by $\nu = \nu_1 \wedge \nu_2$.

1. Show that $\mathbb{P}(\nu_i < +\infty) = 1$ for $i = 1,\ 2$.
2. For all $(i,j) \in \{1,\ 2\}^2$ and all $n \in \mathbb{N}$, define:

$$\begin{aligned} M_n^i &= 2S_n^i - n, \\ M_n^{i,j} &= (2S_n^i - n)(2S_n^j - n) - n\delta_{i,j}, \end{aligned}$$

with $\delta_{i,j} = \mathbb{1}_{\{i=j\}}$. Show that (M_n^i) and $(M_n^{i,j})$ are $\mathcal{F}_n$-martingales, with

$$\mathcal{F}_n = \sigma(X_k^i :\ i \in \{1,2\},\ k \leq n).$$

3. Show that $\mathbb{E}[\nu] \leq 2a$.
4. Show that $\mathbb{E}[M_\nu^{i,j}] = 0$.
5. Show that $\mathbb{E}[|S_\nu^1 - S_\nu^2|] \leq \sqrt{a}$, by looking at the martingale $M_n^{1,1} - 2M_n^{1,2} + M_n^{2,2}$.

Exercise 6.10 (Wald Part II). Let $(X_n)_{n\geq 1}$ be a sequence of *i.i.d.* random variables, centered with finite variance σ^2. Let τ be a L^1 stopping time. Show that

$$\mathrm{Var}\left(\sum_{k=1}^{\tau} X_k\right) = \mathbb{E}\left[\tau\right]\sigma^2 .$$

Exercise 6.11. Let (S_n) be a random walk on $\mathbb{Z}$ such that $S_0 = 0$ and $S_n = U_1 + \cdots + U_n$ for $n \geq 1$, where U_i are *i.i.d.* with $0 < \mathbb{P}\{U_i = 1\} = p = 1 - \mathbb{P}\{U_i = -1\} = 1 - q < 1$.

1. Define $Z_n = (q/p)^{S_n}$. Show that $\{Z_n\}_{n\geq 0}$ is a non-negative martingale.
2. Deduce, from an maximal inequality applied to the martingale $(Z_n)_{n\geq 0}$ that, for any $k \in \mathbb{N}$,

$$\mathbb{P}\left\{\sup_{n\geq 0} S_n \geq k\right\} \leq \left(\frac{p}{q}\right)^k,$$

and when $q > p$

$$\mathbb{E}\left[\sup_{n\geq 0} S_n\right] \leq \frac{p}{q-p}.$$

Exercise 6.12. Let $(X_n)_{n\geq 1}$ be a sequence of *i.i.d.* real-valued random variables with Gaussian law $\mathcal{N}(m,\ \sigma^2)$ and $m < 0$. Define $S_0 = 0$, $S_n = \sum_{i=1}^n X_i$, $\mathcal{F}_n = \sigma(S_0,\ \ldots,\ S_n)$ and

$$W = \sup_{n\geq 0} S_n.$$

The goal of the exercise is to derive some properties on the random variable W.

1. Show that $\mathbb{P}\{W < +\infty\} = +1$.
2. Recall that for $X_1 \sim \mathcal{N}(m,\sigma^2)$, $\mathbb{E}[\mathrm{e}^{\lambda X_1}] = \mathrm{e}^{\lambda^2\sigma^2/2}\mathrm{e}^{\lambda m}$. What is the value of $\mathbb{E}[\mathrm{e}^{\lambda S_{n+1}} \mid \mathcal{F}_n]$?
3. Show that there exists a unique $\lambda_0 > 0$ such that $(\exp(\lambda_0 S_n))_{n\in\mathbb{N}}$ is a martingale.
4. Show that, for any $a > 1$, we have:

$$\mathbb{P}\{\mathrm{e}^{\lambda_0 W} > a\} \leq \frac{1}{a}$$

and for all $t > 0$, $\mathbb{P}\{W > t\} \leq \mathrm{e}^{-\lambda_0 t}$.

5. Show that:
$$\mathbb{E}[e^{\lambda W}] = 1 + \lambda \int_0^{+\infty} e^{\lambda t}\mathbb{P}\{W > t\}dt.$$
Conclude that for $\lambda < \lambda_0$, $\mathbb{E}[e^{\lambda W}] < +\infty$. In particular, the random variable W has moments of every order.

Exercise 6.13. Consider a squared-integrable martingale $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, (M_n)_{n\geq 0}, \mathbb{P})$ such that $M_0 = 0$ a.s. Denote by $A_n = \langle M\rangle_n$ its quadratic variation (*i.e.* the predictable part of the Doob decomposition of $\{M_n^2\}_{n\geq 0}$). Let $a \in \mathbb{R}$, we define $\tau_a = \inf\{n \in \mathbb{N};\ A_{n+1} > a^2\}$.

1. Show that τ_a is a stopping time.
2. Show that:
$$\mathbb{P}\left(\sup_{n\geq 0} |M_{n\wedge\tau_a}| > a\right) \leq a^{-2}\mathbb{E}[A_\infty \wedge a^2].$$
3. Show that:
$$\mathbb{P}\left\{\sup_{n\geq 0} |M_n| > a\right\} \leq \mathbb{P}\left\{A_\infty > a^2\right\} + \mathbb{P}\left\{\sup_{n\in\mathbb{N}} |M_{n\wedge\tau_a}| > a\right\}. \tag{4}$$
4. Let X be a non-negative random variable. Show, using Fubini's theorem, that for any λ
$$\begin{aligned}\int_0^\lambda \mathbb{P}\{X > t\}dt &= \mathbb{E}[X\wedge\lambda],\\ \int_0^{+\infty} \frac{\mathbb{E}[X\wedge a^2]}{a^2}da &= 2\mathbb{E}[\sqrt{X}].\end{aligned}$$
5. Show that$\mathbb{E}[\sup_{n\geq 0}|M_n|] \leq 3\mathbb{E}[\sqrt{A_\infty}]$. We shall integrate (4) with respect to a between 0 and infinity.
6. Let$(Y_n)_{n\geq 1}$ be a sequence of *i.i.d* random variables, centered and squared-integrables. Denote by $S_0 = 0$ and for all $n \geq 1$, $S_n = Y_1 + \ldots + Y_n$. Consider the σ-field $\mathcal{F}_0 = \{\emptyset,\ \Omega\}$ and $\mathcal{F}_n = \sigma(Y_1,\ \ldots,\ Y_n)$ for $n \geq 1$. Show that if τ is a stopping time such that $\mathbb{E}[\sqrt{\tau}] < +\infty$ then $\mathbb{E}[S_\tau] = 0$.

Exercise 6.14. Let $X = (X_t)_{t\in T}$ and $Y = (Y_t)_{t\in T}$ be two real valued processes indexed by the same (arbitrary) set T. Define $U = X + Y$. Suppose that X and Y are U.I. We show in this exercise that U is then also U.I.

1. Let $p, q > 0$. Show that, for all $t \in T$,
$$|X_t|\mathbb{1}_{\{|U_t|>p+q\}} \leq |X_t|\mathbb{1}_{\{|X_t|>p\}} + p\mathbb{1}_{\{|Y_t|>q\}}\ .$$
2. Show that
$$\limsup_{q\to\infty}\sup_{t\in T}\mathbb{P}\left(|Y_t| > q\right) = 0\ .$$
3. For any $p > 0$, Deduce a bound of
$$\limsup_{q\to\infty}\sup_{t\in T}\mathbb{E}\left[|X_t|\mathbb{1}_{\{|U_t|>p+q\}}\right]\ .$$
4. Conclude that U is U.I.

In all the following, if (M_n) is a squared-integrable martingale, then its predictable quadratic variation is denoted by $\langle M\rangle_n$.

Exercise 6.15 (La "Martingale classique"). A player bets on the results of independent tosses of a coin (i.e. the probability of getting tails is equal to 1/2). At each turn, he bets an amount $S > 0$. If the coin lands on Tails, his capital increases by S, if it lands on Heads, the player loses his bet and so his bankroll decreases by S. A popular strategy in France in the 18th century France is called "La Martingale classique". It is defined as as follows:

- the player stops playing the first time he wins, *i.e.* as of the first Tails, his following bets are null and void and **his capital does not evolve anymore**;
- she doubles his bet at each turn, i.e. she bets the sum $S_n = 2^n$ at the n-th round, until she wins.

Let Y_n be the player's capital at time n (i.e. after n tosses of the coin). Let us assume that the initial capital is zero ($Y_0 = 0$, so in the first round he bets $S_1 = 2$ euros that he has to borrow), and that the player has the right to go into debt for an unlimited amount, that is, Y_n can become negative, arbitrarily large in absolute value. We denote by U_n the result of the n-th throw, $U_n = 1$ if the result is Tails and $U_n = 0$ otherwise, and by $\tau = \inf\{ngeq1 : \ U_n = 1\}$ the (random) time of the first "Tails" obtained. We consider the filtration $\mathcal{F} = \mathcal{F}_n\}$ where $\mathcal{F}_n$ denotes the σ-field generated by the (random) results of the n first throws, $\mathcal{F}_n = sigma(U_k : \ k \leq n)$.

1. What is the law of the duration of the game? Show that it is a $\mathcal{F}$-stop time. Can you explain in a heuristic way the strategy of the "classical Martingale"?
2. Show that the (S_n) strategy of "The classical Martingale" is $\mathcal{F}$-predictable.
3. Show that $(Y_n)_{in\mathbb{N}}$ is a martingale and more precisely the sequence stopped at the stopping time τ of a $\mathcal{F}$-martingale which we will specify.
4. Determine the increasing process $\langle Y \rangle_n$. Compute $\mathbb{E}[\langle Y \rangle_n]$ and discuss the convergence of Y_n in L^2.
5. Express Y_n as a function of tau and n, specify its law, and discuss the almost certain convergence of Y_n. What is its almost sure limit?
6. Does the sequence Y_n converge in L^1?
7. Now suppose that the bank does not allow the player to be indebted by more than an limit value L (we can assume that $L = 2^k$ for a $k \geq 1$). Therefore, the player is obliged to stop as soon as his capital at time n is strictly less than $-L + 2^{n+1}$. Let us note Z_n the player's capital at time n.
   (a) Let N be the duration of the game (i.e. the number of times the player bets a non-zero sum). Show that N is a stopping time and specify its law.
   (b) Is the process Z_n a martingale?
   (c) Discuss the almost-sure convergence and in L^1 of Z_n and comment on the results.

Exercise 6.16. Let $X_0 = 1$. We define a sequence $(X_n)_{nin\mathbb{N}}$ recursively by assuming that for all $n \geq 1$, X_n follows the uniform law on $[0, 2X_{n-1}]$ conditionally on $\sigma(X_k : \ k \leq n-1)$, i.e. we can define the sequence (X_n) recursively by: $\forall n \geq 1$, $X_n = 2U_nX_{n-1}$ where U_n are *i.i.d.* with uniform distribution on [0, 1].

1. Show that (X_n) is a square martingale integrable w.r.t. $\mathcal{F}_n = \sigma(U_k : \ k \leq n))$.
2. Compute the increasing process $\langle X \rangle_n$ and $\mathbb{E}[\langle X \rangle_n]$. Discuss the convergence of X_n in L^2. (Hint: X_n will be expressed as a function of $U_1, \ \ldots, \ U_n$ for all $n \geq 1$)
3. Discuss the almost sure convergence of X_n.

4. Determine the almost-sure limit of X_n. (Hint: Consider $Y_n = \log(X_n)$ and consider applying the strong law of large numbers)

5. Discuss the convergence of X_n in L^1.

Exercise 6.17. A player initially has the sum $X_0 = 1$. She plays a game of chance, in which she bets a proportion λ of her capital at each turn, with $0 < \lambda < 1$. She has a 50/50 chance of doubling her bet, otherwise she loses her bet. Precisely, the evolution of the capital X_n according to the time n is described by: $\forall n \in \mathbb{N}$,

$$X_{n+1} = (1-\lambda)X_n + \lambda X_n \xi_{n+1},$$

where the ξ_n are i.i.d., with $\mathbb{P}\{\xi_n = 2\} = \mathbb{P}\{\xi_n = 0\} = 1/2$.

1. Show that (X_n) is a square-integable $\mathcal{F}$-martingale, with $\mathcal{F}_n = \sigma(\xi_k,\ k \leq n)$.

2. Compute $\mathbb{E}[X_n]$.

3. Discuss the almost sure convergence of X_n when $n \to +\infty$.

4. Compute $\mathbb{E}[X_n^2]$ by induction on n.

5. What can we deduce about the convergence in L^2 of (X_n)?

6. Determine the increasing process $\langle X \rangle_n$.

7. We assume that the player bets at each turn the totality of her capital, i.e. $\lambda = 1$.
   (a) Compute explicitly the law of X_n.
   (b) Determine the almost sure limit of (X_n).
   (c) Discuss the convergence of X_n in L^1. Are X_n uniformly integrable?

Exercise 6.18. Let $(F_n, \mathcal{F}_n)_{n\geq 0}$ be a bounded martingale in L^1, $F_n^+ = \max(F_n, 0)$ and $F_n^- = -\min(F_n, 0)$.

1. Show that $(F_n^+)_{n\geq 0}$ and $(F_n^-)_{n\geq 0}$ are submartingales.

2. Let us denote $F_n^+ = M_n + A_n$ and $F_n^- = N_n + B_n$ the Doob-Meyer decomposition of these two submartingales (A_n and B_n are predictable processes with $A_0 = B_0 = 0$ and M_n and N_n are martingales). Show that $A_n = B_n$ $\mathbb{P}$-a.s. [we will show that $A_n - B_n$ is a martingale]

3. Show that $lim_{n\to\infty} A_n$ exists. We note A_∞ this limit. Show that $\mathbb{E}[A_\infty] < \infty$.

4. Let $F_n^\oplus = M_n + \mathbb{E}[A_\infty \mid \mathcal{F}_n]$ and $F_n^\ominus = N_n + \mathbb{E}[A_\infty \mid \mathcal{F}_n]$. Show that $(F_n^\oplus)_{n\geq 0}$ and $(F_n^\ominus)_{n\geq 0}$ are non negative martingales.

5. Show that $F_n = F_n^\oplus - F_n^\ominus$.

6. Show that $F_n^\oplus$ is bounded in L^1.

We have thus shown a result due to Krickeberg (1956). Any martingale bounded in L^1 is the difference of two non-negative martingales bounded in L^1.

Exercise 6.19. Let $(F_n)_{n\geq 0}$ be an adapted process and $(V_n)_{n\geq 1}$ a predictable predictable process, both with respect to the same filtration $(\mathcal{F}_n)_{n\geq 0}$. We define

$$(V \cdot F)_n = \sum_{k=1}^{n} V_k (F_k - F_{k-1}) \ ,$$

the martingale transform of $(F_k)_{k\geq 0}$ (a discrete time equivalent of the stochastic integral). For $n \geq 1$, $D_n = F_n - F_{n-1}$. We assume in the following that $F_0 = 0$ and $\sup_k |V_k| \leq 1$, $\mathbb{P}$-a.s.

1. First assume that $(F_n)_{n\geq 0}$ is a bounded martingale in L^2. Show that $\mathbb{E}\left[(V\cdot F)_n^2\right] \leq \mathbb{E}\left[F_n^2\right]$. Deduce that $(V\cdot F)_n$ converges $\mathbb{P}$-a.s. and in L^2.

2. We assume that $(F_n)_{n\geq 0}$ is a non-negative bounded submartingale, i.e. there exists $M > 0$ such that $\sup_n |F_n| \leq M$, $\mathbb{P}$-a.s.

   (a) Show that $\mathbb{E}\left[F_n^2\right] \geq \mathbb{E}\left[F_{n-1}^2\right] + \mathbb{E}\left[D_n^2\right]$.

   (b) Deduce that $\sum_{k=1}^n \mathbb{E}\left[D_k^2\right] \leq \mathbb{E}\left[F_n^2\right]$.

   (c) Let $\hat{F}_n = \sum_{k=1}^n \hat{D}_k$ where $\hat{D}_1 = D_1$ and for $k \geq 2$, $\hat{D}_k = D_k - \mathbb{E}\left[D_k | \mathcal{F}_{k-1}\right]$. Show that $(\hat{F}_n)_{n\geq 0}$ and $((V\cdot\hat{F})_n)_{n\geq 0}$ are martingales.

   (d) Show that $\mathbb{E}\left[\hat{F}_n^2\right] \leq \mathbb{E}\left[F_n^2\right]$ and $\mathbb{E}\left[((V\cdot\hat{F})_n)^2\right] \leq \mathbb{E}\left[F_n^2\right]$. Deduce that $(\hat{F}_n)_{n\geq 0}$ and $((V\cdot\hat{F})_n)_{n\geq 0}$ converge as. and in L^2. Deduce that $\sum_{k=1}^n \mathbb{E}\left[D_k | \mathcal{F}_{k-1}\right]$ converges a.s. to a finite random variable p.s. then that $\sum_{k=1}^n V_k \mathbb{E}\left[D_k | \mathcal{F}_{k-1}\right]$ converges a.s. to a a.s. finite random variable.

   (e) Deduce that $((V\cdot F)_n)_{n\geq 0}$ converges a.s.

3. Let $(F_n)_{n\geq 0}$ be a non -negative bounded submartingale in L^1 and with bounded increases: *i.e.* there exists $M > 0$ such that $\sup_n |D_n| \leq M$, $\mathbb{P}$-a.s. For $c \geq 0$, we denote by

$$\tau_c = \inf\left\{n \geq 0, F_n \geq c\right\} \ .$$

   We denote $F_n^{\tau_c} = F_{n\wedge\tau_c}$.

   (a) Show that $((V\cdot F^{\tau_c})_n)_{n\geq 0}$ converges p.s. Deduce that $((V\cdot F)_n)_{n\geq 0}$ converges p.s. on the event $\{\tau_c = \infty\}$.

   (b) Show that for any sequence $(c_n)_{n\geq 0}$ of positive numbers such that $lim_{n\to\infty} c_n = \infty$,

$$\mathbb{P}\bigcup_{n=1}^{\infty}\{\tau_{c_n} = \infty\} = 1 \ .$$

   (c) Deduce that $((V\cdot F)_n)_{n\geq 0}$ converges a.s.

4. Show using the results of exercise 6.18 that $((V\cdot F)_n)_{n\geq 0}$ converges a.s. if $(F_n)_{n\geq 0}$ is a bounded martingale in L^1.

Exercise 6.20. Let $(S_n)_{n\geq 0}$ be a symmetric random walk, $S_0 = 0$ and $S_n = S_{n-1} + \epsilon_n$, where $(\epsilon_n)_{n\geq 0}$ is a sequence of random variables i.i.d. with $\mathbb{P}(\epsilon_1 = 1) = 1/2$, $\mathbb{P}(\epsilon_1 = -1) = 1/2$. Let on the other hand $a \in \mathbb{N}^*$.

1. Show that $S_n^2 - n$ is a martingale.

2. Let $T = \inf\{n \geq 0, S_n \notin]-a, a[\}$. Show that $\mathbb{E}\left[T\wedge n\right] \leq a^2$ and deduce that $\mathbb{E}\left[S_T^2\right] = \mathbb{E}\left[T\right] = a^2$.

3. Compute the constants b and c such that

$$Y_n = S_n^4 - 6nS_n^2 + bn^2 + cn$$

   is a martingale.

4. Deduce an expression for $\mathbb{E}\left[T^2\right]$.

Exercise 6.21. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration $\{\mathcal{F}_n\}_{n\geq 0}$ and $X_1,\ X_2,\ \dots$ be a sequence of square integrable random variables on Ω such that: $\forall n \in \mathbb{N}$, $\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = 0$. Let $V_0 = 0$ and, for $n \geq 1$, $V_n = \mathbb{E}[X_n^2 \mid \mathcal{F}_{n-1}]$. Let $(Z_n,\ n \geq 1)$ be r.v. such that Z_n is $\mathcal{F}_{n-1}$ measurable. We set $M_0 = A_0 = 0$ and, for $n \geq 1$,

$$M_n = \sum_{k=1}^{n} Z_k X_k, \quad A_n = \sum_{k=1}^{n} Z_k^2 V_k, \quad A_\infty = \lim \uparrow A_n.$$

It is assumed that, for all n, $\mathbb{E}[A_n] < +\infty$.

1. Show that M_n is a squared-integrable martingale.
2. Show that $Y_n = M_n^2 - A_n$ is a martingale.
3. Show that if ν is a bounded stopping time, $\mathbb{E}[M_\nu] = 0$ and $\mathbb{E}[M_\nu^2] = \mathbb{E}[A_\nu]$.
4. Let ν be a stopping time such that $\mathbb{E}[A_\nu] < +\infty$.
   (a) Show that $M_{\nu\wedge n}$ converges a.s. and in L^2. Deduce that, on $\{\nu = +\infty\}$, $M_\infty = \lim M_n$ exists a.s.
   (b) Show that $\mathbb{E}[M_\nu] = 0$ and that $\mathbb{E}[M_\nu^2] = \mathbb{E}[A_\nu]$.
   (c) Show that, for all $\rho > 0$, $\mathbb{P}\{\sup_{n\leq\nu} |M_n| \geq \rho\} \leq \frac{1}{\rho^2}\mathbb{E}[A_\nu]$.

Exercise 6.22 (0-1 law). Let $(X_n)_{n\in\mathbb{N}}$ be a process of independent random variables. Define its natural tail σ-field by

$$\mathcal{T} = \bigcap_{n\in\mathbb{N}} \sigma(X_p, p \geq n)\ .$$

1. Compute $\mathbb{E}\left[\mathbb{1}_A \middle| \mathcal{F}_n^X\right]$ for all $n \in \mathbb{N}$ and $A \in \mathcal{T}$.
2. Deduce that any $A \in \mathcal{T}$ has probability 0 or 1.

7 Markov Chains: basic definitions

Exercise 7.1. Prove Lemma 7.2.6.

Exercise 7.2. Let $(\mathsf{X}, \mathcal{X})$ and $(\mathsf{Y}, \mathcal{Y})$ be two measurable spaces. Let P and Q be two Markov and probability kernels on $\mathsf{X} \times \mathcal{X}$ and $\mathsf{X} \times \mathcal{Y}$, respectively. Let $(\mathcal{F}_k)_{k\in\mathbb{N}}$ be a filtration, $(X_k, \mathcal{F}_k)_{k\in\mathbb{N}}$ be a homogeneous Markov chain with kernel P and suppose that $(Y_k)_{k\in\mathbb{N}}$ is $(\mathcal{F}_k)_{k\in\mathbb{N}}$-adapted and such that, for all $k \geq 0$, the conditional distribution of Y_k given $\mathcal{F}_{k-1} \vee \sigma(X_k)$ is $Q(X_k, \cdot)$. Set $Z_k = (X_k, Y_k)$ for all $k \geq 0$.

1. In which case $(X_k)_{k\in\mathbb{N}}$ is i.i.d., depending on the initial distribution μ taken for X_0 ?
2. Show that $(Z_k)_{k\in\mathbb{N}}$ is a homogeneous Markov with kernel R defined by
$$R((x, y), A) = P \otimes Q(x, A) \quad x \in \mathsf{X},\ y \in \mathsf{Y},\ A \in \mathcal{X} \otimes \mathcal{Y} .$$
3. Determine all possible kernels P and Q when $\mathsf{X} = \mathsf{Y} = \{0, 1\}$ using $(a, b), (c, d) \in [0, 1]^2$.

We let p and q denote the corresponding Markov matrices on $\{0, 1\}^2$.

4. Show that for all $k \geq 1$, $\mathbb{E}\left[Y_k \middle| X_{k-1}, Y_{k-1}\right] = pq(X_{k-1}, 1)$.
5. Is $(Y_k, \mathcal{F}_k)_{k\in\mathbb{N}}$ a Markov chain ?

Exercise 7.3. Let $(X_k)_{k\in\mathbb{N}}$ be a homogeneous Markov chain valued in $\mathbb{R}$ with Markov kernel P. Define F on $\mathbb{R}^2$ by
$$F(x, x') = P(x, (-\infty, x']) .$$

1. Write $F(x, x')$ using $\mathbb{P}_{P,x}$ and X_1.
2. Show that, for all $(x, x') \in \mathbb{R}^2$, $F(x, x') = \inf_{q\in\mathbb{Q}} \left(F(x, q)\mathbb{1}_{\{x'\leq q\}} + \mathbb{1}_{\{x'>q\}}\right)$.
3. Deduce that F is a Borel function.

Exercise 7.4. Soit $(Z_n)_{n\geq0}$ une suite i.i.d. de variables aléatoires à valeurs dans $\mathbb{N}$, de loi μ. On considère une suite de variables aléatoires $(X_n)_{n\geq0}$ à valeurs dans $\mathbb{N}$ definie pour $n \geq 1$ par :
$$X_{n+1} = \begin{cases} X_n - 1 & \text{si } X_n \geq 1 \\ Z_n + 1 & \text{si } X_n = 0 , \end{cases}$$
avec X_0 indépendant de $(Z_n)_{n\geq0}$. Montrer que $(X_n)_{n\geq0}$ est une chaîne de Markov homogène sur $\mathbb{N}$ de matrice de transition à déterminer.

Exercise 7.5. Let f and g be two measurable functions defied on $\mathbb{R}$ and respectively valued in $\mathbb{R}$ and $\mathbb{R}_+^*$. Let $(Z_n)_{n\geq0}$ be an i.i.d. sequence of real valued random variables with probability distribution μ. Let us define for all $n \geq 1$,
$$X_n = f(X_{n-1}) + g(X_{n-1})Z_n ,$$
given some real valued random variable X_0 independent of $(Z_n)_{n\geq0}$.

1. Show that $(X_n)_{n\geq0}$ is a Markov chain on $\mathbb{R}$, and determine its transition kernel Q.
2. Show that if μ admits a density with respect to the Lebesgue measure, then Q admits a transition density function and determine this density function.

8 Stationary Markov chains

Exercise 8.1. Let P be a Markov kernel on a general state space $(E, \mathcal{F})$. Assume that there exist a probability measure μ on $(E, \mathcal{F})$ and a mapping $\lambda : \mathcal{F} \to [0, 1]$ such that,

$$\lim_{n\to\infty} \sup_{A\in\mathcal{F}} |\mu P^n(A) - \lambda(A)| = 0 \; . \tag{5}$$

1. Show that λ is a probability measure on $(E, \mathcal{F})$.

We can show that (5) holds if and only if

$$\lim_{n\to\infty} \sup_{|g|\leq 1} \left| \int_E g(x)\mu P^n(\mathrm{d}x) - \int_E g(y)\lambda(\mathrm{d}y) \right| = 0 \; . \tag{6}$$

2. Show that, if (6) holds, then λ is invariant for P.
3. Assume now that all probability measure μ on $(E, \mathcal{F})$ satisfies (5), with the same λ. Show that P admits a unique invariant probability measure.

Exercise 8.2. Let $d \in \mathbb{N}^*$. Let P be a Markov kernel on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Assume that there exists a probability measure μ on $\mathbb{R}^d$, such that the sequence of probability measures $(\mu P^n)_{n\geq 0}$ converges weakly to some probability measure λ_μ. Assume moreover that for all f continuous and bounded, then Pf is also continuous and bounded.

1. Show that λ_μ is invariant for P.
2. What can be said if P admits a unique invariant probability measure.

Assume now that for all probability measure μ on $\mathbb{R}^d$, $(\mu P^n)_{n\geq 0}$ weakly converges to the same probability measure λ.

3. Show that P admits a unique invariant probability measure.

Exercise 8.3. Let $(\xi_n)_{n\in\mathbb{N}}$ denote the canonical process valued in a countable state space E. Consider the sequence of successive return time to $y \in E$ defined by $\sigma_y^{(0)} = \inf\{n > 0 \; : \; \xi_n = y\}$ and for $p \geq 1$,

$$\sigma_y^{(p)} = \begin{cases} \sigma_y^{(p-1)} + \sigma_y^{(0)} \circ \mathrm{S}^{\sigma_y^{(p-1)}} & \text{if } \sigma_y^{(p-1)} < \infty \; , \\ \infty & \text{otherwise} \;\; , \end{cases}$$

where S is the shift operator. Also, denote by $N_y = \sum_{n\geq 0} \mathbb{1}_y(\xi_n)$.

1. Show that we can write, for all $p \geq 1$,

$$\sigma_y^{(p)} = \inf\{n > \sigma_y^{(p-1)} \; : \; \xi_n = y\} \; .$$

2. Show that for all $y \in E$, $\left(\mathcal{F}_{\sigma_y^{(p)}}\right)_{p\geq 0}$ is a filtration.

Let P be a Markov kernel/transition matrix on E.

3. Show that, for all y and all probability ν on E, under $\mathbb{P}_{P,\nu}$, the sequence of random variables $(\sigma_y^{(p)})_{p\in\mathbb{N}}$ is a homogeneous Markov chain with respect to $\left(\mathcal{F}_{\sigma_y^{(p)}}\right)_{p\geq 0}$, valued in $\overline{\mathbb{N}^*} = \mathbb{N}^* \cup \{\infty\}$.
4. Determine the corresponding transition matrix using the distribution μ_y of $\sigma_y^{(0)}$.

We now assume that $y \in E$ with $\mathbb{P}_{P,y}(\sigma_y^{(0)} < \infty) = 1$.

5. Show that for all $p \geq 1$, $\mathbb{P}_{P,y}(\sigma_y^{(p)} < \infty) = 1$. Deduce from this result that $\mathbb{P}_{P,y}(N_y = \infty) = 1$.

6. Show that under $\mathbb{P}_{P,y}$, the sequence of random variables $(\sigma_y^{(p+1)} - \sigma_y^{(p)})_{p \geq 0}$ is i.i.d. , with common distribution μ_y.

Exercise 8.4. We use the same definition and assumptions as in Exercise 8.3.

1. Show for all $x, y \in E$, $x \neq y$,

$$\mathbb{P}_{P,x}\left(N_y = m\right) = \begin{cases} \mathbb{P}_{P,x}\left(\sigma_y^{(0)} < \infty\right) \mathbb{P}_{P,y}\left(\sigma_y^{(0)} = \infty\right) \left(\mathbb{P}_{P,y}\left(\sigma_y^{(0)} < \infty\right)\right)^{m-1} & \text{if } m \geq 1 \\ \mathbb{P}_{P,x}\left(\sigma_y^{(0)} = \infty\right) & \text{otherwise} \end{cases} .$$

2. Show for all $x \in E$,

$$\mathbb{P}_{P,x}\left(N_x = m\right) = \begin{cases} \mathbb{P}_{P,x}\left(\sigma_x^{(0)} = \infty\right) \left(\mathbb{P}_{P,x}\left(\sigma_x^{(0)} < \infty\right)\right)^{m-1} & \text{if } m \geq 1 \\ 0 & \text{otherwise} \end{cases} .$$

Exercise 8.5. Let $\xi = (\xi_n)_{n \in \mathbb{N}}$ denote the canonical process valued in $E = \{0, 1\}$ and the transition matrix on E, K given by

$$K = \begin{pmatrix} 1/4 & 3/4 \\ 1/2 & 1/2 \end{pmatrix}$$

1. Show that K admits a unique invariant probability measure μ.

Let F defined on $E^{\mathbb{N}}$ by

$$F(x) = x_0 + x_1 .$$

2. Compute for all $i \in E$, $\mathbb{E}_{K,i}[F(\xi)]$.

3. Deduce from this result that for all $n \geq 1$ and $i \in E$, $\mathbb{E}_{K,i}[F(\mathrm{S}^n) \mid \mathcal{F}_n^{\xi}]$.

Let $\sigma = \inf\{n \geq 1 \mid \xi_n = 1\}$.

4. Compute $\mathbb{P}_{K,i}(\sigma = n)$ pour $i = 0, 1$ and all $n \geq 1$.

5. Justify that for all $i \in E$, $\mathbb{P}_{K,i}(\sigma < \infty) = 1$.

6. Compute $\mathbb{E}_{K,i}[F(\mathrm{S}^{\sigma}) \mid \mathcal{F}_{\sigma}^{\xi}]$.

Exercise 8.6. Let $\xi = (\xi)_{n \in \mathbb{N}}$ denote the canonical process valued in $E = \{-1, 0, 1\}$ and let the transition matrix K be given by

$$K = \begin{pmatrix} 1/4 & 1/4 & 1/2 \\ 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

Let F defined on $E^{\mathbb{N}}$ by

$$F(x) = x_1 - x_2 .$$

1. Compute for all $i \in E$, $\mathbb{E}_{K,i}[F(\xi)]$

2. Deduce from this result that for all $n \geq 1$ and $i \in E$, $\mathbb{E}_{K,i}[F(\xi \circ \mathrm{S}^n) \mid \mathcal{F}_n^{\xi}]$.

3. Let $\sigma = \inf\{n \geq 1 \mid \xi_n = 1\}$. Compute for all $i \in E$, $\mathbb{E}_{K,i}[\mathbb{1}_{\{\sigma < \infty\}} F(\xi \circ \mathrm{S}^{\sigma}) \mid \mathcal{F}_{\sigma}^{\xi}]$.

Exercise 8.7. On considère ici l'espace d'état fini $E = \{1, 2, 3\}$, et la matrice de transition

$$K = \begin{pmatrix} 0 & 1/3 & 2/3 \\ 1/2 & 1/4 & 1/4 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

Soit F la fonctionnelle définie sur $E^{\mathbb{N}}$ par

$$F(x) = \mathbb{1}_{\{\sum_{j=1}^{3} \mathbb{1}_{\{1\}}(x_j)=1\}} \, .$$

1. Que représente F ?
2. Calculer pour tout $i \in E$, $\mathbb{E}_{K,i}[F(\xi)]$.
3. En déduire pour tout $n \geq 1$ et $i \in E$, $\mathbb{E}_{K,i}[F(\xi \circ \mathrm{S}^n) \mid \mathcal{F}_n^{\xi}]$.
4. Soit $\sigma = \inf\{n \geq 1 \mid \xi_n = 1\}$. Justifier que pour tout $i \in E$, $\mathbb{P}_{K,i}(\sigma < \infty) = 1$, et calculer $\mathbb{E}_{K,i}[F(\xi \circ \mathrm{S}^{\sigma}) \mid \mathcal{F}_{\sigma}^{\xi}]$.

9 Solutions

Solution of Exercise 1.1 1. We have, since 1 has norm 1, $\mathrm{proj}\left(X \middle| \mathrm{Span}\,(1)\right) = \langle X, 1\rangle\, 1 = \mathbb{E}\left[X\right]$.

2. Writing $\mathrm{Span}\,(1, Y) = \mathrm{Span}\,(1, Y - \mathbb{E}\left[Y\right])$, we find

$$\begin{aligned} Z &= \mathrm{proj}\left(X \middle| \mathrm{Span}\,(1, Y)\right) \\ &= \mathbb{E}\left[X\right] + \frac{\langle X, Y - \mathbb{E}\left[Y\right]\rangle}{\|Y - \mathbb{E}\left[Y\right]\|^2}\left(Y - \mathbb{E}\left[Y\right]\right) \\ &= \mathbb{E}\left[X\right] + \frac{\mathrm{Cov}\,(X, Y)}{\mathrm{Var}\,(Y)}\left(Y - \mathbb{E}\left[Y\right]\right)\,, \end{aligned}$$

with the convention $0/0 = 0$.

Solution of Exercise 1.2 1. Taking the integral in (1.10) out of the two sums $\frac{1}{n}\sum_{k=0}^{n-1}$ and $\sum_{k=-n}^{n}$ and, we get the result with

$$J_n(u) = \frac{1}{n}\sum_{k=0}^{n-1}\sum_{j=-n}^{n}\frac{1}{2\pi}\mathrm{e}^{\mathrm{i}ju}\,.$$

2. The first formula is obtain by exchanging the order of summation in the previous display. The second formula is obtained by developing the square.

3. $J_n \geq 0$ is obvious from the second formula. Since $\int_{\mathbb{T}} \mathrm{e}^{\mathrm{i}2\pi kt}\mathrm{d}t = 0$ for $k \neq 0$, we get $\int_{\mathbb{T}} J_n(t)\mathrm{d}t = 1$ from the first formula. Now, let $\epsilon \in (0, \pi]$. For all $t \in \mathbb{T} \setminus [-\epsilon, \epsilon]$, since $|\sin(\pi t)| \geq \sin(\epsilon/2)$, we also get, for all $t \in \mathbb{T} \setminus [-\delta, \delta]$,

$$J_n(t) = \frac{1}{2\pi n}\left|\frac{\sin(nt/2)}{\sin(t/2)}\right|^2 \leq \frac{1}{2\pi n}\frac{1}{\sin^2(\epsilon/2)}\,.$$

Hence the result.

4. Using $\int_{\mathbb{T}} J_n = 1$, we have

$$\left|\int_{\mathbb{T}} f(t) J_n(x - t)\ \mathrm{d}t - f(x)\right| = \left|\int_{\mathbb{T}} \left[f(t) - f(x)\right]\ J_n(x - t)\ \mathrm{d}t\right|\,.$$

We thus only need to prove that

$$\sup_{x\in\mathbb{R}}\left|\int_{\mathbb{T}} \left[f(x - u) - f(x)\right]\ J_n(u)\ \mathrm{d}u\right| \to 0\,.$$

Since f is continuous on $\mathbb{R}$ and periodic, f is uniformly continuous. Let $\delta > 0$. There exists $\epsilon \in (0, \pi)$ such that $|f(x) - f(t)| \leq \delta$ for $|x - t| \leq \epsilon$. We thus obtain that

$$\left|\int_{\mathbb{T}} \left[f(x - u) - f(x)\right]\ J_n(u)\ \mathrm{d}u\right| \leq 2\sup(|f|)\int_{\epsilon<|u|<\pi} |J_n(u)|\ \mathrm{d}u + \delta\int_{-\epsilon}^{\epsilon} |J_n(u)|\ \mathrm{d}u\,.$$

Observe that this bound is independent of x, that the first term tends to 0 as $n \to \infty$ thanks to the previous question and the second term is bounded by δ using

$$\int_{-\epsilon}^{\epsilon} |J_n(u)|\ \mathrm{d}u \leq \int_{\mathbb{T}} J_n = 1\,.$$

Theorem 1.3.1 is proved.

5. Clearly f_n converges to $\mathbb{1}_F$ pointwise. Since $0 \leq f_n \leq 1$, we conclude by dominated convergence.

6. Let F be a closed set such that $A \subset F$ and $\mu(F) \leq \mu(A) + \epsilon/2$. Then

$$\int |\mathbb{1}_A - \mathbb{1}_F| \, \mathrm{d}\mu = \mu(F) - \mu(A) \leq \epsilon/2 \,.$$

   We conclude by taking $g_\epsilon = f_n$ with n large enough and f_n as in the previous question.

7. This follows by observing that the case where $\mathbb{1}_A$ is replaced by a simple measurable function follows by linearity of the integral and since for any measurable function f we can find a sequence of simple functions f_n converging to f pointwise and such that $|f_n|$ is non-decreasing.

8. Let $f \in L^2(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mu)$. We can assume $f \geq 0$ without loss of generality. Let $f_n = f \wedge n$. Hence $f_n \in L^1(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mu)$ and converges to f in the L^2 sense. For any n let $g_n \in L^1(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mu)$ such that the L^1-norm $\|f_n - g_n\|_1 \leq 2^{-n}$. Let $h_n = g_{n+} \wedge n$. Then we have the following bound of the L^2-norm:

$$\|f_n - h_n\|_2 \leq (n\|f_n - h_n\|_1)^{1/2} \leq (n\|f_n - g_n\|_1)^{1/2} \leq \left(n2^{-n}\right)^{1/2} \,.$$

   This shows that and set included in $L^2(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mu)$ that is dense in $L^1(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mu)$ is also dense in $L^2(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mu)$.

9. Let $\epsilon > 0$, $f \in L^2(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mu)$ and f_n a 2π-periodic functions such that $\int |f - f_n|^2 \mathrm{d}\mu \leq \epsilon/2$. By Theorem 1.3.1 we can find for an element of Span $(\phi_n,\, n \in \mathsf{Z})$. g_ϵ such that $\sup |g_\epsilon - f_n|^2 \leq \epsilon/(2\mu(\mathbb{T}))$. It follows that

$$\int |f - g_\epsilon|^2 \mathrm{d}\mu \leq \int |f - f_n|^2 \mathrm{d}\mu + \int |f_n - g_\epsilon|^2 \mathrm{d}\mu \leq \epsilon/2 + \sup |g_\epsilon - f_n|^2 \mu(\mathbb{T}) \leq \epsilon \,.$$

Solution of Exercise 2.1 First note that by definition for all $t \in (0,1)$ and $x \in \mathbb{R}$, we have

$$F(x) \geq t \text{ if and only if } F^{-1}(t) \leq x \; . \tag{7}$$

1. Let $t \in (0,1)$. Taking $x = F^{-1}(t)$ in (7), we have

$$F \circ F^{-1}(t) \geq t \; . \tag{8}$$

   Assume that F is continuous on $\mathbb{R}$. Since $\lim_{y \to -\infty} F(y) = 0$, $\lim_{y \to +\infty} F(y) = 1$, by the intermediate value theorem, for all $t \in]0,1[$, there exists $x \in \mathbb{R}$ such that $F(x) = t$. On one hand, we have by definition $F^{-1}(t) \leq x$ and on the other hand by (8), we have $F \circ F^{-1}(t) \geq t = F(x)$. Since F is nondecreasing, we get $F \circ F^{-1}(t) \geq t = F(x)$.

2. Let $x \in \mathbb{R}$. Taking $t = F(x)$ in (7), we have

$$F^{-1} \circ F(x) \leq x \; .$$

   The proof of the second statement is by contradiction. Assume that $F^{-1} \circ F(x) < x$ and F is strictly increasing on $(x - \epsilon, x)$ for $\epsilon > 0$. Then $F \circ F^{-1} \circ F(x) < F(x)$. But using (8) we get $F(x) < F(x)$.

3. Let X be a random variable associated with F. Assume that $F^{-1} \circ F(X) < X$. Then by the proof of 2, there exist $a, b \in \mathbb{R}$ such that $X \in (a, b]$ such that F is constant on $(a, b]$.

   Now consider the set

$$E = \{u \in (0,1) | \text{ there exist } a, b \in \mathbb{R}, \text{ for all } x \in (a,b], \, F(x) = u\} \; .$$

   We deduce from the previous definitions that $F^{-1} \circ F(X) < X$ if and only if there exists $u \in E$ such that $X \in (a_u, b_u)$ where we defined

$$a_u = \inf\{z \in \mathbb{R} \, : \, F(z) = u\} \; , \text{ and } b_u = \sup\{z \in \mathbb{R} \, : \, F(z) = u\} \; .$$

   Hence to conclude the proof we only need to prove that

   (i) E is at most countable.

   (ii) For all $u \in E$, $\mathbb{P}(X \in (a_u, b_u)) = 0$.

   Let $u \in E$. Then there exist $a, b \in \mathbb{R}$ and $q_u \in (a, b] \cup \mathbb{Q}$ such that $F(q_u) = u$. Then the map $u \mapsto q_u$ is clearly an injection from E to $\mathbb{Q}$ and (i) is proved.

   Now, clearly, by definition of a_u and b_u, for all $a_u < a < b < b_u$, we have $F(b) - F(a) = 0$. Assertion (ii) follows.

4. Assume that F is continuous. Therefore by 1, for all $u \in (0,1)$ there exists $x \in \mathbb{R}$ satisfying $F(x) = u$. Then we get for all $u \in (0,1)$

$$\mathbb{P}\left(F(X) \leq u\right) = \mathbb{P}\left(X \leq x\right) = F(x) = u \; .$$

   It yields that $F(X)$ is a uniform random variable on $[0,1]$.

   Let $n \in \mathbb{N}$, a simple calculation gives $\mathbb{E}\left[F^n(X)\right] = 1/(n+1)$.

5. Let U be a uniform random variable on $[0,1]$. By (7), we have

$$\mathbb{P}\left(F^{-1}(U) \leq x\right) = \mathbb{P}\left(U \leq F(x)\right) = F(x) \; .$$

   Thus, $F^{-1}(U)$ is a random variable associated with the cumulative distribution function F.

6. A simple example is

$$F(s) = (s/2)\mathbb{1}_{[0,2]}(s) \ , \quad G(s) = s\mathbb{1}_{[0,1]}(s) \ ,$$

associated with the uniform distribution on $[0,2]$ and $[0,1]$ respectively.

7. Assume that $G \leq_{sto} F$. Therefore for all $s \in \mathbb{R}$, $F(s) \leq G(s)$. Now let $t \in (0,1)$, we have using 1, $G(F^{-1}(t)) \geq F(F^{-1}(t)) \geq t$. Therefore by definition, we get

$$G^{-1}(t) \leq F^{-1}(t) \ . \tag{9}$$

Let U be a uniform random variable on $[0,1]$ and define $X = F^{-1}(U)$ and $Y = G^{-1}(U)$. By 5, X and Y are associated with F and G respectively. In addition by (9), we get almost surely,

$$Y = G^{-1}(U) \leq F^{-1}(U) = X \ .$$

Now let X and Y two random variables associated with F and G respectively such that $X \geq Y$ almost surely. Therefore for all $x \in \mathbb{R}$, we have

$$G(s) = \mathbb{P}\left(Y \leq s\right) \geq \mathbb{P}\left(Y \leq X \leq s\right) = \mathbb{P}\left(X \leq s\right) = F(s) \ .$$

8. Let Y be a uniform random variable on $[0,1]$ and b be Bernoulli random variable with parameter $5/6$. Then define

$$X = b + (1-b)Y/2 \ .$$

A simple calculation ensures that for all $x \in \mathbb{R}$, $\mathbb{P}\left(X \leq x\right) \leq \mathbb{P}\left(Y \leq x\right)$. However, we get

$$\mathbb{P}\left(X \leq Y\right) = \mathbb{P}\left(b = 0\right) = 1/6 > 0 \ .$$

9. This question is straightforward using 6.

Solution of Exercise 2.2 By Proposition 2.1.5-(h), we have $\mathbb{E}\left[\mathbb{E}\left[Y|\mathcal{B}\right]\right] = \mathbb{E}\left[Y\right]$ and $\mathbb{E}\left[\mathbb{E}\left[Y^2\middle|\mathcal{B}\right]\right] = \mathbb{E}\left[Y^2\right]$. Therefore, we get

$$\begin{aligned}
\sigma^2(Y) &= \mathbb{E}\left[Y^2\right] - \{\mathbb{E}\left[Y\right]\}^2 \\
&= \mathbb{E}\left[\mathbb{E}\left[Y^2\middle|\mathcal{B}\right]\right] - \mathbb{E}\left[\{\mathbb{E}\left[\mathbb{E}\left[Y|\mathcal{B}\right]\right]\}^2\right] + \mathbb{E}\left[\{\mathbb{E}\left[\mathbb{E}\left[Y|\mathcal{B}\right]\right]\}^2\right] - \{\mathbb{E}\left[\mathbb{E}\left[Y|\mathcal{B}\right]\right]\}^2 \\
&= \mathbb{E}\left[\sigma^2(Y|\mathcal{B})\right] + \sigma^2(\mathbb{E}[Y \mid \mathcal{B}]) \ .
\end{aligned}$$

If Y is independent of $\mathcal{B}$, we have $\sigma^2(Y) = \sigma^2(Y|\mathcal{B})$.

Solution of Exercise 2.3 (a) Let $A \in \mathcal{G}$. Then by Lemma 2.1.1 and linearity we have

$$\mathbb{E}\left[\mathbb{1}_A(aX+bY)\right] = a\mathbb{E}\left[\mathbb{1}_A\mathbb{E}\left[X|\mathcal{G}\right]\right] + b\mathbb{E}\left[\mathbb{1}_A\mathbb{E}\left[Y|\mathcal{G}\right]\right] = \mathbb{E}\left[\mathbb{1}_A\{a\mathbb{E}\left[X|\mathcal{G}\right] + b\mathbb{E}\left[Y|\mathcal{G}\right]\}\right] \ .$$

By uniqueness of the conditional expectation Lemma 2.1.1-(i), we get the result.

(b) Assume X is $\mathcal{G}$-measurable. Then the proof follows from a direct application of Lemma 2.1.1-(i).

(c) If $\mathcal{G} = \{\emptyset, \Omega\}$, for all $A \in \mathcal{G}$, $\mathbb{E}\left[\mathbb{1}_A X\right] = \mathbb{E}\left[\mathbb{1}_A\mathbb{E}\left[X\right]\right]$. Then the result is an application of Lemma 2.1.1-(i).

(d) Assume X is independent of $\mathcal{G}$. Then for all $A \in \mathcal{G}$, $\mathbb{E}\left[\mathbb{1}_A X\right] = \mathbb{E}\left[\mathbb{1}_A\right]\mathbb{E}\left[X\right] = \mathbb{E}\left[\mathbb{1}_A\mathbb{E}\left[X\right]\right]$. Then the result is an application of Lemma 2.1.1-(i).

(e) Considering $Y - X$, note that we just need to show the result for $X = 0$ by Proposition 2.1.5-(a). Define the set $A \in \mathcal{G}$ by $A = \{\mathbb{E}\left[Y|\mathcal{G}\right] \leq 0\}$. Then since $Y \geq 0$ and $\mathbb{E}\left[\mathbb{1}_A Y\right] = \mathbb{E}\left[1_A\mathbb{E}\left[Y|\mathcal{G}\right]\right]$, we get $\mathbb{E}\left[1_A\mathbb{E}\left[Y|\mathcal{G}\right]\right] = 0$. Therefore, $1_A\mathbb{E}\left[Y|\mathcal{G}\right] = 0$, $\mathbb{P}$-almost surely and $\mathbb{E}\left[Y|\mathcal{G}\right] = \mathbb{1}_{A^c}\mathbb{E}\left[Y|\mathcal{G}\right] \geq 0$, $\mathbb{P}$-almost surely.

(f) Note that $X \leq X \vee Y$ and $Y \leq X \vee Y$. Then by (e), we have $\mathbb{E}[X|\mathcal{G}] \vee \mathbb{E}[Y|\mathcal{G}] \leq \mathbb{E}[X \vee Y|\mathcal{G}]$. We deduce from the last result taking $Y = 0$ that $\mathbb{E}[X|\mathcal{G}]_+ \leq \mathbb{E}[X_+|\mathcal{G}]$. Moreover, since for all $a \in \mathbb{R}$, $|a| = a_+ + (-a)_+$, we have $|\mathbb{E}[X|\mathcal{G}]| \leq \mathbb{E}[|X||\mathcal{G}]$.

(g) Let $\mathcal{H}$ be a σ-field, $\mathcal{G} \subset \mathcal{H} \subset \mathcal{F}$. Then for all $A \in \mathcal{G}$, since $A \in \mathcal{H}$ we have

$$\mathbb{E}[1_A \mathbb{E}[\mathbb{E}[X|\mathcal{H}]|\mathcal{G}]] = \mathbb{E}[1_A \mathbb{E}[X|\mathcal{H}]] = \mathbb{E}[1_A X] \ .$$

Then the proof follows from a direct application of Lemma 2.1.1-(i).

(h) It suffices to take $A = \Omega$ in (2.1).

(i) Assume that X is $\mathcal{G}$-measurable. Let $A \in \mathcal{G}$. Then we have by Lemma 2.1.4, since $\mathbb{1}_A X$ is $\mathcal{G}$-measurable,

$$\mathbb{E}[\mathbb{1}_A XY] = \mathbb{E}[\mathbb{1}_A X \mathbb{E}[Y|\mathcal{G}]] \ .$$

Then the proof follows from a direct application of Lemma 2.1.1-(i).

(j) Let $\mathcal{H}$ be a σ-field, $\mathcal{H} \subset \mathcal{F}$ and assume that $\sigma(X) \vee \mathcal{H}$ is independent of $\mathcal{G}$. We want to show that $\mathbb{E}[X|\mathcal{H} \vee \mathcal{G}] = \mathbb{E}[X|\mathcal{H}]$. By Lemma 2.1.1-(i), we just need to prove that for all $A \in \mathcal{H} \vee \mathcal{G}$, we have

$$\mathbb{E}[\mathbb{1}_A X] = \mathbb{E}[\mathbb{1}_A \mathbb{E}[X|\mathcal{H}]] \ .$$

We first consider A of the form $B \cap C$, with $B \in \mathcal{H}$ and $C \in \mathcal{G}$. Indeed, using the assumption, for such measurable set, we get since $1_B X$ is $\sigma(X) \vee \mathcal{H}$-measurable, $\mathbb{1}_B \mathbb{E}[X|\mathcal{H}]$ is $\mathcal{H}$-measurable,

$$\mathbb{E}[\mathbb{1}_A X] = \mathbb{E}[\mathbb{1}_C \mathbb{1}_B X] = \mathbb{E}[\mathbb{1}_C] \mathbb{E}[\mathbb{1}_B X] = \mathbb{E}[\mathbb{1}_C] \mathbb{E}[\mathbb{1}_B \mathbb{E}[X|\mathcal{H}]] = \mathbb{E}[\mathbb{1}_A \mathbb{E}[X|\mathcal{H}]] \ . \tag{10}$$

Now consider $\mathcal{E} \subset \mathcal{F}$ and $\mathcal{C} \subset \mathcal{F}$ defined by

$$\mathcal{E} = \{A \in \mathcal{F} : \mathbb{E}[\mathbb{1}_A X] = \mathbb{E}[\mathbb{1}_A \mathbb{E}[X|\mathcal{H}]]\} \ , \quad \mathcal{C} = \{B \cap C : B \in \mathcal{H} \text{ and } C \in \mathcal{G}\} \ .$$

By (10) we get that $\mathcal{C} \subset \mathcal{E}$. It is straightforward to check that $\mathcal{C}$ is stable by finite intersection, contains Ω and $\sigma(\mathcal{C}) = \mathcal{H} \vee \mathcal{G}$. Therefore it is a π-system. Then we just need to show that $\mathcal{E}$ is a λ-system since by the π-λ theorem, it will imply that $\sigma(\mathcal{C}) = \mathcal{H} \vee \mathcal{G} \subset \mathcal{E}$.

Let $A \in \mathcal{E}$. Using that $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|\mathcal{H}]]$, we get that $A^c \in \mathcal{E}$. Consider now a sequence $(A_n)_{n\in\mathbb{N}} \in \mathcal{E}^{\mathbb{N}}$ such that for all $n < p$, $A_n \cap A_p = \emptyset$. Then for all $N \in \mathbb{N}$,

$$\mathbb{E}\left[\mathbb{1}_{\cup_{k=0}^N A_k} X\right] = \sum_{n=1}^{N} \mathbb{E}[\mathbb{1}_{A_k} X] = \sum_{n=1}^{N} \mathbb{E}[\mathbb{1}_{A_k} \mathbb{E}[X|\mathcal{H}]] = \mathbb{E}\left[\mathbb{1}_{\cup_{k=0}^N A_k} \mathbb{E}[X|\mathcal{H}]\right] \ .$$

Setting $A = \cup_{k=0}^N A_k$ and using the dominated convergence theorem, we get

$$\mathbb{E}[\mathbb{1}_A X] = \mathbb{E}[\mathbb{1}_A \mathbb{E}[X|\mathcal{H}]] \ .$$

Therefore $A \in \mathcal{E}$ and $\mathcal{E}$ is a λ-system.

(k) We first show the result when F is the identity. Namely, for all $A \in \mathcal{B}(\mathbb{R}^{p+q})$ and all $\omega \in \Omega$, we prove

$$\mathbb{P}^{(Y,Z)|\mathcal{G}}(\omega, A) = \mathbb{P}((Y(\omega), Z) \in A) = \int_{\Omega} \mathbb{1}_A(Y(\omega), Z)\mathbb{P}(\mathrm{d}\tilde{\omega}) \ . \tag{11}$$

Consider first A of the form $A = B \times C$, with $B \in \mathcal{B}(\mathbb{R}^p)$ and $C \in \mathcal{B}(\mathbb{R}^q)$. Then for all $D \in \mathcal{G}$, we have since $\mathbb{1}_D Y$ is $\mathcal{G}$ measurable and Z is independent of $\mathcal{G}$

$$\begin{aligned}\mathbb{E}[\mathbb{1}_D \mathbb{1}_A(Y,Z)] &= \mathbb{E}[\mathbb{1}_D \mathbb{1}_B(Y)\mathbb{1}_C(Z)] = \mathbb{E}[\mathbb{1}_D \mathbb{1}_B(Y)] \mathbb{E}[\mathbb{1}_C(Z)] \\ &= \mathbb{E}[\mathbb{1}_D \mathbb{1}_B(Y)] \mathbb{P}(Z \in C) = \mathbb{E}[\mathbb{1}_D \mathbb{1}_B(Y)\mathbb{P}(Z \in C)] \ .\end{aligned}$$

Therefore, we have almost surely

$$\mathbb{P}^{(Y,Z)|\mathcal{G}}(\omega, A) = \mathbb{E}\left[\,\mathbb{1}_A(Y,Z)|\,\mathcal{G}\right](\omega) = \mathbb{1}_B(Y(\omega))\mathbb{P}\left(Z \in C\right) = \int_\Omega \mathbb{1}_A(Y(\omega), Z(\tilde{\omega}))\mathbb{P}(\mathrm{d}\tilde{\omega}) \ . \quad (12)$$

Consider now the two set $\mathcal{E}$ and $\mathcal{C}$ contained in $\mathcal{F} = \mathcal{B}(\mathbb{R}^{p+q})$ defined by

$$\mathcal{E} = \left\{ A \in \mathcal{F} \,:\, \mathbb{P}^{(Y,Z)|\mathcal{G}}(\omega, A) = \int_\Omega \mathbb{1}_A(Y(\omega), Z(\tilde{\omega}))\mathbb{P}(\mathrm{d}\tilde{\omega}) \ ,\omega\text{-almost surely}\right\}$$

$$\mathcal{C} = \{B \cap C \,:\, B \in \mathcal{B}(\mathbb{R}^p) \text{ and } C \in \mathcal{B}(\mathbb{R}^q)\} \ .$$

By (12), we get that $\mathcal{C} \subset \mathcal{E}$. It is straightforward to check that $\mathcal{C}$ is stable by finite intersection, contains Ω and $\sigma(\mathcal{C}) = \mathcal{H} \vee \mathcal{G}$. Therefore it is a π-system. Then we just need to show that $\mathcal{E}$ is a λ-system since by the π-λ theorem, it will imply that $\sigma(\mathcal{C}) = \mathcal{H} \vee \mathcal{G} \subset \mathcal{E}$.

Let $A \in \mathcal{E}$, A it is clear by definition that $A^c \in \mathcal{E}$. Consider now a sequence $(A_n)_{n\in\mathbb{N}} \in \mathcal{C}^{\mathbb{N}}$ such that for all $n < p$, $A_n \cap A_p = \emptyset$. Then by defition for all $N \in \mathbb{N}$, we have almost surely

$$\mathbb{P}^{(Y,Z)|\mathcal{G}}\left(\omega, \bigcup_{k=0}^N A_k\right) = \int_\Omega \mathbb{1}_{\bigcup_{k=0}^N A_k}(Y(\omega), Z(\tilde{\omega}))\mathbb{P}(\mathrm{d}\tilde{\omega}) \ . \quad (13)$$

Therefore almost surely for all $N \in \mathbb{N}$ (note the difference here), we get that (13) holds. Setting $A = \cup_{k=0}^N A_k$ and using the monotone convergence theorem, we get

$$\mathbb{P}^{(Y,Z)|\mathcal{G}}(\omega, A) = \int_\Omega \mathbb{1}_A(Y(\omega), Z(\tilde{\omega}))\mathbb{P}(\mathrm{d}\tilde{\omega}) \ . \quad (14)$$

Then $A \in \mathcal{E}$ and $\mathcal{E}$ is a λ-system. So we have shown (11).

Let now $F : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^m$ be a Borel function and $X = F(Y,Z)$. Then for all $A \in \mathcal{B}(\mathbb{R}^m)$ and $B \in \mathcal{G}$, we have

$$\begin{aligned}\mathbb{E}\left[\mathbb{1}_B \mathbb{1}_A(X)\right] = \mathbb{E}\left[\mathbb{1}_B \mathbb{1}_{F^{-1}(A)}(Y,Z)\right] &= \int_\Omega \mathbb{1}_B(\omega) \int_\Omega \mathbb{1}_{F^{-1}(A)}(Y(\omega), Z(\tilde{\omega}))\mathbb{P}(\mathrm{d}\tilde{\omega})\mathbb{P}(\mathrm{d}\omega) \\ &= \int_\Omega \mathbb{1}_B(\omega) \int_\Omega \mathbb{1}_A(F(Y(\omega), Z(\tilde{\omega})))\mathbb{P}(\mathrm{d}\tilde{\omega})\mathbb{P}(\mathrm{d}\omega) \ .\end{aligned}$$

Therefore, we get the expected result for all $A \in \mathcal{B}(\mathbb{R}^m)$ almost surely:

$$\mathbb{P}^{X|\mathcal{G}}(\omega, A) = \mathbb{P}\left(F(Y(\omega), Z) \in A\right) \ .$$

Solution of Exercise 2.4 The Hilbert space H of all $\mathbb{R}^p$-valued L^2 random variables is endowed with the scalar product

$$\langle U, V\rangle = \mathbb{E}\left[U^T V\right]$$

In this context, $\mathrm{Span}\,(1, \mathbf{Y})$ is seen as the linear space in H obtained by a linear transformation of the random variables 1 and $\mathbf{Y}$, that is, we have

$$\begin{aligned}\mathrm{Span}\,(1, \mathbf{Y}) &= \{a + A\mathbf{Y} \ :\ a \in \mathbb{R}^p,\ A \in \mathbb{R}^p \times \mathbb{R}^q\} \\ &= \{b + A(\mathbf{Y} - \mathbb{E}\left[\mathbf{Y}\right]) \ :\ b \in \mathbb{R}^p,\ A \in \mathbb{R}^p \times \mathbb{R}^q\} \ ,\end{aligned}$$

where we set $b = a - A\mathbb{E}\left[\mathbf{Y}\right]$.

1. Note that $\mathrm{Span}\,(1, \mathbf{Y})$ is a finite dimensional linear subspace of H, hence is closed. So by Theorem 1.4.1, $\widehat{\mathbf{X}} := \mathrm{proj}\,(\mathbf{X}|\,\mathrm{Span}\,(1, \mathbf{Y}))$ is given by

$$\widehat{\mathbf{X}} = \hat{b} + \hat{A}(\mathbf{Y} - \mathbb{E}\left[\mathbf{Y}\right]) \ ,$$

with $\hat{b} \in \mathbb{R}^p,\ \hat{A} \in \mathbb{R}^p \times \mathbb{R}^q$ such that

$$\left\langle \mathbf{X} - (\hat{b} + \hat{A}(\mathbf{Y} - \mathbb{E}[\mathbf{Y}]), b + A(\mathbf{Y} - \mathbb{E}[\mathbf{Y}]) \right\rangle = 0 \text{ for all } b \in \mathbb{R}^p,\ A \in \mathbb{R}^p \times \mathbb{R}^q ,$$

which is equivalent to the two conditions

$$\left\langle \mathbf{X} - \hat{b}, b \right\rangle = 0,\ \left\langle \mathbf{X} - \hat{A}(\mathbf{Y} - \mathbb{E}[\mathbf{Y}]), A(\mathbf{Y} - \mathbb{E}[\mathbf{Y}]) \right\rangle = 0 \text{ for all } b \in \mathbb{R}^p,\ A \in \mathbb{R}^p \times \mathbb{R}^q .$$

This clearly yields $\hat{b} = \mathbb{E}[\mathbf{X}]$ for the first condition and since

$$\mathbb{E}\left[(\mathbf{X} - \hat{A}(\mathbf{Y} - \mathbb{E}[\mathbf{Y}])^T A(\mathbf{Y} - \mathbb{E}[\mathbf{Y}])\right] = \mathrm{Trace}\left(A\mathbb{E}\left[(\mathbf{Y} - \mathbb{E}[\mathbf{Y}])(\mathbf{X} - \hat{A}(\mathbf{Y} - \mathbb{E}[\mathbf{Y}])^T\right]\right) ,$$

the second condition gives

$$\mathrm{Cov}\left(\mathbf{Y}, \mathbf{X} - \hat{A}\mathbf{Y}\right) = \mathbb{E}\left[(\mathbf{Y} - \mathbb{E}[\mathbf{Y}])(\mathbf{X} - \hat{A}(\mathbf{Y} - \mathbb{E}[\mathbf{Y}])^T\right] = 0 .$$

which yields $\hat{A} = \mathrm{Cov}(\mathbf{Y}, \mathbf{X})\,\mathrm{Cov}(\mathbf{Y})^{-1}$. Hence, as a result,

$$\widehat{\mathbf{X}} = \mathbb{E}[\mathbf{X}] + \mathrm{Cov}(\mathbf{X}, \mathbf{Y})\,\mathrm{Cov}(\mathbf{Y})^{-1}\ (\mathbf{Y} - \mathbb{E}[\mathbf{Y}]) .$$

Now observe that

$$\begin{aligned}
\mathrm{Cov}(\mathbf{X} - \widehat{\mathbf{X}}) &= \mathrm{Cov}(\mathbf{X} - \mathrm{Cov}(\mathbf{X}, \mathbf{Y})\,\mathrm{Cov}(\mathbf{Y})^{-1}\mathbf{Y}) \\
&= \mathrm{Cov}\left(\mathbf{X} - \mathrm{Cov}(\mathbf{X}, \mathbf{Y})\,\mathrm{Cov}(\mathbf{Y})^{-1}\mathbf{Y}, \mathbf{X}\right) \\
&= \mathrm{Cov}(\mathbf{X}) - \mathrm{Cov}(\mathbf{X}, \mathbf{Y})\,\mathrm{Cov}(\mathbf{Y})^{-1}\mathrm{Cov}(\mathbf{Y}, \mathbf{X}) .
\end{aligned}$$

Hence we have (iii) of Proposition 2.1.11.

2. Let us write

$$\mathbf{X} = \widehat{\mathbf{X}} + \left(\mathbf{X} - \widehat{\mathbf{X}}\right) ,$$

and observe that since $(\mathbf{X}, \mathbf{Y})$ is Gaussian, so is $(\mathbf{Y}, \left(\mathbf{X} - \widehat{\mathbf{X}}\right))$, which is obtained by a linear transform of it. Moreover since $\mathrm{Cov}\left(\mathbf{Y}, \mathbf{X} - \widehat{\mathbf{X}}\right) = 0$ by definition of $\widehat{\mathbf{X}}$, they are independent. In the above decomposition, $\widehat{\mathbf{X}}$ is $\sigma(\mathbf{Y})$-measurable and $(\mathbf{X} - \widehat{\mathbf{X}})$ is independent of $\sigma(\mathbf{Y})$. This immediately gives

$$\mathbb{E}[\mathbf{X}|\,\mathbf{Y}] = \widehat{\mathbf{X}} = \mathrm{proj}\left(\mathbf{X}|\,\mathrm{Span}\,(1, \mathbf{Y})\right) ,$$

that is, (i) of Proposition 2.1.11. Moreover, we can apply Property (b). This gives that, for all $\omega \in \Omega$ and all $A \in \mathcal{B}(\mathbb{R}^p)$,

$$\mathbb{P}^{\mathbf{X}|\mathbf{Y}}(\mathbf{Y}(\omega), A) = \int \mathbb{1}_A(\widehat{\mathbf{X}}(\omega) + \mathbf{X}(\omega') - \widehat{\mathbf{X}}(\omega'))\,\mathbb{P}(\mathrm{d}\omega') .$$

But since $\mathbf{X} - \widehat{\mathbf{X}}$ is a linear transform of $(\mathbf{X}, \mathbf{Y})$, it is a Gaussian vector, moreover it has mean 0, hence, for all $\omega \in \Omega$, the random vector $\omega' \mapsto \widehat{\mathbf{X}}(\omega) + \mathbf{X}(\omega') - \widehat{\mathbf{X}}(\omega')$ is $\mathcal{N}(\widehat{\mathbf{X}}, \mathrm{Cov}(\mathbf{X} - \widehat{\mathbf{X}}))$. Hence we obtain (ii) of Proposition 2.1.11.

Solution of Exercise 2.5 1. For all $n \in \mathbb{N}^*$ set $A_n = \{|\phi| \geq 1/n\}$. Then for all $n \in \mathbb{N}^*$,

$$0 = \int_{A_n} \psi\ \mathrm{d}\lambda\ ,\ \text{and}\ \left|\int_{A_n} \psi\ \mathrm{d}\lambda\right| \geq (2n)^{-1}\lambda(A_n) .$$

Therefore for all $n \in \mathbb{N}^*$, $\lambda(A_n) = 0$, and the monotone convergence theorem concludes the proof.

2. Let $A \in \mathcal{F}$, $\lambda(A) = 0$. Then by the definition of the Lebesgue integral, $\mu(A)$ is equal to 0.

3. Applying the formula with $A = \{\phi = 0\}$ we immediately get $\mu(\{\phi = 0\}) = 0$.

4. For example consider λ is the Lebesgue measure and $\phi = \mathbb{1}_{[0,1]}$.

5. Since λ and μ are σ-finite, there exists an increasing sequence $(A_n)_{n\in\mathbb{N}} \in \mathcal{F}^{\mathbb{N}}$ such that $\cup_{n\in\mathbb{N}} A_n = \Omega$ and for all $n \in \mathbb{N}$, $\mu(A_n) + \lambda(A_n) < \infty$.

   The proof is by contradiction. Assume that $\lambda(\{\phi = \infty\}) \neq 0$. Therefore by the monotone convergence theorem, there exists $N \in \mathbb{N}$ such that $\lambda(\{\phi = \infty\} \cap A_N) \neq 0$. Then for all $M > 0$, we have

$$\mu(A_N) \geq M\lambda(\{\phi = \infty\} \cap A_N) \ ,$$

   and taking the limit as $M \to \infty$, we get $\mu(A_N) = \infty$. We have a contradiction.

6. If another function ϕ' satisfies for all $A \in \mathcal{F}$, $\mu(A) = \int_A \phi' \, \mathrm{d}\lambda$ then

$$\int_A (\phi - \phi') \, \mathrm{d}\lambda = 0$$

   whenever $\mu(A) < \infty$. Hence by Question 1, we have that $\lambda(A_n \cap \{\phi \neq \phi'\}) = 0$ for all $n \geq 1$. Since $\mathrm{cum}_n A_n = \Omega$, we get the result.

Solution of Exercise 2.7 1. Let X and Y denote the countable sets in which $\mathbf{X}$ and $\mathbf{Y}$ take their values, respectively. Then $(\mathbf{X}, \mathbf{Y})$ takes its values in $\mathsf{X} \times \mathsf{Y}$ which is also countable. Moreover, letting μ_E denote the counting measure on E, that is

$$\mu_E = \sum_{x\in E} \delta_x$$

where δ_x is the Dirac point mass at x, we note that

$$\mu_{\mathsf{X}\times\mathsf{Y}} = \mu_{\mathsf{X}} \otimes \mu_{\mathsf{Y}} \ .$$

We conclude that $(\mathbf{X}, \mathbf{Y})$ admits a density

$$f : (x, y) \mapsto \mathbb{P}(\mathbf{X} = x, \mathbf{Y} = y)$$

with respect to the product measure $\mu_{\mathsf{X}} \otimes \mu_{\mathsf{Y}}$. Hence the definition of the conditional density of $\mathbf{X}$ given $\mathbf{Y}$ applies. We observe that in this case

$$f(x|y) = \frac{f(x, y)}{\sum_{x'} f(x', y)} = \frac{\mathbb{P}(\mathbf{X} = x, \mathbf{Y} = y)}{\mathbb{P}(\mathbf{Y} = y)} = \mathbb{P}(\mathbf{X} = x|\mathbf{Y} = y) \ ,$$

which is the usual conditional probability of event $\{\mathbf{X} = x\}$ given the event $\{\mathbf{Y} = y\}$.

2. Denote by λ_p the Lebesgue measure on $\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$. Let p and q be the dimensions of the random vectors $\mathbf{X}$ and $\mathbf{Y}$, respectively. Then $(\mathbf{X}, \mathbf{Y})$ admits a density $f : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}_+$ with respect to $\lambda_{p+q} = \lambda_p \otimes \lambda_q$. Hence the definition of the conditional density of $\mathbf{X}$ given $\mathbf{Y}$ applies.

3. Let X and Y denote the sets in which $\mathbf{X}$ and $\mathbf{Y}$ take their values, respectively. Here only Y is assumed to be countable, hence endowed with the σ-field $\mathcal{Y} = \mathcal{P}(\mathsf{Y})$, and we let X be endowed with a σ-field $\mathcal{X}$. Let A be a set of $\mathcal{X} \otimes \mathcal{Y}$. Then we have

$$\mathbb{P}((\mathbf{X}, \mathbf{Y}) \in A) = \sum_{y\in\mathsf{Y}} \mathbb{P}((\mathbf{X}, y) \in A, \mathbf{Y} = y) \leq \sum_{y\in\mathsf{Y}} \mathbb{P}((\mathbf{X}, y) \in A) \ .$$

Let $\xi' = \mu_{\mathsf{Y}}$ and $\xi = \mathbb{P}^{\mathbf{X}}$. Then we have

$$\xi \otimes \xi'(A) = \sum_{y \in \mathsf{Y}} \int \mathbb{1}_A(x, y)\ P^{\mathbf{X}}(\mathrm{d}x) = \sum_{y \in \mathsf{Y}} \mathbb{P}((\mathbf{X}, y) \in A)\ .$$

Hence we obtain from the last two displays that $\mathbb{P}^{(\mathbf{X},\mathbf{Y})} \ll \xi \otimes \xi'$. Hence, by Radon-Nikodym's thorem, $(\mathbf{X}, \mathbf{Y})$ admits a density $f : \mathsf{X} \times \mathsf{Y} \to \mathbb{R}_+$ with respect to $\xi \otimes \xi'$ and the definition of the conditional density of $\mathbf{X}$ given $\mathbf{Y}$ applies. It is not so easy to determine f. But it is interesting. To this end, we first observe that f is characterized by the equality

$$\mathbb{P}(\mathbf{X} \in B, \mathbf{Y} = y) = \int_{B \times \{y\}} f\ \mathrm{d}\xi \otimes \xi' \quad \text{for all } B \in \mathcal{X}, y \in \mathsf{Y},$$

by the characterization theorem of probability measures on π-systems. Developing the right-hand side with the definition of ξ and ξ', we get that

$$\mathbb{P}(\mathbf{X} \in B, \mathbf{Y} = y) = \mathbb{E}[\mathbb{1}_B(X) f(X, y)] \quad \text{for all } B \in \mathcal{X}, y \in \mathsf{Y},$$

We can interpret this by writing that, for all $y \in \mathsf{Y}$,

$$f(\mathbf{X}, y) = \mathbb{P}(\mathbf{Y} = y | \mathbf{X}) \quad \mathbb{P}\text{-a.s.}$$

Moreover since the density of $\mathbf{Y}$ with respect to $\xi' = \mu_{\mathsf{Y}}$ is $y \mapsto \mathbb{P}(\mathbf{Y} = y)$, the conditional density of $\mathbf{X}$ given $\mathbf{Y}$ reads

$$f(x|y) = \frac{f(x, y)}{\mathbb{P}(\mathbf{Y} = y)}\ .$$

Remark 9.1. *The choice of ξ above corresponds to ξ of the form*

$$\xi(B) = \sum_y w_y \mathbb{P}(\mathbf{X} \in B, \mathbf{Y} = y)$$

with $w_y = 1$. We note that, for any $w_y \geq 0$ satisfying

$$\mathbb{P}(\mathbf{Y} = y) > 0 \Rightarrow w_y > 0\ ,$$

we have

$$\mathbb{P}((\mathbf{X}, y) \in A) w_y = 0 \Rightarrow \mathbb{P}((\mathbf{X}, y) \in A, \mathbf{Y} = y) = 0\ .$$

Hence we have a wide possibility of choices for ξ, but this variety of choices is not so helpful. More interestingly, we observe that another dominating measure of $\mathbb{P}^{(\mathbf{X},\mathbf{Y})}$ is simply $\mathbb{P}^{\mathbf{X}} \otimes \mathbb{P}^{\mathbf{Y}}$. In this very special case, the joint density, the conditional density of X given Y and the conditional density of Y given X all coincide!

4. Once a conditional density $f(x|y)$ is defined (from a density f with respect to the product measure $\xi \otimes \xi'$), we can apply the formula

$$\mathbb{E}[\mathbf{X}|\mathbf{Y}] = \int x\, f(x|\mathbf{Y})\ \mathrm{d}\xi(x)\ .$$

In the first case, we get

$$\mathbb{E}[\mathbf{X}|\mathbf{Y}] = g(\mathbf{Y}) \quad \text{with} \quad g(y) = \sum_{x \in \mathsf{X}} x \mathbb{P}(\mathbf{X} = x | \mathbf{Y} = y)\ .$$

In the second case case, we can directly write

$$\mathbb{E}[\mathbf{X}|\mathbf{Y}] = \int x\, f(x|\mathbf{Y})\ \mathrm{d}x\ ,$$

with

$$f(x|y) = \frac{f(x,y)}{f^{\mathbf{Y}}(y)} = \frac{f(x,y)}{\int f(x',y)\,\mathrm{d}x'} .$$

In the third case, we get that

$$\mathbb{E}[\mathbf{X}|\mathbf{Y}] = \int x\, f(x|\mathbf{Y})\;\mathrm{d}\xi(x) = g(\mathbf{Y}) ,$$

with

$$\begin{aligned}
g(y) &= \int x\, f(x|y)\;\mathrm{d}\xi(x) \\
&= \mathbb{E}[\mathbf{X}\, f(\mathbf{X}|y)] \\
&= \mathbb{E}\left[\mathbf{X}\, \frac{\mathbb{P}(\mathbf{Y}=y|\mathbf{X})}{\mathbb{P}(\mathbf{Y}=y)}\right] \\
&= \frac{1}{\mathbb{P}(\mathbf{Y}=y)}\mathbb{E}\left[\mathbb{E}\left[\mathbf{X}\,\mathbb{1}_{\{\mathbf{Y}=y\}}\middle|\,\mathbf{X}\right]\right] \\
&= \frac{\mathbb{E}\left[\mathbf{X}\,\mathbb{1}_{\{\mathbf{Y}=y\}}\right]}{\mathbb{P}(\mathbf{Y}=y)} .
\end{aligned}$$

5. Let $\begin{bmatrix}\mathbf{X}^T & \mathbf{Y}^T\end{bmatrix}^T \sim \mathcal{N}(\begin{bmatrix}m & m'\end{bmatrix}^T, \Gamma)$ with Γ invertible. Then $(\mathbf{X},\mathbf{Y})$ admits a density proposrtional to

$$f : (x,y) \mapsto \exp\left(-\frac{1}{2}\begin{bmatrix}(x-m)\\(y-m')\end{bmatrix}^T \Gamma^{-1}\begin{bmatrix}(x-m)\\(y-m)\end{bmatrix}\right) .$$

Let us write the positive definite symmetric matrix Γ^{-1} as the block matrix

$$\Gamma^{-1} = \begin{bmatrix}A & B\\ B^T & C\end{bmatrix} .$$

Then we have, for a well chosen positive function h,

$$f(x|y) = h(y)\exp\left(-\frac{1}{2}\left((x-m)^T A(x-m) + 2(x-m)^T B(y-m')\right)\right) .$$

Changing h adequately into h_1, we get

$$f(x|y) = h_1(y)\exp\left(-\frac{1}{2}\left((x-m+A^{-1}B(y-m'))^T A(x-m+A^{-1}B(y-m'))\right)\right) .$$

At a fixed y, as a function of x, we recognize $x \mapsto f(x|y)$ as the density of a Gaussian vector with mean $m-A^{-1}B(y-m')$ and covariance matrix A^{-1}. Hence we find that the conditional distribution of $\mathbf{X}$ given $\mathbf{Y}$ is a Gaussian one with mean $m - A^{-1}B(\mathbf{Y}-m')$ and covariance A^{-1}. In particular, the conditional expectation of $\mathbf{X}$ given $\mathbf{Y}$ is a linear function of 1 and $\mathbf{Y}$, and since $\mathbf{X}$ is L^2, it is the L^2 projection on $\mathrm{Span}\,(1,\mathbf{Y})$.

Solution of Exercise 2.8 1. First, note since f is symmetric, X and $-X$ have the same law. Let $A, B \in \mathcal{B}(\mathbb{R}^d)$, we have

$$\mathbb{E}\left[\mathbb{1}_{\{X\in A\}}\mathbb{1}_{\{|X|\in B\}}\right] = (1/2)\mathbb{E}\left[(\mathbb{1}_{\{X\in A\}} + \mathbb{1}_{\{-X\in A\}})\mathbb{1}_{\{|X|\in B\}}\right] .$$

Since

$$\mathbb{1}_{\{X\in A\}} + \mathbb{1}_{\{-X\in A\}} = \mathbb{1}_{\{|X|\in A\}} + \mathbb{1}_{\{-|X|\in A\}} ,$$

we get

$$\mathbb{E}\left[\mathbb{1}_{\{X\in A\}}\mathbb{1}_{\{|X|\in B\}}\right] = (1/2)\mathbb{E}\left[(\mathbb{1}_{\{|X|\in A\}} + \mathbb{1}_{\{-|X|\in A\}})\mathbb{1}_{\{|X|\in B\}}\right] .$$

Thus, the conditional law of X given $|X|$ is given by

$$\mathbb{P}^{X\,|\,|X|}(|X|\,,A) = (1/2)(\delta_{|X|}(A) + \delta_{-|X|}(A)) .$$

2. Let Y be random variable on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ as X and taking value in Y. Denote by $\mathbb{P}_Y$ the law of Y on Y. Let $A \in \mathcal{B}(\mathbb{R}^d)$. Let $A \in \mathcal{B}(\mathbb{R})$. By definition and since X has a density with respect to the Lebesgue measure, we have

$$0 = \mathbb{P}(X \in A) = \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\{X \in A\}}\middle| Y\right]\right] = \int_\Omega \mathbb{P}^{X|Y}(Y(\omega), A)\mathrm{d}\mathbb{P}(\omega)$$
$$= \int_{\mathrm{Y}} \mathbb{P}^{X|Y}(y, A)\mathrm{d}\mathbb{P}_Y(y) .$$

It implies that y-$\mathbb{P}_Y$ almost surely, $\mathbb{P}^{X|Y}(y, A) = 0$.

3. It is wrong that $\mathbb{P}^{X|Y}(y, A)$ admits a density with respect to the Lebesgue measure $\mathbb{P}_Y$ almost surely. A counterexample is given by the first question.

Solution of Exercise 2.9 1. The density with respect to the Lebesgue measure of the uniform distribution on the triangle $(0,0),(1,0),(0,1)$ in $\mathbb{R}^2$ is

$$(x, y) \mapsto 2\mathbb{1}_{\{x\geq 0\}}\mathbb{1}_{\{y\geq 0\}}\mathbb{1}_{\{x+y\leq 1\}} .$$

Therefore the conditional density of X given Y is for $y \in [0, 1)$

$$(x, y) \mapsto 2(1-y)^{-1}\mathbb{1}_{\{x\geq 0\}}\mathbb{1}_{\{x\leq 1-y\}} ,$$

and its conditional expectation is given by

$$\mathbb{E}[X|Y] = 2(1-Y)^{-1}\int_{\mathbb{R}} x\mathbb{1}_{\{x\geq 0\}}\mathbb{1}_{\{x\leq 1-Y\}}\mathrm{d}x = 1 - Y .$$

2. The density with respect to the Lebesgue measure of the uniform distribution on the square $(0,0),(1,0),(1,1),(0,1)$ in $\mathbb{R}^2$ is

$$(x, y) \mapsto \mathbb{1}_{\{0\leq x\leq 1\}}\mathbb{1}_{\{0\leq y\leq 1\}} .$$

Therefore the conditional density of X given Y is for $y \in [0, 1)$

$$(x, y) \mapsto \mathbb{1}_{\{0\leq x\leq 1\}} ,$$

and its conditional expectation is given by

$$\mathbb{E}[X|Y] = \int_{\mathbb{R}} x\mathbb{1}_{\{0\leq x\leq 1\}}\mathrm{d}x = 1/2 .$$

These results are not surprising since if U and V are uniformly distributed on $[0, 1]$, then (U, V) is uniformly distributed on the square.

Solution of Exercise 2.10 Let us find first the law of (X, Y). Let $A \subset \mathbb{N}$ and $B \in \mathcal{B}(\mathbb{R}_+)$. Then using that Y is $\sigma(Y)$-measurable and the conditional distribution of X given Y is a Poisson distribution with parameter Y we get by Fubini's theorem

$$\begin{aligned}\mathbb{P}(X \in A, Y \in B) &= \mathbb{E}[\mathbb{1}_A(X)\mathbb{1}_B(Y)] = \mathbb{E}[\mathbb{1}_B(Y)\mathbb{E}[\mathbb{1}_A(X)|Y]] \\ &= \mathbb{E}\left[\mathbb{1}_B(Y)\,\mathbb{P}^{X|Y}(A)\right] = \mathbb{E}\left[\mathbb{1}_B(Y)\,\mathbb{P}^{X|Y}(A)\right] = \mathbb{E}\left[\mathbb{1}_B(Y)\,\mathrm{e}^{-Y}\sum_{x\in A} Y^x/(x!)\right] \\ &= \sum_{x\in A}\mathbb{E}\left[\mathbb{1}_B(Y)\,\mathrm{e}^{-Y}Y^x/(x!)\right] = \sum_{x\in A}\int_B (y^x/(x!))\mathrm{e}^{-2y}\mathrm{d}y .\end{aligned}$$

Therefore the law of (X, Y) has density with respect to $\mu \otimes \lambda$

$$(x, y) \mapsto (y^x/(x!))\mathrm{e}^{-2y} ,$$

where μ is the counting measure on $\mathbb{N}$ and λ is the Lebesgue measure on $\mathbb{R}_+$.

If we take the marginal with respect to X, we get that the distribution of X has for density

$$x \mapsto 2^{-x-1} ,$$

therefore X follows a geometric distribution with parameter $1/2$.

Finally, the conditional density of Y given X is given by for all $y \geq 0$ and $x \in \mathbb{N}$ by

$$2^{-x-1}(y^x/\,(x!))\mathrm{e}^{-2y} .$$

We recognize the Gamma distribution with parameters $x+1$ and $1/2$. Besides the conditional expectation of Y given X is therefore

$$\mathbb{E}\,[Y|\,X] = (X+1)/2 .$$

Solution of Exercise 2.11

1. First note that $S_p \sim \mathcal{P}(\lambda_1 + \cdots, \lambda_p)$. Then for all $k_1, \cdots, k_p \in \mathbb{N}$ and $j \in \mathbb{N}$, we have

$$\mathbb{P}\left(X_1 = k_1, \cdots, X_p = k_p, S_p = j\right) = \mathbb{1}_{\{\sum_{i=1}^p k_i = j\}} \mathbb{P}\left(X_1 = k_1, \cdots, X_p = j - \sum_{i=1}^{p-1} k_i\right)$$
$$\mathbb{1}_{\{\sum_{i=1}^p k_i = j\}} \prod_{i=1}^p \mathrm{e}^{-\lambda_i} \lambda_i^{k_i}/(k_i!) .$$

   Therefore, the conditional law of $(X_1, \cdots, X_p)$ given S_p is

$$\begin{aligned}\mathbb{P}\left(X_1 = k_1, \cdots, X_p = k_p | S_p = j\right) &= \frac{\mathbb{P}\left(X_1 = k_1, \cdots, X_p = k_p, S_p = j\right)}{\mathbb{P}\left(S_p = j\right)} \\ &= \mathbb{1}_{\{\sum_{i=1}^p k_i = j\}} \frac{j!}{(\sum_{i=1}^p \lambda_i)^j} \prod_{i=1}^p \mathrm{e}^{-\lambda_i} \lambda_i^{k_i}/(k_i!)\end{aligned}$$

   It is a multinomial distribution.

2. The distribution of X_1 given $X_1 + X_2$ is a binomial with parameter $(X_1 + X_2, \lambda_1/(\lambda_1 + \lambda_2))$. Hence $\mathbb{E}\,[X_1|\,X_1 + X_2] = (X_1 + X_2)\lambda_1/(\lambda_1 + \lambda_2)$.

Solution of Exercise 2.12

1. Since (X_i) are i.i.d., we have that for all permutation $\sigma \in \mathcal{S}_n$, $(X_{\sigma(1)}, \cdots, X_{\sigma(n)})$ has the same distribution than $(X_1, \cdots, X_n)$. Therefore for all $x_1, \cdots, x_n \in \mathbb{R}$,

$$\begin{aligned}&\mathbb{P}\left(X_{(1)} \leq x_1, \cdots, X_{(n)} \leq x_n\right) \\ &= \sum_{\sigma \in \mathcal{S}_n} \mathbb{P}\left(X_{\sigma(1)} \leq \cdots \leq X_{\sigma(n)}, X_{\sigma(1)} \leq x_1, \cdots, X_{\sigma(n)} \leq x_n\right) \\ &= (n!) \int_{\mathbb{R}^n} \mathbb{1}_{\{y_1 \leq \cdots \leq y_n\}} \mathbb{1}_{\{y_1 \leq x_1\}} \cdots \mathbb{1}_{\{y_n \leq x_n\}} f(y_1) \cdots f(y_n) \mathrm{d}y_1 \cdots \mathrm{d}y_n .\end{aligned}$$

   So $(X_{(1)}, \cdots, X_{(n)})$ has a density with the d-dimensional Lebesgue measure given by $(n!)\mathbb{1}_{\{y_1 \leq \cdots \leq y_n\}} f(y_1) \cdots f(y$

2. We deduce that $(X_{(1)}, X_{(n)})$ has density

$$(x, y) \mapsto f^{(X_{(1)}, X_{(n)})}(x, y) = n!\, f(x) f(y) \mathbb{1}_{\{x \leq y\}} \int_{\mathbb{R}^{n-2}} \mathbb{1}_{\{x \leq y_2 \leq \cdots \leq y_{n-1} \leq y\}}\, f(y_2) \cdots f(y_{n-}) \mathrm{d}y_2 \cdots \mathrm{d}y_{n-1} .$$

   The latter integral can be found also in the computation of

$$\mathbb{P}(x \leq X_{(1,n-2)} \leq \cdots \leq X_{(n-2,n-2)} \leq y) = \int_{\mathbb{R}^{n-2}} \mathbb{1}_{\{x \leq y_2 \leq \cdots \leq y_{n-1} \leq y\}}\, (n-2)!\, f(y_2) \cdots f(y_{n-}) \mathrm{d}y_2 \cdots \mathrm{d}y_{n-1} ,$$

where $X_{(i,n-2)}$ is the i-th order statistic of $X_1, \ldots, X_{n-2}$. Hence we obtain

$$\int_{\mathbb{R}^{n-2}} \mathbb{1}_{\{x \le y_2 \le \cdots \le y_{n-1} \le y\}} \, f(y_2) \cdots f(y_{n-}) \mathrm{d}y_2 \cdots \mathrm{d}y_{n-1} = \frac{(F(y) - F(x))^{n-2}}{(n-2)!} \; .$$

And, finally,

$$f^{(X_{(1)}, X_{(n)})}(x, y) = n(n-1) \, f(x) f(y) \mathbb{1}_{\{x \le y\}} \, (F(y) - F(x))^{n-2} \; .$$

Similarly, we show that $X^{(n)}$ has for density

$$y \mapsto n f(y) (F(y))^{n-1} \; .$$

Therefore, the conditional law of $X_{(1)}$ given $X_{(n)}$ is given by the conditional density

$$f^{X_{(1)}|X_{(n)}}(x|y) = \frac{\mathbb{1}_{\{x \le y\}} (n-1) f(x) (F(y) - F(x))^{n-2}}{(F(y))^{n-1}} \; .$$

3. We obtain

$$\mathbb{E}\left[X_{(1)} \middle| X_{(n)} \right] = \int_{-\infty}^{X_{(n)}} x(n-1) f(x) \frac{(F(X_{(n)}) - F(x))^{n-2}}{(F(X_{(n)}))^{n-1}} \, \mathrm{d}x \; .$$

Solution of Exercise 2.13 1. Tout d'abord, on peut définir σ comme étant la variable aléatoire à valeurs dans $\mathcal{S}_n$ telle que

$$X_{\sigma(1)} \le \cdots \le X_{\sigma(n)} \; .$$

Or par définition on a :

$$\begin{aligned} 1 &= \mathbb{P}\left(X_{\sigma(1)} \le \cdots \le X_{\sigma(n)} \right) \\ &= \mathbb{P}\left(\bigcup_{i,j=1; i \ne j}^{n} \{X_i = X_j\} \right) + \mathbb{P}\left(X_{\sigma(1)} < \cdots < X_{\sigma(n)} \right) \\ &= \mathbb{P}\left(X_{\sigma(1)} < \cdots < X_{\sigma(n)} \right) \; , \end{aligned}$$

car $(X_1, \cdots, X_n)$ admet une densité f sur $\mathbb{R}^n$.
Remarquez que pour tout $i \in \{1, \cdots, n\}$, $R(i)$ représente le nombre de $(X_j)_j$ plus petit que X_i, donc $R = \sigma^{-1}$.

2. On a pour toute fonction mesurable bornée $g : \mathbb{R}^n \to \mathbb{R}$,

$$\begin{aligned} &\mathbb{E}\left[g(X_{\sigma(1)}, \cdots, X_{\sigma(n)}) \mathbb{1}_{\{R=r\}} \right] = \mathbb{E}\left[g(X_{r^{-1}(1)}, \cdots, X_{r^{-1}(n)}) \mathbb{1}_{\{X_{r^{-1}(1)} \le \cdots \le X_{r^{-1}(n)}\}} \right] \\ &= \int_{\mathbb{R}^n} g(x_{r^{-1}(1)}, \cdots, x_{r^{-1}(n)}) \mathbb{1}_{\{x_{r^{-1}(1)} \le \cdots \le x_{r^{-1}(n)}\}} f(x_1, \cdot, x_n) \mathrm{d}x_1 \cdots \mathrm{d}x_n \\ &= \int_{\mathbb{R}^n} g(x_1, \cdots, x_n) \mathbb{1}_{\{x_1 \le \cdots \le x_n\}} f(x_{r(1)}, \cdot, x_{r(n)}) \mathrm{d}x_1 \cdots \mathrm{d}x_n \; , \end{aligned}$$

par permutation des coordonnées associées à r (C^1-difféomorphisme avec un jacobien de valeurs absolue 1).

3. Si on note Leb^n la mesure de Lebesgue sur $\mathbb{R}^n$ et μ la mesure de comptage sur $\mathcal{S}_n$, on en déduit que $(X_{\sigma(1)}, \cdots, X_{\sigma(n)}, R)$ admet une densité par rapport à $\mathrm{Leb}^n \otimes \mu$ donnéé par

$$f^{(X_{\sigma(1)}, \cdots, X_{\sigma(n)}, R)}(x_1, \cdots, x_n, r) = \mathbb{1}_{\{x_1 \le \cdots \le x_n\}} f(x_{r(1)}, \cdot, x_{r(n)}) \; . \tag{15}$$

Aussi comme $(X_{\sigma(1)}, \cdots, X_{\sigma(n)})$ admet une densité par rapport à la mesure de Lebesgue Leb^n

$$f^{(X_{\sigma(1)}, \cdots, X_{\sigma(n)})}(x_1, \cdots, x_n) = \sum_{\tau \in \mathcal{S}_n} \mathbb{1}_{\{x_1 \leq \cdots \leq x_n\}} f(x_{\tau(1)}, \cdots, x_{\tau(n)}), \tag{16}$$

on en déduit par le théoreme 2.2.5 que la densité conditionnelle de R sachant $(X_{\sigma(1)}, \cdots, X_{\sigma(n)})$ par rapport à la mesure de comptage est donnée par

$$f^{R \mid (X_{\sigma(1)}, \cdots, X_{\sigma(n)})}(r) = \frac{f^{(X_{\sigma(1)}, \cdots, X_{\sigma(n)}, R)}(x_1, \cdots, x_n, r)}{f^{(X_{\sigma(1)}, \cdots, X_{\sigma(n)})}(x_1, \cdots, x_n)} \ . \tag{17}$$

4. Dans le cas où les X_i sont i.i.d. de même densité ϕ sur $\mathbb{R}$, (15) et (16) deviennent

$$f^{(X_{\sigma(1)}, \cdots, X_{\sigma(n)})}(x_1, \cdots, x_n) = \mathbb{1}_{\{x_1 \leq \cdots \leq x_n\}} \phi(x_1) \cdots \phi(x_n)$$
$$f^{(X_{\sigma(1)}, \cdots, X_{\sigma(n)})}(x_1, \cdots, x_n) = (n!) \mathbb{1}_{\{x_1 \leq \cdots \leq x_n\}} \phi(x_1) \cdots \phi(x_n) \ .$$

Donc (17) peut se réécrire simplement:

$$f^{R \mid (X_{\sigma(1)}, \cdots, X_{\sigma(n)})}(r) = 1/(n!) \ .$$

Donc R est indépendant de $(X_{\sigma(1)}, \cdots, X_{\sigma(n)})$ et suit une loi uniforme sur $\mathcal{S}_n$.

Solution of Exercise 3.1 1. Note that the mean of p_θ is given by

$$\int_\theta^\infty x\mathrm{e}^{\theta-x}\mathrm{d}x = \int_\theta^\infty (x-\theta)\mathrm{e}^{\theta-x}\mathrm{d}x + \int_\theta^\infty \theta\mathrm{e}^{\theta-x}\mathrm{d}x = 1+\theta\ .$$

Therefore an unbiased estimator is

$$\hat{\theta}_n = n^{-1}\sum_{i=1}^n (X_i - 1)\ .$$

2. Since $\hat{\theta}$ is unbiased, we have

$$\begin{aligned}
\mathbb{E}_\theta\left[(\hat{\theta}_n-\theta)^2\right] &= n^{-2}\sum_{i=1}^n \mathrm{Var}_\theta\left\{X_i-1\right\} \\
&= n^{-2}\sum_{i=1}^n \int_\theta^\infty (x-1-\theta)^2\mathrm{e}^{\theta-x}\mathrm{d}x \\
&= n^{-2}\sum_{i=1}^n \int_0^\infty (u-1)^2\mathrm{e}^{-u}\mathrm{d}u \\
&\text{where we use the change of variable } x-\theta = u \\
&= n^{-1}\ .
\end{aligned}$$

Observe that we have used that the variace of an exponential random variable of mean 1 is 1.

3. We show now that under $\mathbb{P}_\theta$, $\tilde{\theta}_n$ has a density with respect to the Lebesgue measure. Let $f:\mathbb{R}\to\mathbb{R}$ be a bounded measurable function. Then denoting by $m(x)=\min\{x_i\ 1\le i\le n\}$ we have

$$\begin{aligned}
\mathbb{E}_\theta\left[f(\tilde{\theta}_n)\right] &= \int_{\mathbb{R}^n} f(m(x))\mathrm{e}^{n\theta-\sum_{i=1}^n x_i}\mathbb{1}_{\{m(x)\ge\theta\}}\mathrm{d}x_1\cdots\mathrm{d}x_n \\
&= \sum_{j=1}^n \int_{\mathbb{R}^n} f(x_j)\mathrm{e}^{n\theta-\sum_{i=1}^n x_i}\mathbb{1}_{\{m(x)=x_j\}}\mathbb{1}_{\{x_j\ge\theta\}}\mathrm{d}x_1\cdots\mathrm{d}x_n \\
&= \sum_{j=1}^n \int_{\mathbb{R}} f(x_j)\mathrm{e}^{-(x_j-\theta)}\mathbb{1}_{\{x_j\ge\theta\}} \\
&\quad\times\int_{\mathbb{R}^{n-1}} \mathrm{e}^{(n-1)\theta-\sum_{i=1;i\ne j}^n x_i}\mathbb{1}_{\{m(x)=x_j\}}\mathbb{1}_{\{x_j\ge\theta\}}\mathrm{d}x_1\cdots\mathrm{d}x_{j-1}\mathrm{d}x_{j+1}\cdots\mathrm{d}x_n\mathrm{d}x_j \\
&\text{using Fubini's theorem} \\
&= \sum_{j=1}^n \int_{\mathbb{R}} f(x_j)\mathrm{e}^{-(x_j-\theta)}\mathbb{1}_{\{x_j\ge\theta\}}A(x_j)^{n-1}\mathrm{d}x_j\ ,
\end{aligned}$$

where

$$A(x_j) = \int_{x_j}^\infty \mathrm{e}^{\theta-t}\mathrm{d}t = \mathrm{e}^{\theta-x_j}\ .$$

Finally, it follows that under $\mathbb{P}_\theta$, $\tilde{\theta}_n$ has a density given by

$$f_{\tilde{\theta}_n}(u) = n\mathrm{e}^{-n(u-\theta)}\mathbb{1}_{\{u\ge\theta\}}\ .$$

4. A simple calculation gives that

$$\mathbb{E}_\theta\left[\tilde{\theta}_n\right] = \int_\theta^\infty un\mathrm{e}^{-n(u-\theta)}\mathrm{d}u = \theta+n^{-1}\ .$$

Thus, the estimator $\tilde{\theta}_n$ has bias n^{-1}

5. By the previous questions, we have

$$\mathbb{E}_\theta\left[(\theta-\tilde{\theta}_n)^2\right] = n^{-2} + \mathrm{Var}_\theta\left\{\theta-\tilde{\theta}_n\right\}$$
$$= 2/n^2 \ .$$

6. The conclusion is that even if $\tilde{\theta}_n$ is biased, its mean square error is much smaller than the one of $\hat{\theta}_n$, which is unbiased yet.

Solution of Exercise 3.2 1. Note that a Bernouilli random variable with parameter θ has density with respect to the counting measure on $\{0,1\}$ given by $p_\theta(x) = \theta^x(1-\theta)^{1-x}$. So, for all $\theta \in (0,1)$, $\mathbb{P}_\theta$ has a density with respect to the counting measure ν on $\{0,1\}^n$ given by

$$\frac{\mathrm{d}\mathbb{P}_\theta}{\mathrm{d}\nu}(x_{1:n}) = \prod_{i=1}^n p_\theta(x_i) = \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i} \ .$$

Therefore by the Fisher's factorization theorem, $S_n = \sum_{i=1}^n X_i$ is a sufficient statistics. We can also compute explicitly in that case the conditionnal law of X_i given S_n for all i, it is given by

$$\mathbb{E}_\theta\left[\mathbb{1}_{\{X_i=1\}} \mid S_n\right] = \mathbb{E}_\theta\left[X_i \mid S_n\right]$$
$$= n^{-1}\sum_j \mathbb{E}_\theta\left[X_j \mid S_n\right]$$
$$\text{since the } (X_i)_i \text{ are i.i.d.see question 2}$$
$$= n^{-1}\mathbb{E}_\theta\left[S_n \mid S_n\right]$$
$$= n^{-1}S_n \ .$$

Therefore, the conditionnal law of X_i given S_n is

$$K(S_n,\cdot) = n^{-1}S_n\delta_1 + (1-n^{-1}S_n)\delta_0 \ ,$$

therefore it does not depend on θ, and S_n is a sufficient statistics.

2. Set $\hat{\theta} = \alpha(X_1 - X_2)^2$. We are looking for α such that the mean of $\hat{\theta}$ under $\mathbb{P}_\theta$ is the variance of X_1, $\theta(1-\theta)$. But a simple calculation lead to

$$\mathbb{E}_\theta\left[\hat{\theta}\right] = \mathbb{E}_\theta\left[\alpha(X_1-\theta+\theta-X_2)^2\right] = 2\alpha\theta(1-\theta) \ .$$

Thus, we need to take $\alpha = 1/2$.

3. First remark: $X_i^2 = X_i$!
Note that since the $(X_i)_i$ are independent, then for all i, j, $i \neq j$,

$$\mathbb{E}_\theta\left[X_iX_j \mid S_n\right] = \mathbb{E}_\theta\left[X_1X_2 \mid S_n\right] \ . \tag{18}$$

Indeed, for all i, j, $i \neq j$ and all $k \in \mathbb{N}$,

$$\mathbb{E}_\theta\left[X_iX_j\mathbb{1}_{\{S_n=k\}}\right] = \mathbb{E}_\theta\left[X_1X_2\mathbb{1}_{\{S_n=k\}}\right] \ .$$

Similarly, we have for all i, $\mathbb{E}_\theta\left[X_i \mid S_n\right] = \mathbb{E}_\theta\left[X_1 \mid S_n\right]$, and therefore

$$\mathbb{E}_\theta\left[X_i \mid S_n\right] = n^{-1}S_n \ . \tag{19}$$

Note also, that

$$S_n^2 = \sum_{i\neq j} X_iX_j + \sum_{i=1}^n X_i^2 = \sum_{i\neq j} X_iX_j + \sum_{i=1}^n X_i = \sum_{i\neq j} X_iX_j + S_n \ . \tag{20}$$

Therfore using (18), we have

$$\mathbb{E}_\theta\left[X_1X_2 \mid S_n\right] = (n(n-1))^{-1}(S_n^2 - S_n) \ .$$

Using this result and (19), we have

$$\mathbb{E}_\theta\left[\hat{\theta} \mid S_n\right] = (1/2)(2n^{-1}S_n + 2(n(n-1))^{-1}(S_n^2 - S_n)) \ .$$

Solution of Exercise 3.3 1. This question is easy and left to the reader. Just the beginning of the proof is for all $k \in \mathbb{N}$,

$$\mathbb{P}\left(Y_1 + Y_2 = k\right) = \sum_{n \geq 0} \mathbb{P}\left(Y_1 + Y_2 = k, Y_1 = n\right) = \sum_{n \geq 0} \mathbb{P}\left(Y_2 = k - n\right) \mathbb{P}\left(Y_1 = n\right) \ .$$

2. For all $\theta > 0$, the Poisson distribution with parameter θ has a density with respect to the counting measure on $\mathbb{N}$ given by $p_\theta(x) = \theta^x \mathrm{e}^{-\theta}/(x!)$ Therefore, for all θ, $\mathbb{P}_\theta$ has a density with respect to the counting measure on $\mathbb{N}^n$ given by

$$f_\theta(x_{1:n}) = \prod_{i=1}^n p_\theta(x_i) = \theta^{\sum_{i=1}^n x_i} \mathrm{e}^{-n\theta} \Big/ \left(\prod_{i=1}^n x_i!\right) \ .$$

And, it is on the form $h \times \tilde{f}_\theta \circ g$ for

$$h(x_{1:n}) = \left(\prod_{i=1}^n x_i!\right)^{-1} \ , \ g(x_{1:n}) = \sum_{i=1}^n x_i \ , \ \tilde{f}_\theta(t) = \theta^t \mathrm{e}^{-n\theta} \ .$$

By the Fisher's factorization theorem, S_n is a sufficient statistics.

3. For all $\theta > 0$, $k, j \in \mathbb{N}$, we have since $X_1 \sim \theta$ and $S_n - X_1 \sim \mathcal{P}((n-1)\theta)$ are independent,

$$\begin{aligned}\mathbb{P}_\theta\left(X_1 = k, S_n = j\right) &= \mathbb{P}_\theta\left(X_1 = k, S_n - X_1 = j - k\right) \\ &= \mathbb{1}_{\{j \geq k\}} \mathbb{P}_\theta\left(X_1 = k\right) \mathbb{P}_\theta\left(S_n - X_1 = j - k\right) \\ &= \mathbb{1}_{\{j \geq k\}} \theta^k \binom{j}{k} (j!)^{-1} ((n-1)\theta)^{j-k} \mathrm{e}^{-n\theta} \ .\end{aligned}$$

Also since $\mathbb{P}_\theta(S_n = j) = (n\theta)^j \mathrm{e}^{-n\theta}/(j!)$, we get

$$\mathbb{P}_\theta\left(X_1 \geq x_0 \mid S_n = j\right) = \sum_{k \geq x_0} \mathbb{1}_{\{j \geq k\}} \binom{j}{k} \frac{((n-1))^{j-k}}{n^j} \ .$$

So the conditional distribution of X_1 given S_n is a Binomial distribution with parameter $(1/n, S_n)$.

4. For $x_0 = 0$, we find 1 and for $x_0 = 1$, we find $1 - (1 - 1/n)^{S_n}$.

Solution of Exercise 4.1 Consider the matrix

$$B = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} .$$

So, $(X_1, X_2) = BX$ and it is a Gaussian vector by assumption on X. Besides denote by μ and Σ the mean and the covariance matrix of X respectively. Then, (X_1, X_2) has mean $B\mu$ and covariance matrix

$$B\Sigma B^{\mathrm{T}} = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \Sigma \begin{pmatrix} A_1^{\mathrm{T}} & A_2^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \begin{pmatrix} \Sigma A_1^{\mathrm{T}} & \Sigma A_2^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} A_1 \Sigma A_1^{\mathrm{T}} & A_1 \Sigma A_2^{\mathrm{T}} \\ A_2 \Sigma A_1^{\mathrm{T}} & A_2 \Sigma A_2^{\mathrm{T}} \end{pmatrix}$$

Then X_1 and X_2 are independent if and only if $A_1 \Sigma A_2^{\mathrm{T}} = 0$.

Solution of Exercise 4.2 Let $f : \mathbb{R} \to \mathbb{R}$ be a bounded measurable function. Since U is independent of X, we have

$$\mathbb{E}[f(Y)] = \mathbb{E}[\mathbf{1}_{\{U=1\}} f(X)] + \mathbb{E}[\mathbf{1}_{\{U=0\}} f(-X)] = 2^{-1}(\mathbb{E}[f(X)] + \mathbb{E}[f(-X)]) = \mathbb{E}[f(X)].$$

where we have used that X and $-X$ has the same distribution. Therefore $Y \sim \mathcal{N}(0, 1)$.

Let us study Cov (X, Y). Using that U is independent of X again, we get

$$\text{Cov}\,(X, Y) = \mathbb{E}[YX] = \mathbb{E}[\mathbf{1}_{\{U=1\}} X^2] - \mathbb{E}[\mathbf{1}_{\{U=0\}} X^2] = 0.$$

Nevertheless X and Y are not independent since

$$\mathbb{P}(X \in [0, 1], Y \in [3, 4]) = 0 \neq \mathbb{P}(X \in [0, 1])\mathbb{P}(Y \in [3, 4])$$

Solution of Exercise 4.3

1. Since Γ is a nonnegative definite Hermitian matrix, there exists a unitary matrix U such that

$$U^H \Gamma U = \text{diag}(\lambda_1, \cdots, \lambda_n) ,$$

where $(\lambda_i)_{1 \leq i \leq n}$ are the eigenvalues of Γ. Now let X be n-dimensional zero-mean Gaussian random variable with covariance matrix $\text{diag}(\lambda_1, \cdots, \lambda_n)$. Then the covariance matrix of $Z = UX$ is by the linearity of the expectation,

$$\mathbb{E}\left[ZZ^H\right] = U\mathbb{E}\left[XX^T\right] U^H = \Gamma .$$

2. First, since Γ is Hermitian, we have that $\Gamma^H = \Gamma$, which implies that

$$\text{Im}(\Gamma) = -\text{Im}(\Gamma)^T . \tag{21}$$

Therefore, Σ is symmetric. Let us show that it is positive definite. Let $u \in \mathbb{R}^{2n}$, we need to show that $u^T \Sigma u > 0$ if $u \neq 0$. But if we write $u^T = [u_1^T, u_2^T]^T$, using (21) and the assumption on Γ,

$$\begin{aligned} u\Sigma u &= (1/2)\left(u_1^T \text{Re}(\Gamma) u_1 + u_2^T \text{Re}(\Gamma) u_2 - u_1^T \text{Im}(\Gamma) u_2 + u_2^T \text{Im}(\Gamma) u_1\right) \\ &= (1/2)[u_1^T - \mathrm{i} u_2^T]\Gamma[u_1^T + \mathrm{i} u_2^T]^T > 0 . \end{aligned}$$

3. A simple calculation leads to $\mathbb{E}\left[ZZ^H\right] = \Gamma$.

4. A simple calculation leads to $\mathbb{E}\left[ZZ^T\right] = 0$.

5. To apply Kolmogorov's theorem, we need to define a family of compatible (or consistent) probability measure $\{\nu_I \mid I \subset T, I \text{ finite }\}$ such that for all $I \subset T, I$ finite, ν_I is a probability measure on $\mathbb{C}^I$, for all $J \subset I$, we have (this is the definition that the family is compatible)

$$\nu_J = \nu_I \circ \Pi_{I,J}^{-1} , \tag{22}$$

Furthermore, we need to have

$$\text{for all } t \in T, X_t \sim \nu_t \ , \mathbb{E}\left[X_t\right] = \mu(t) \ , \tag{23}$$
$$\text{and for all } s,t \in T, (X_s, X_t) \sim \nu_{s,t} \ , \mathrm{Cov}(X_s, X_t) = \gamma(s,t) \ . \tag{24}$$

Let us define this family of probability measure. For all $I \subset T, I$ finite, consider,

$$\mu_I = \{\mu(i_j) \mid i_j \in I\} \ , \ \Gamma_I = \{\gamma(i_j, i_k) \mid i_j, i_k \in I\} \ ,$$
$$\Sigma_I = (1/2)\begin{pmatrix} \mathrm{Re}(\Gamma_I) & -\mathrm{Im}(\Gamma_I) \\ \mathrm{Im}(\Gamma_I) & \mathrm{Re}(\Gamma_I) \ , \end{pmatrix}$$

and $(X_I, Y_I) \sim \mathcal{N}(0, \Sigma_I)$, $Z_I = \mu_I + X_I + \mathrm{i}Y_I$. Then define ν_I to be the distribution of Z_I. Then, we have that for all $J \subset I$, $\nu_I \circ \Pi_{I,J}$ is the distribution of Z_J which is by definition ν_J. Therefore, (22) holds and $\{\nu_I \mid I \subset T, I \text{ finite }\}$ is compatible. Therefore we can apply Kolmogorov's theorem. Furthermore, (23) easily holds and an application of 3. for $I = \{s,t\}$ gives (24).

Solution of Exercise 4.4 See solution of Exercice 5.2.

Solution of Exercise 4.6

1. If Y is $\mathcal{F}_\tau$-measurable, then for all $A \in \mathcal{B}(\mathbb{R})$, we have $\{Y \in A\} \in \mathcal{F}_\tau$. Therefore, for all $A \in \mathcal{B}(\mathbb{R})$

$$\{Y \mathbb{1}_{\{\tau \le n\}} \in A\} = (\{Y \in A\} \cap \{\tau \le n\}) \cup (\{0 \in A\} \cap \{\tau > n\}) \ , \tag{25}$$

with the convention $\{0 \in A\} = E$ if $0 \in A$ and $\emptyset$ otherwise. Since by definition of $\mathcal{F}_\tau$, $(\{Y \in A\} \cap \{\tau \le n\}) \in \mathcal{F}_n$ and $\{\tau > n\} \in \mathcal{F}_n$, it follows that $\{Y \mathbb{1}_{\{\tau \le n\}} \in A\} \in \mathcal{F}_n$.
Now if $Y \mathbb{1}_{\{\tau \le n\}}$ is $\mathcal{F}_n$-measurable for all $n \in \mathbb{N}$, then using (25) again and since the union is disjoint, we can write for all $A \in \mathcal{F}$,

$$\{Y \in A\} \cap \{\tau \le n\} = \{Y \mathbb{1}_{\{\tau \le n\}} \in A\} \setminus (\{0 \in A\} \cap \{\tau > n\}) \ ,$$

and therefore since $\{\tau > n\} \in \mathcal{F}_n$, it follows that for all $n \ge 0$, $\{Y \in A\} \cap \{\tau \le n\} \in \mathcal{F}_n$. By definition, Y is $\mathcal{F}_\tau$-measurable.

2. This question is easy and left to the reader.

3. Denote by $Z = \mathbb{1}_{\{\tau = n\}} Y$. Then by the first question, it enough to show that for all $k \ge 0$, $Z\mathbb{1}_{\{\tau \le k\}}$ is $\mathcal{F}_k$-measurable. But for all $k < n$, $Z\mathbb{1}_{\{\tau \le k\}} = 0$ is, and for all $k \ge n$, $Z\mathbb{1}_{\{\tau \le k\}} = Z$ also, since Z is $\mathcal{F}_n$-measurable and $\mathcal{F}_n \subset \mathcal{F}_k$.

4. For all $A \in \mathcal{F}_\tau$, we have by the second question

$$\begin{aligned}\mathbb{E}\left[X \mathbb{1}_{\{\tau = n\}} \mathbb{1}_A\right] &= \mathbb{E}\left[X \mathbb{1}_{A \cap \{\tau = n\}}\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{F}_n\right] \mathbb{1}_{A \cap \{\tau = n\}}\right] \ .\end{aligned}$$

Since by the previous question $\mathbb{E}\left[X \mid \mathcal{F}_n\right] \mathbb{1}_{\{\tau = n\}}$ is $\mathcal{F}_\tau$-measurable, the proof is concluded.

Solution of Exercise 5.1 We need to show that the process $(Z_t)_{t\in\mathbb{Z}}$ is a second order stationary process, where $(X_t)_{t\in\mathbb{Z}}$ and $(Y_t)_{t\in\mathbb{Z}}$ are uncorrelated second order stationary processes. First note for all $t \in \mathbb{Z}$

$$\mathbb{E}[Z_t] = \mathbb{E}[X_t] + \mathbb{E}[Y_t] = \mathbb{E}[X_0] + \mathbb{E}[Y_0] = \mathbb{E}[Z_0] \ .$$

Second, for all $t \in \mathbb{Z}$ and $h \in \mathbb{Z}$, we have since $(X_t)_{t\in\mathbb{Z}}$ and $(Y_t)_{t\in\mathbb{Z}}$ are uncorrelated:

$$\begin{aligned}&\mathbb{E}\left[(Z_t - \mathbb{E}[Z_0])(Z_{t+h} - \mathbb{E}[Z_0])^H\right] \\ &= \mathbb{E}\left[(X_t - \mathbb{E}[X_0])(X_{t+h} - \mathbb{E}[X_0])^H\right] + \mathbb{E}\left[(Y_t - \mathbb{E}[Y_0])(Y_{t+h} - \mathbb{E}[Y_0])^H\right] = \gamma_X(h) + \gamma_Y(h) \ ,\end{aligned}$$

where γ_X and γ_Y is the autocovariance function of (X_t) and (Y_t). In the case where these two processes are complex valued, by the Herglotz theorem, we have two measures ν_X and ν_Y on $\mathbb{T}$ such that for all $h \in \mathbb{Z}$,

$$\gamma_X(h) + \gamma_Y(h) = \int_{\mathbb{T}} \mathrm{e}^{\mathrm{i}h\lambda} \left\{\nu_X + \nu_Y\right\}(\mathrm{d}\lambda) \ .$$

Therefore, the spectral measure of $(Z_t)_{t\in\mathbb{Z}}$ is $\nu_X + \nu_Y$.

Solution of Exercise 5.2

1. $(Y_t)_{t\in\mathbb{Z}}$ is a linear filter of $(\varepsilon_t)_{t\in\mathbb{Z}}$ which is shift invariant so it is strongly stationnary if $(\varepsilon_t)_{t\in\mathbb{Z}}$ is.
For all $t \in \mathbb{Z}$,
$$\mathbb{E}[Y_t] = a \ .$$
For all $t \in \mathbb{Z}$ and $h \in \mathbb{Z}$, we have
$$\mathbb{E}[Y_t Y_{t+h}] = \begin{cases} \sigma^2 b^2 + \sigma^2 c^2 & \text{if } h = 0 \\ \sigma^2 cb & \text{if } h = \pm 1 \\ 0 & \text{otherwise} \end{cases}$$

2. $Y_t = a + b\varepsilon_t + c\varepsilon_{t+1}$.
$(Y_t)_{t\in\mathbb{Z}}$ is a linear filter of $(\varepsilon_t)_{t\in\mathbb{Z}}$ which is shift invariant so it is strongly stationnary if $(\varepsilon_t)_{t\in\mathbb{Z}}$ is.
For all $t \in \mathbb{Z}$,
$$\mathbb{E}[Y_t] = a \ .$$
For all $t \in \mathbb{Z}$ and $h \in \mathbb{Z}$, we have
$$\mathbb{E}\left[Y_t \overline{Y}_{t+h}\right] = \begin{cases} \sigma^2 b^2 + \sigma^2 c^2 & \text{if } h = 0 \\ \sigma^2 \bar{c} b & \text{if } h = -1 \\ \sigma^2 c\bar{b} & \text{if } h = 1 \\ 0 & \text{otherwise} \end{cases}$$

3. $(Y_t)_{t\in\mathbb{Z}}$ is a linear filter of $(\varepsilon_t)_{t\in\mathbb{Z}}$ which is shift invariant so it is strongly stationnary if $(\varepsilon_t)_{t\in\mathbb{Z}}$ is.
For all $t \in \mathbb{Z}$, since $\rho < 1$, by the usual integration theorems,
$$\mathbb{E}[Y_t] = \sum_{j=0}^{\infty} \rho^j \mathbb{E}[\varepsilon_{t-j}] = 0 \ .$$
For all $t \in \mathbb{Z}$ and $h \in \mathbb{N}$, we have by the usual integration theorems
$$\mathbb{E}[Y_t Y_{t+h}] = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \rho^{j+k} \mathbb{E}[\varepsilon_{t-j}\varepsilon_{t+h-k}] = \sum_{j=0}^{\infty} \rho^{2j+h}\sigma^2 = \rho^h \sigma^2 (1-\rho^2)^{-1} \ .$$

4. $(Y_t)_{t\in\mathbb{Z}}$ is a (non-linear) filter of $(\varepsilon_t)_{t\in\mathbb{Z}}$ which is shift invariant so it is strongly stationnary if $(\varepsilon_t)_{t\in\mathbb{Z}}$ is.
   For all $t \in \mathbb{Z}$,
   $$\mathbb{E}\left[Y_t\right] = 0 \; .$$
   For all $t \in \mathbb{Z}$ and $h \in \mathbb{Z}$, we have by the usual integration theorems
   $$\mathbb{E}\left[Y_t Y_{t+h}\right] = \begin{cases} \sigma^4 & \text{if } h = 0 \\ 0 & \text{otherwise} \; . \end{cases}$$

5. This time, $(Y_t)_{t\in\mathbb{Z}}$ is not strongly stationnary if $(\varepsilon_t)_{t\in\mathbb{Z}}$ is. Indeed, if $t \in \mathbb{Z}$ is odd, $Y_t = 0$ and if $t \in \mathbb{Z}$ is even, $Y_t = 2\varepsilon_t$. But, if $(Y_t)_{t\in\mathbb{Z}}$ is strongly stationnary, for all $t \in \mathbb{Z}$, Y_t and Y_{t+1} have the same distribution.
   Following the same lines, we have the autocovariance function depends on t, so it is not a weakly stationary process neither.

Solution of Exercise 5.3 1. Since $(\chi(h))_{h\in\mathbb{Z}} \in \ell_2(\mathbb{Z})$, by the Herglotz theorem, we have that the spectral density is given by
$$f(\lambda) = \frac{1}{2\pi}\sum_{h\in\mathbb{Z}} \chi(h)\mathrm{e}^{-\mathrm{i}h\lambda} = \frac{1}{2\pi}\left\{1 + 2\rho\cos(\lambda)\right\} \; .$$
A necessary condition then for χ to be an autocovariance function is that f is nonnegative. which is the case if and only if $|\rho| \leq 1/2$.

2. Let $(\varepsilon_t)_{t\in\mathbb{Z}}$ be i.i.d. real standard Gaussian random variable, and
   $$X_t = a\varepsilon_t + b\varepsilon_{t-1} \; ,$$
   where $a, b \in \mathbb{R}$. Then by the correction of Example 5.2.4, χ is the autocovariance function of X if $a^2 + b^2 = 1$ and $ab = \rho$. Therefore,
   $$a = (1/2)(\sqrt{1+2\rho} - \sqrt{1+2\rho})$$
   $$b = (1/2)(\sqrt{1+2\rho} + \sqrt{1+2\rho}) \; .$$
   We find the condition $|\rho| \leq 1/2$ again.

Solution of Exercise 5.4 1. We can write for all $t \geq 1$,
$$\Sigma_t = (1-\rho)\mathrm{Id}_t + [\rho] \; ,$$
where $[\rho]$ is the t-dimensional matrix whose entries are all ρ. Therefore, the eigenvalues of Σ_t are $(1-\rho)$ and $1 + (t-1)\rho$. So, Σ_t is symmetric nonnegative positive matrix if and only if $\rho \in [0, 1]$.

2. Let $(\epsilon_t)_{t\in\mathbb{Z}}$ be i.i.d. real standard Gaussian random variable, Z be an other real standard Gaussian random variable independent of $(\epsilon_t)_{t\in\mathbb{Z}}$, and consider
   $$X_t = a\epsilon_t + bZ \; .$$
   Then, for all t, $\mathbb{E}\left[X_t\right] = 0$ and for all $t \in \mathbb{Z}$, $h \in \mathbb{Z}$,
   $$\mathbb{E}\left[X_t X_{t+h}\right] = \begin{cases} a^2 + b^2 & \text{if } h = 0 \\ b^2 & \text{otherwise} \; . \end{cases}$$
   Therefore, Σ_t is the finite autocovariance matrix for $(X_t)_{t\in\mathbb{Z}}$ if $a^2 + b^2 = 1$ and $b^2 = \rho$. We find the condition that $\rho \in [0, 1]$ again.

Solution of Exercise 5.5 1. By the characterization of the projection on a finite dimensional sub-space of L^2, we have since X, Y are centered

$$\mathrm{proj}(X \mid \mathrm{Span}(Y)) = \rho Y \ .$$

2. By the properties of the orthogonal projection and X, Y are centered, we have

$$\mathbb{E}\left[(X - \mathrm{proj}(X \mid \mathrm{Span}(Y)))^2\right] = \mathbb{E}\left[(X - \mathrm{proj}(X \mid \mathrm{Span}(Y)))X\right] \\ \mathrm{Var}\,(X) - \rho^2\,\mathrm{Var}\,(Y) \ .$$

Solution of Exercise 5.6 Let Γ_t be the autocovariance matrix of $(Y_1, \cdots, Y_t)$. Then, λ is an eigenvalue of Γ_t if and only is there exists $x \in \mathbb{R}^t$

$$x^H \Gamma_t x = \lambda\,|x|^2 \ . \tag{26}$$

But

$$\begin{aligned} x^H \Gamma_t x &= \sum_{p,j=1}^{t} \bar{x}_p \gamma(p-j) x_j \\ &= \sum_{p,j=1}^{t} \bar{x}_p \left\{ \int_{\mathbb{T}} \mathrm{e}^{\mathrm{i}\lambda(p-j)} f(\lambda)\mathrm{d}\lambda \right\} x_j = \int_{\mathbb{T}} f(\lambda)\varphi(\lambda)\mathrm{d}\lambda \ , \end{aligned}$$

where

$$\varphi(\lambda) = \overline{\left(\sum_{p=1}^{t} x_p \mathrm{e}^{-\mathrm{i}\lambda p}\right)} \left(\sum_{j=1}^{t} x_p \mathrm{e}^{-\mathrm{i}\lambda j}\right) = \left|\left(\sum_{p=1}^{t} x_p \mathrm{e}^{-\mathrm{i}\lambda p}\right)\right|^2 \ .$$

Since

$$|x|^2 = (2\pi)^{-1} \int_{\mathbb{T}} \mathrm{e}^{\mathrm{i}\lambda(p-j)} \varphi(\lambda)\mathrm{d}\lambda \ ,$$

(26) holds if and only if

$$(2\pi)^{-1}\lambda \int_{\mathbb{T}} \mathrm{e}^{\mathrm{i}\lambda(p-j)} \varphi(\lambda)\mathrm{d}\lambda = \int_{\mathbb{T}} f(\lambda)\varphi(\lambda)\mathrm{d}\lambda \ .$$

Since f takes values in $[m, M]$, this implies that $\lambda \in [2\pi m, 2\pi M]$.

Solution of Exercise 5.7 1. If we take $a = -1$, then by the Herglotz theorem, we have that $(X_t)_{t\in\mathbb{Z}}$ has a spectral density f given by

$$f(\lambda) = \pi^{-1}(1 - \cos(\lambda)) \ ,$$

which satisfies the assumptions of the exercise.

2. By the spectral representation, we have for all $t \in \mathbb{Z}$, $X_t = \int \mathrm{e}^{\mathrm{i}t\lambda} \mathrm{d}\hat{X}(\lambda)$, so we have

$$g_t(\lambda) = \sum_{k=0}^{t} \mathrm{e}^{-k\mathrm{i}\lambda} \ .$$

3. Since for all $n \geq 1$, we have

$$\int_{\mathbb{T}} \left| \frac{1}{n} \sum_{k=1}^{n} \mathrm{e}^{-\mathrm{i}k\lambda} \right|^2 \mathrm{d}\lambda = n^{-1} \ ,$$

the limit as $n \to \infty$ is 0.

4. We need to show that $\phi : \lambda \mapsto (1-\mathrm{e}^{-\mathrm{i}\lambda})^{-1}$ belongs to $L^2(\mathbb{T},\nu)$ where ν is the spectral measure of $(X_t)_{t\in\mathbb{Z}}$. So, we need to that $\lambda \mapsto f(\lambda)|\phi|^2(\lambda)$ is integrable near 0. But by the Taylor-Lagrange theorem, we can write for all $\lambda \in \mathbb{T}$,
$$f(\lambda) = f'(0)\lambda + f''(\epsilon)\lambda^2/2\ ,\ \epsilon \in [0,\lambda]\ , \tag{27}$$
and we have
$$|\phi(\lambda)|^{-1} = 2 - 2\cos(\lambda)\ . \tag{28}$$
Using (27) and (28), it follows that $\lambda \mapsto f(\lambda)|\phi|^2(\lambda)$ can be continuously extended on $\mathbb{T}$. So, Z is well defined.

5. We show that $n^{-1}\sum_{j=0}^{n} g_j(\lambda) = \sum_{j=0}^{T}\sum_{k=0}^{j} \mathrm{e}^{-\mathrm{i}k\lambda}$ converges to ϕ in $L^2(\mathbb{T},\nu)$. But since for all $j \geq 0$, we have
$$\sum_{k=0}^{j} \mathrm{e}^{\mathrm{i}k\lambda} = \frac{1-\mathrm{e}^{-\mathrm{i}(j+1)\lambda}}{1-\mathrm{e}^{\mathrm{i}\lambda}}\ ,$$
it follows that
$$n^{-1}\sum_{j=0}^{n} g_j(\lambda) - \phi(\lambda) = n^{-1}\sum_{j=0}^{n}\sum_{k=0}^{j} \mathrm{e}^{-\mathrm{i}k\lambda} - (1-\mathrm{e}^{\mathrm{i}\lambda})^{-1} = -n^{-1}\sum_{j=0}^{n} \frac{\mathrm{e}^{-\mathrm{i}(j+1)\lambda}}{1-\mathrm{e}^{\mathrm{i}\lambda}}\ .$$
Since $f(0) = 0$ and f is continuouly differentiable, $f(\lambda)(1-\mathrm{e}^{\mathrm{i}\lambda})^{-1}$ is continuous on $\mathbb{T}$, Question 3 implies that
$$\lim_{n\to\infty}\int_{\mathbb{T}} \left| n^{-1}\sum_{j=0}^{n} g_j(\lambda) - \phi(\lambda)\right|^2 f(\lambda)\mathrm{d}\lambda = 0\ .$$
It comes that $n^{-1}\sum_{j=0}^{n} g_j(\lambda)$ converges to ϕ in $L^2(\mathbb{T},\nu)$. And by the spectral representation theorem, we have $n^{-1}\sum_{j=0}^{n} Y_j$ converges to Z in $L^2(\Omega)$.

6. Since
$$Y_t = \epsilon_0 - \epsilon_{-t}\ ,$$
we have that $n^{-1}\sum_{j=0}^{n} Y_j = \epsilon_0 - n^{-1}\sum_{k=-n-1}^{-1}\epsilon_k$, so converges to ϵ_0 in L^2.

Solution of Exercise 5.8 1. The proof is by induction. Since for all $j \geq 1$, Γ_j is the Gram matrix associated with the family $(X_1,\cdots,X_j)$ in $L^2(\Omega)$, if Γ_k is invertible then $(X_1,\cdots,X_k)$ is linearly independent. Also, if besides Γ_{k+1} is not invertible, it yields that $X_{k+1} \in \mathrm{Span}\,(X_1,\cdots,X_k)$, since this time $(X_1,\cdots,X_{k+1})$ is linearly dependent.
Assume now that for all $n \in \{k+1,\cdots,N\}$, for $N \geq k+1$, $X_n \in \mathrm{Span}\,(X_1,\cdots,X_k)$. Then, since the process is weakly stationary, Γ_k is also the Gram matrix associated with the family $(X_{N-k},\cdots,X_N)$ in $L^2(\Omega)$, if Γ_k is invertible then $(X_{N-k},\cdots,X_N)$ is linearly independent. Also, if besides Γ_{k+1} is not invertible, it yields that $X_{N+1} \in \mathrm{Span}\,(X_{N-k},\cdots,X_N)$, since this time $(X_{N+1-k},\cdots,X_N)$ is linearly dependent. Since by the induction hypothesis, $\mathrm{Span}\,(X_{N-k},\cdots,X_N) \subset \mathrm{Span}\,(X_1,\cdots,X_k)$, it follows that $X_{n+1} \in \mathrm{Span}\,(X_1,\cdots,X_k)$.

2. For all $j \geq 1$, we have
$$\gamma(0) = \mathbb{E}\left[|X_{j+k}|^2\right] = \sum_{p,q=1}^{k} \bar{\alpha}_p^{j+k}\mathbb{E}\left[\bar{X}_p X_q\right]\alpha_q^{j+k} = (\alpha^{j+k})^H\Gamma_k\alpha^{j+k} \geq \lambda_\star\|\alpha^{j+k}\|^2\ . \tag{29}$$
where $\lambda_\star = \min\{\lambda \mid \lambda \in \mathrm{Spec}\,\Gamma_k\}$. Moreover, since Γ_k is invertible, it is a positive definite matrix and $\lambda_\star$ is positive. Thus by (29), for all $j \geq 1$ $\|\alpha^{j+k}\|^2$ is uniformly bounded by $\gamma(0)/\lambda_\star$.

3. The proof is by contradiction. Assume that there exists $N \geq 1$ such that Γ_N is non invertible, and consider $k = \max\{j \in \{1, \cdots, N\} \mid \Gamma_k \text{ is invertible }\}$. Then by assumption, k is well defined, and by definition Γ_k is invertible, Γ_{k+1} is not. Now for all $j \geq 1$, we have
$$\gamma(0) = \mathbb{E}\left[|X_{j+k}|^2\right] = \sum_{q=1}^{k} \alpha_q^{j+k} \mathbb{E}\left[\overline{X_{j+k}} X_q\right] = \sum_{q=1}^{k} \alpha_q^{j+k} \mathbb{E}\left[\overline{X_{j+k}} X_q\right] = \sum_{q=1}^{k} \alpha_q^{j+k} \gamma(j+k-q) \,. \tag{30}$$
Since for all $j \geq 1$ $\|\alpha^{j+k}\|^2$ is uniformly bounded by $\gamma(0)/\lambda_\star$, $\sum_{q=1}^{k} \alpha_q^{j+k}\gamma(j+k-q)$ goes to 0 as $j \to \infty$. therefore by (30), this implies that $\gamma(0) = 0$ which is excluded by assumption.

4. Assume that X is linearly predictable, and consider $k = \min\{j \geq 0 \mid \Gamma_j \text{ is invertible }\}$. Then if $k = 0$, we have that X is the constant process equals to 0. Otherwise, the exercise leads to a contradiction.

Solution of Exercise 5.9 1. Since $(X_t)_{t\in\mathbb{Z}}$ and Y_0 are uncorrelated and the scalar product is continuous, we have that for all $t \in \mathbb{Z}$, that $\mathcal{H}_t^X$ is orthogonal to $\mathrm{Span}\,(Y_0) = \mathcal{H}_{-\infty}^Y = \mathcal{H}_t^Y$. Therefore since for all t, $\mathcal{H}_t^Z \subset \mathcal{H}_t^X \overset{\perp}{\oplus} \mathcal{H}_{-\infty}^Y$, we have $\mathcal{H}_{-\infty}^Z \subset \mathcal{H}_{-\infty}^X \overset{\perp}{\oplus} \mathcal{H}_{-\infty}^Y$. Finally, since $(X_t)_{t\in\mathbb{Z}}$ is a weak noise, it is purely non-deterministic so $\mathcal{H}_{-\infty}^X = \{0\}$ and the proof follows.

2. We have for all $z \in \mathbb{Z}$ and $n \geq 1$, $T_{t,n} = Y_0 + n^{-1}\sum_{k=1}^{n} X_{t-n}$. Since $(X_t)_{t\in\mathbb{Z}}$ is a weak noise, we have
$$\mathbb{E}\left[\|n^{-1}\sum_{k=1}^{n} X_{t-n}\|^2\right] = \mathrm{Var}\left(n^{-1}\sum_{k=1}^{n} X_{t-n}\right) = n^{-1}\sigma^2 \,,$$
and it yields that $\lim_{n\to\infty} T_{t,n} = Y_0$ for all $t \in \mathbb{Z}$.

3. From the previous question, we get that for all $t \in \mathbb{Z}$, $Y_0 \in \mathcal{H}_t^Z$ and therefore $Y_0 \in \mathcal{H}_{-\infty}^Z$. By the first question $\mathrm{Span}\,(Y_0) = \mathcal{H}_{-\infty}^Z$.

Solution of Exercise 5.10 1. Let $t \in \mathbb{Z}$. First for all $s \leq t$, we have
$$\mathcal{H}_s^X \overset{\perp}{\otimes} \mathcal{H}_{s\to t}^\epsilon = \mathcal{H}_t^X \,,$$
Second by Theorem 1.4.4, we have for all $h \in \mathcal{H}_t^X$,
$$\lim_{s\to-\infty} \mathrm{proj}\left(h \middle| \mathcal{H}_s^X\right) = \mathrm{proj}\left(h \middle| \mathcal{H}_{-\infty}^X\right) \,,$$
$$\lim_{s\to-\infty} \mathrm{proj}\left(h \middle| \mathcal{H}_{s\to t}^\epsilon\right) = \mathrm{proj}\left(h \middle| \mathcal{H}_t^\epsilon\right) \,,$$
and we have
$$\mathcal{H}_t^X = \mathcal{H}_{-\infty}^X + \mathcal{H}_t^\epsilon \,.$$
Now we show that this two spaces are orthogonal. For all $x \in \mathcal{H}_{-\infty}^X$, and for all $s \in \mathbb{Z}$, we have for all $y \in \mathcal{H}_{s\to t}^\epsilon$ by definition $x \perp y$. Therefore, since the scalar product is continuous and $\mathcal{H}_t^\epsilon = \overline{\bigcup_{s=-\infty}^{t} \mathcal{H}_{s\to t}^\epsilon}$n it follows that for all $x \in \mathcal{H}_{-\infty}^X$, $x \perp y$ for all $y \in \mathcal{H}_t^\epsilon$.

2. By definition $U_t = \mathrm{proj}\left(X_t \middle| \mathcal{H}_t^\epsilon\right)$, therefore we have $V_t = \mathrm{proj}\left(X_t \middle| \mathcal{H}_{-\infty}^X\right)$. It follows then that $(V_t)_{t\in\mathbb{Z}}$ and $(U_t)_{t\in\mathbb{Z}}$ are uncorrelated since for all $t \in \mathbb{Z}$, $\mathcal{H}_t^\epsilon \perp \mathcal{H}_{-\infty}^X$.

3. By the previous question we have for all $t \in \mathbb{Z}$, $\mathcal{H}_t^U \subset \mathcal{H}_t^\epsilon$ and $\mathcal{H}_t^V \subset \mathcal{H}_{-\infty}^X$. Next, we have by the previous question again,
$$\mathcal{H}_t^\epsilon \overset{\perp}{\otimes} \mathcal{H}_{-\infty}^X = \mathcal{H}_t^X \subset \mathcal{H}_t^U \overset{\perp}{\otimes} \mathcal{H}_t^V \,.$$
This implies that $\mathcal{H}_t^U = \mathcal{H}_t^\epsilon$ and $\mathcal{H}_t^V \subset \mathcal{H}_{-\infty}^X$.

4. Since for all $t \in \mathbb{Z}$, $\mathcal{H}_t^U = \mathcal{H}_t^\epsilon$ and $(\epsilon_t)_{t\in\mathbb{Z}}$ is a weak noise, so is purely non-deterministic, we get that $(U_t)_{t\geq 0}$ is also purely non-determistic, this means

$$\bigcap_{t\in\mathbb{Z}} \mathcal{H}_t^U = \{0\} \; .$$

V is deterministic since for all $t \in \mathbb{Z}$, $s \leq t$, we have

$$\mathrm{proj}\left(V_t \middle| \mathcal{H}_s^V \right) = \mathrm{proj}\left(V_t \middle| \mathcal{H}_{-\infty}^X \right) = V_t \; ,$$

so the innovation process associated with $(V_t)_{t\in\mathbb{Z}}$ is equals to 0 in L^2.

Solution of Exercise 7.1 The first two points are straightforward. To prove (iii), it is easy to show that any $f : \mathsf{Y} \times \mathsf{Z} \times \mathsf{U} \to \mathbb{R}_+$,

$$(M \otimes N) \otimes P(f) = M \otimes (N \otimes P)(f) = \int \left(\int (f(y,z,u)\ P(z, \mathrm{d}u))\ N(y, \mathrm{d}z) \right)\ M(x, \mathrm{d}y)\ .$$

Solution of Exercise 7.2

1. Suppose that $(X_k)_{k \geq 0}$ are i.i.d. and denote by μ the common distribution of this sequence. Then for all $f : \mathbb{R} \to \mathbb{R}$, measurable and bounded, we have since X_0 and X_1 are independent

$$\mathbb{E}\left[f(X_1) \middle| X_0 \right] = \mathbb{E}\left[f(X_1) \right]\ . \tag{31}$$

On the other hand, we have

$$\begin{aligned} \mathbb{E}\left[f(X_1) \middle| X_0 \right] &= \int_{\mathsf{X}} f(x_1) P(X_0, \mathrm{d}x_1) \\ \mathbb{E}\left[f(X_1) \right] &= \int_{\mathsf{X}} f(x_1) \mu(\mathrm{d}x_1)\ . \end{aligned}$$

Using (31), it follows that μ almost surely, $P(X_0, \cdot) = \mu(\cdot)$. The converse is easy and left to the reader.

2. By the monotone class theorem, we can assume that $A \in \mathcal{X} \times \mathcal{Y}$ is on the form $A = B \times C$ with $B \in \mathcal{X}$ and $C \in \mathcal{Y}$. Also by assumption, for all $k \geq 0$ we have

$$\begin{aligned} &\mathbb{P}\left[X_{k+1} \in B\ \ Y_{k+1} \in C \,\middle|\, \mathcal{F}_k \right] \\ &\qquad = \mathbb{E}\left[\mathbb{1}_{\{X_{k+1} \in B\}} \mathbb{E}\left[\mathbb{1}_{\{\ Y_{k+1} \in C\}} \middle| \mathcal{F}_k \vee \sigma(X_{k+1}) \right] \middle| \mathcal{F}_k \right] \\ &\qquad = \mathbb{E}\left[\mathbb{1}_{\{X_{k+1} \in B\}} Q(X_{k+1}, C) \middle| \mathcal{F}_k \right] \\ &\qquad = \int_{\mathsf{X}} \mathbb{1}_{\{x_{k+1} \in B\}} Q(x_{k+1}, C) P(x_k, \mathrm{d}x_{k+1}) \\ &\qquad = \int_B \int_C Q(x_{k+1}, \mathrm{d}y) P(x_k, \mathrm{d}x_{k+1})\ . \end{aligned}$$

Therefore, it follows that (X_k, Y_k) is a Markov chain with kernel $P \otimes Q$ on $\mathsf{X} \times (\mathcal{X} \otimes \mathcal{Y})$.

3. In the case where $\mathsf{X} = \mathsf{Y} = \{0, 1\}$, then P (and Q) can be written on the form of a matrix

$$\begin{pmatrix} a & 1-a \\ 1-b & b \end{pmatrix}$$

with $a, b \in [0, 1]$.

4. Using the second question, we have for all $k \geq 1$,

$$\begin{aligned} \mathbb{E}\left[Y_k \middle| X_{k-1}, Y_{k-1} \right] &= \mathbb{P}\left[Y_k = 1 \,\middle|\, X_{k-1}, Y_{k-1} \right] \\ &= \mathbb{P}\left[X_k \in \{0,1\}, Y_k = 1 \,\middle|\, X_{k-1}, Y_{k-1} \right] \\ &= \sum_{x=0}^{1} p(X_{k-1}, x) q(x, 1) = pq(X_{k-1}, 1)\ . \end{aligned}$$

5. We have by definition that $Y_0 \sim q(X_0, \cdot)$, therefore for $k \in \{0, 1\}$

$$\mathbb{P}\left(Y_0 = k \right) = \sum_{m=0}^{1} \mu_X(m) q(m, k)\ . \tag{32}$$

Therefore for all $i, k \in \{0,1\}$,

$$\mathbb{P}(X_0 = i \mid Y_0 = k) = \frac{\mathbb{P}(X_0 = i\ ,\ Y_0 = k)}{\mathbb{P}(Y_0 = k)} = \frac{\mu_X(i)q(i,k)}{\sum_{m=0}^{1}\mu_X(m)q(m,k)} . \tag{33}$$

Since

$$\mathbb{E}\left[pq(X_0,\cdot)\middle|\, Y_0\right] = \sum_{i=0}^{1} pq(i,\cdot)\mathbb{P}(X_0 = i \mid Y_0) .$$

we have that the law of $\mathbb{E}\left[pq(X_0,\cdot)\middle|\, Y_0\right]$ is $\mathbb{P}(Y_0 = 0)\,\delta_{a_0} + \mathbb{P}(Y_0 = 1)\,\delta_{a_1}$ where

$$a_0 = \sum_{i=0}^{1} pq(i,\cdot)\mathbb{P}(X_0 = i \mid Y_0 = 0)\ , a_1 = \sum_{i=0}^{1} pq(i,\cdot)\mathbb{P}(X_0 = i \mid Y_0 = 1) .$$

Using (32) and (33), we have that the law of $\mathbb{E}\left[pq(X_0,\cdot)\middle|\, Y_0\right]$ is

$$\sum_{m=0}^{1} \mu_X(m)q(m,0)\delta_{a_0} + \sum_{m=0}^{1} \mu_X(m)q(m,1)\delta_{a_1}$$

where

$$a_0 = \sum_{i=0}^{1} pq(i,\cdot)\frac{\mu_X(i)q(i,0)}{\sum_{m=0}^{1}\mu_X(m)q(m,0)}\ , a_1 = \sum_{i=0}^{1} pq(i,\cdot)\frac{\mu_X(i)q(i,0)}{\sum_{m=0}^{1}\mu_X(m)q(m,1)} .$$

6. The law of $pq(X_0,\cdot)$ is simply

$$\mu_X(0)\delta_{pq(0,\cdot)} + \mu_X(1)\delta_{pq(1,\cdot)} .$$

7. On the one hand, we have by the question 2 and 4 that for all $k \geq 0$,

$$\mathbb{P}\left[Y_{k+1} \in \cdot \,\middle|\, \mathcal{F}_k^Z\right] = \mathbb{P}\left[Y_{k+1} \in \cdot \,\middle|\, X_k, Y_k\right] = pq(X_k,\cdot) . \tag{34}$$

On the other hand, we have

$$\mathbb{P}\left[Y_{k+1} \in \cdot \,\middle|\, Y_k\right] = \mathbb{E}\left[\mathbb{P}\left[Y_{k+1} \in \cdot \,\middle|\, X_k, Y_k\right]\middle|\, Y_k\right] = \mathbb{E}\left[pq(X_k,\cdot)\middle|\, Y_k\right] . \tag{35}$$

For $(Y_k, \mathcal{F}_k^Z)_{k\in\mathbb{N}}$ to be a Markov chain, this sequence must satisfy for all $k \geq 0$,

$$\mathbb{P}\left[Y_{k+1} \in \cdot \,\middle|\, \mathcal{F}_k^Z\right] = \mathbb{P}\left[Y_{k+1} \in \cdot \,\middle|\, Y_k\right] .$$

So by (34) and (35),

$$\mathbb{E}\left[pq(X_k,\cdot)\middle|\, Y_k\right] = pq(X_k,\cdot) ,$$

8. A simple counterexample by the previous question is provided is $\mathbb{P}(Y_0 = 0) \neq \mathbb{P}(X_0 = 0)$ and $\mathbb{P}(Y_0 = 0) \neq \mathbb{P}(X_0 = 1)$.

Solution of Exercise 7.3 1. We have by definition:

$$F(x,x') = P(x,(-\infty,x']) = \mathbb{P}_x(X_1 \leq x') .$$

2. Denote for all $x, x' \in \mathbb{R}$ and $q \in \mathbb{Q}$ by $A(x,x',q) = F(x,q)\mathbb{1}_{\{x' \leq q\}} + \mathbb{1}_{\{x' > q\}}$. For all $q \in \mathbb{Q}$, we have $F(x,x') \leq A(x,x',q)$, and thus

$$F(x,x') \leq \inf_{q\in\mathbb{Q}} A(x,x',q) . \tag{36}$$

First if $F(x,x') = 1$. Since for all $q \in \mathbb{Q}$, $A(x,x',q) \leq 1$. we have trivially $F(x,x') = \inf_{q\in\mathbb{Q}} A(x,x',q)$ Now assume that $F(x,x') < 1$. Since let $(q_n)_{n\in\mathbb{N}}$ be a sequence of rational number converging to x' and greater than x'. Then since $y \to F(x,y)$ is right-continuous, we have $\lim_{n\to+\infty} F(x,q_n) = F(x,x')$ and therefore,

$$F(x,x') = \lim_{n\to+\infty} A(x,x',q_n) .$$

The proof is concluded using this result and (36).

3. The set $\{(x, x') \to A(x, x', q) \| \, q \in \mathbb{Q}\}$ is a countable set of Borel functions on $\mathbb{R}^2$. Therefore, the infimum of this set is also a Borel function. Therefore by the question 2, F is a Borel function.

Solution of Exercise 7.4 Let $f : \mathbb{N} \to \mathbb{R}$ be a bounded function. Then for all $k \geq 0$, we have

$$\begin{aligned}\mathbb{E}\left[\, f(X_{k+1}) \middle|\, \mathcal{F}_k^X\right] &= \mathbb{E}\left[\, \mathbb{1}_{\{X_k \geq 1\}} f(X_k - 1) + \mathbb{1}_{\{X_k = 0\}} f(Z_k + 1) \middle|\, \mathcal{F}_k^X\right] \\ &= \mathbb{1}_{\{X_k \geq 1\}} f(X_k - 1) \mathbb{1}_{\{X_k = 0\}} \sum_{z \in \mathbb{N}} f(z + 1) \mu_z \ ,\end{aligned}$$

where we have used that $(Z_k)_{k \geq 0}$ is independent of X_0 for the last line, so by clear induction Z_k is independent of $\mathcal{F}_k^X$. Therefore, $(X_n)_{n \geq 0}$ is a homogeneous Markov chain with transition matrix given by

$$M(x, A) = \mathbb{1}_{\{x \geq 1\}} \delta_{x-1}(A) + \mathbb{1}_{\{x=0\}} \sum_{z \in A} \mu_z \ .$$

Solution of Exercise 7.5 1. Let $A \in \mathcal{B}(\mathbb{R})$. Then for all $k \geq 0$,

$$\begin{aligned}\mathbb{P}\left[X_{k+1} \in A \,|\, \mathcal{F}_k^X\right] &= \mathbb{P}\left[\, f(X_k) + g(X_k) Z_k \in A \,|\, \mathcal{F}_k^X\right] \\ &= \int_{\mathbb{R}} \mathbb{1}_{\{f(X_k) + g(X_k) z \in A\}} \mu(\mathrm{d}z) \ ,\end{aligned}$$

where we used that $(Z_k)_{k \geq 0}$ is independent of X_0, so Z_k is independent of $\mathcal{F}_k^X$ and Exercise 2.3 (k). Therefore $(X_n)_{n \geq 0}$ is a homogeneous Markov chain in $\mathbb{R}$ with Markov kernel given by for all $x \in \mathbb{R}$ and $A \in \mathcal{B}(\mathbb{R})$

$$Q(x, A) = \int_{\mathbb{R}} \mathbb{1}_{\{f(x) + g(x) z \in A\}} \mu(\mathrm{d}z)$$

2. If μ admits a density ϕ with respect to the Lebesgue measure $\mathrm{d}z$, we have using a change of variable,

$$Q(x, A) = \int_{\mathbb{R}} \mathbb{1}_{\{f(x) + g(x) z \in A\}} \phi(z) \mathrm{d}z = \int_A \phi((z - f(x)) g(x)^{-1}) \mathrm{d}z$$

Solution of Exercise 8.1 1. Let us verify that λ is a probability measure. First, taking $A = E$ or $A = \emptyset$, it follows that $\lambda(E) = 1, \lambda(\emptyset) = 0$. For all $A \in \mathcal{F}$, $\lambda(A) = \lim_{n\to\infty} \mu P^n(A) \geq 0$. Also, let $(A_k)_{k\geq 0}$ be pairwise disjoint sets in $\mathcal{F}$. First, for all $p \geq 0$ we have by (5)

$$\lambda\left(\cup_{k=0}^{p} A_k\right) = \lim_{n\to\infty} \mu P^n\left(\cup_{k=0}^{p} A_k\right) = \sum_{k=0}^{p} \lim_{n\to\infty} \mu P^n\left(A_k\right) = \sum_{k=0}^{p} \lambda\left(A_k\right) \ . \tag{37}$$

Then for all $p \geq 0$ and all $n \geq 0$

$$\begin{aligned}
\left|\lambda\left(\cup_{k=0}^{\infty} A_k\right) - \sum_{k=0}^{p} \lambda\left(A_k\right)\right| &= \left|\lambda\left(\cup_{k=0}^{\infty} A_k\right) - \lambda\left(\cup_{k=0}^{p} A_k\right)\right| \\
&\leq \left|\lambda\left(\cup_{k=0}^{\infty} A_k\right) - \mu P^n\left(\cup_{k=0}^{\infty} A_k\right)\right| + \left|\mu P^n\left(\cup_{k=0}^{\infty} A_k\right) - \mu P^n\left(\cup_{k=0}^{p} A_k\right)\right| \\
&\qquad + \left|\mu P^n\left(\cup_{k=0}^{p} A_k\right) - \lambda\left(\cup_{k=0}^{p} A_k\right)\right| \\
&\leq 2\sup_{A\in\mathcal{F}} \left|\mu P^n(A) - \lambda(A)\right| + \left|\mu P^n\left(\cup_{k=0}^{\infty} A_k\right) - \mu P^n\left(\cup_{k=0}^{p} A_k\right)\right| \ .
\end{aligned}$$

Also for $\epsilon > 0$, by (5) there exists $n \geq 0$ such that the first term in the right hand side is smaller than ϵ. Besides, for this fixed n, since μP^n is a measure, there exists $N \geq 0$ such that for all $p \geq N$, the second term is smaller than ϵ. Therefore for all $p \geq N$, $\left|\lambda\left(\cup_{k=0}^{\infty} A_k\right) - \sum_{k=0}^{p} \lambda\left(A_k\right)\right| \leq \epsilon$. Therefore,

$$\lambda\left(\cup_{k=0}^{\infty} A_k\right) = \sum_{k=0}^{\infty} \lambda\left(A_k\right) \ ,$$

which shows the σ-additivity of λ and concludes that λ is a probability measure.

2. For all $A \in \mathcal{F}$, we have taking $g(x) = P(x, A)$ in (6)

$$\lambda P(A) = \int_E P(x, A)\lambda(\mathrm{d}x) = \lim_{n\to\infty} \int_E P(x, A)\mu P^n(\mathrm{d}x) = \lim_{n\to\infty} \mu P^{n+1}(A) = \lambda(A) \ ,$$

which shows that λ is invariant for P.

3. Assume that P admits two invariant probability measures λ_1 and λ_2. Then for all $A \in \mathcal{F}$, we have
$$\lambda_1(A) = \lim_{n\to\infty} \lambda_1 P^n(A) = \lambda(A) = \lim_{n\to\infty} \lambda_2 P^n(A) = \lambda_2(A) \ .$$
So $\lambda_1 = \lambda_2 = \lambda$.

Solution of Exercise 8.2 1. Since for all continuous and bounded function f, Pf is also continuous and bounded, we have for all such function

$$\begin{aligned}
\int_{\mathbb{R}^d} f(x)\lambda_\mu(\mathrm{d}x) = \lim_{k\to\infty} \int_{\mathbb{R}^d} f(x)\mu P^{k+1}(\mathrm{d}x) = \lim_{k\to\infty} \int_{\mathbb{R}^d} Pf(x)\mu P^k(\mathrm{d}x) \\
= \int_{\mathbb{R}^d} Pf(x)\lambda_\mu(\mathrm{d}x) \ .
\end{aligned} \tag{38}$$

Therefore for all continuous and bounded function f,

$$\int_{\mathbb{R}^d} f(x)\lambda_\mu(\mathrm{d}x) = \int_{\mathbb{R}^d} f(x)\lambda_\mu P(\mathrm{d}x) \ ,$$

therefore $\lambda_\mu = \lambda_\mu P$.

2. If P admits a unique probability measure λ, then $\lambda_\mu = \lambda$.

3. Assume that P admits two invariant probability measure λ_1 and λ_2. Then, for all bounded and continuous function f we have

$$\int_{\mathbb{R}^d} f(x)\lambda_1(\mathrm{d}x) = \lim_{n\to\infty}\int_{\mathbb{R}^d} f(x)\lambda_1 P^n(\mathrm{d}x)$$
$$= \int_{\mathbb{R}^d} f(x)\lambda(\mathrm{d}x) = \lim_{n\to\infty}\int_{\mathbb{R}^d} f(x)\lambda_2 P^n(\mathrm{d}x) = \int_{\mathbb{R}^d} f(x)\lambda_2(\mathrm{d}x)\ .$$

Therefore, $\lambda_1 = \lambda_2 = \lambda$.

Solution of Exercise 8.3 1. Let $p \geq 1$. If $\sigma_y^{(p-1)} = \infty$, then the relation clearly holds. Otherwise,

$$\sigma_y^{(0)} \circ \mathrm{S}^{\sigma_y^{(p-1)}} = \inf\{n > 0\ ; \xi_{\sigma_y^{(p-1)}+n} = y\} = \inf\{k > \sigma_y^{(p-1)}\ ; \xi_k = y\} - \sigma_y^{(p-1)}\ .$$

2. By the first question, for all $p \geq 1$, $\sigma_p^{(y)}$ is the p-th return time to y. Therefore, for all $n \geq 0$, we have

$$\{\sigma_p^{(y)} \leq n\} = \left\{\sum_{i=1}^{n} \mathbb{1}_{\{\xi_i = y\}} \geq p\right\} \in \mathcal{F}_n^{\xi}\ ,$$

which implies that $\sigma_p^{(y)}$ is a stopping time.

3. Let $p \geq 1$, we need to show that $\mathcal{F}_{\tau_y^{(p)}} \subset \mathcal{F}_{\tau_y^{(p+1)}}$. Let $A \in \mathcal{F}_{\tau_y^{(p)}}$. Then for all $n \geq 0$, we have since $\tau_y^{(p+1)} \geq \tau_y^{(p)}$, $\{\tau_y^{(p+1)} \leq n\} \subset \{\tau_y^{(p)} \leq n\}$ so

$$A \cap \{\tau_y^{(p+1)} \leq n\} = A \cap \{\tau_y^{(p)} \leq n\} \cap \{\tau_y^{(p+1)} \leq n\}\ .$$

But since $A \in \mathcal{F}_{\tau_y^{(p)}}$, $A \cap \{\tau_y^{(p)} \leq n\} \in \mathcal{F}_n$. Moreover, since $\tau_y^{(p+1)}$ is a stopping time $\{\tau_y^{(p+1)} \leq n\} \in \mathcal{F}_n$. Then, if follows by definition of a σ-field that $A \cap \{\tau_y^{(p+1)} \leq n\} \in \mathcal{F}_n$ and it implies that $A \in \mathcal{F}_{\tau_y^{(p+1)}}$.

4. Let $f : \overline{\mathbb{N}^*} \to \mathbb{R}$ be a bounded function. Then for all $p \geq 2$, we have

$$\mathbb{E}\left[f(\sigma_y^{(p)})\middle|\mathcal{F}_{\sigma_y^{(p-1)}}\right] = \sum_{i=1}^{\infty}\mathbb{E}\left[\mathbb{1}_{\{\sigma_y^{(p-1)}=i\}} f(i + \sigma_y^{(0)} \circ \mathrm{S}^i)\middle|\mathcal{F}_{\sigma_y^{(p-1)}}\right] + f(\infty)\mathbb{1}_{\{\sigma_y^{(p-1)}=\infty\}}\ ,$$

where we used that $\mathbb{1}_{\{\sigma_y^{(p-1)}=\infty\}}$ is $\mathcal{F}_{\sigma_y^{(p-1)}}$-measurable. By the question 4 of exercise 1, it follows that

$$\mathbb{E}\left[f(\sigma_y^{(p)})\middle|\mathcal{F}_{\sigma_y^{(p-1)}}\right] = \sum_{i=1}^{\infty}\mathbb{1}_{\{\sigma_y^{(p-1)}=i\}}\mathbb{E}\left[f(i + \sigma_y^{(0)} \circ \mathrm{S}^i)\middle|\mathcal{F}_i\right] + f(\infty)\mathbb{1}_{\{\sigma_y^{(p-1)}=\infty\}}\ .$$

Using the simple Markov property, we get

$$\mathbb{E}\left[f(\sigma_y^{(p)})\middle|\mathcal{F}_{\sigma_y^{(p-1)}}\right] = \sum_{i=1}^{\infty}\mathbb{1}_{\{\sigma_y^{(p-1)}=i\}}\mathbb{E}_y\left[f(i + \sigma_y^{(0)})\right] + f(\infty)\mathbb{1}_{\{\sigma_y^{(p-1)}=\infty\}}\ . \tag{39}$$

Therefore, $\mathbb{E}\left[f(\sigma_y^{(p)})\middle|\mathcal{F}_{\sigma_y^{(p-1)}}\right] = Kf(\sigma_{p-1}^{(y)})$ where K is Markov kernel on $\overline{\mathbb{N}}^*$ given by

$$K(i,j) = \begin{cases} \mathbb{P}_y\left(\sigma_y^{(0)} = j - i\right) = \mu_y(j-i) & \text{if } i,j < \infty \text{ and } j \geq i \\ 0 & \text{if } i,j \in \overline{\mathbb{N}} \text{ or } j < i \\ \mathbb{P}_y\left(\sigma_y^{(0)} = \infty\right) = \mu_y(\infty) & \text{if } i < \infty \text{ and } j = \infty \\ 1 & \text{if } i = j = \infty\ . \end{cases} \tag{40}$$

5. By (40), we have for all $p \geq 1$,

$$\mathbb{P}_y\left(\sigma_y^{(p+1)} < \infty \,\middle|\, \mathcal{F}_{\sigma_y^{(p)}}\right) = \sum_{j=0}^{\infty} K(\sigma_y^{(p)}, j) = \mathbb{1}_{\{\sigma_y^{(p)} < \infty\}} \sum_{j=0}^{\infty} \mathbb{P}_y\left(\sigma_y^{(0)} = j\right)$$
$$= \mathbb{1}_{\{\sigma_y^{(p)} < \infty\}} \mathbb{P}_y\left(\sigma_y^{(0)} < \infty\right) .$$

Taking the expectation, we get

$$\mathbb{P}_y\left(\sigma_y^{(p+1)} < \infty\right) = \mathbb{P}_y\left(\sigma_y^{(0)} < \infty\right) \mathbb{P}_y\left(\sigma_y^{(p)} < \infty\right) ,$$

and a easy induction concludes the proof.
Since $\{N_y = \infty\} = \bigcap_{p=1}^{\infty} \{\sigma_y^{(p)} < \infty\}$, we have

$$\mathbb{P}_y\left(N_y = \infty\right) = \mathbb{P}_y\left(\bigcap_{p=1}^{\infty} \{\sigma_y^{(p)} < \infty\}\right) = \lim_{p \to \infty} \downarrow \mathbb{P}_y\left(\sigma_y^{(p)} < \infty\right) = 1 .$$

6. Let $p \geq 1$ and $i_1, \cdots, i_p \in \mathbb{N}$. Successively conditioning on $\mathcal{F}_{\sigma_k^{(y)}}$ for $k = p, \cdots, 1$ and using the strong Markov property, we have

$$\begin{aligned}
&\mathbb{P}_y\left(\sigma_y^{(2)} - \sigma_y^{(0)} = i_1, \cdots, \sigma_y^{(p+1)} - \sigma_y^{(p)} = i_p\right) \\
&\qquad = \mathbb{E}_y\left[\mathbb{1}_{\{\sigma_y^{(2)} - \sigma_y^{(0)} = i_1, \cdots, \sigma_p^{(y)} - \sigma_{p-1}^{(y)} = i_p\}} \mathbb{P}_y\left(\sigma_{p+1}^{(y)} - \sigma_p^{(y)} = i_p \,\middle|\, \mathcal{F}_{\sigma_p^{(y)}}\right)\right] \\
&\qquad = \mathbb{E}_y\left[\mathbb{1}_{\{\sigma_y^{(2)} - \sigma_y^{(0)} = i_1, \cdots, \sigma_p^{(y)} - \sigma_{p-1}^{(y)} = i_p\}} \mathbb{P}_y\left(\sigma_y^{(0)} \circ \mathrm{S}^{\sigma_p^{(y)}} = i_p \,\middle|\, \mathcal{F}_{\sigma_p^{(y)}}\right)\right] \\
&\qquad = \mathbb{E}_y\left[\mathbb{1}_{\{\sigma_y^{(2)} - \sigma_y^{(0)} = i_1, \cdots, \sigma_p^{(y)} - \sigma_{p-1}^{(y)} = i_p\}} \mathbb{P}_y\left(\sigma_y^{(0)} = i_p\right)\right] \\
&\qquad = \mathbb{P}_y\left(\sigma_y^{(0)} = i_p\right) \mathbb{E}_y\left[\mathbb{1}_{\{\sigma_y^{(2)} - \sigma_y^{(0)} = i_1, \cdots, \sigma_p^{(y)} - \sigma_{p-1}^{(y)} = i_p\}}\right] \\
&\qquad = \cdots \\
&\qquad = \prod_{k=1}^{p} \mathbb{P}_y\left(\sigma_y^{(0)} = i_k\right) .
\end{aligned}$$

Therefore $(\sigma_y^{(p+1)} - \sigma_y^{(p)})_{p \geq 1}$ is i.i.d. , with common distribution μ_y.

Solution of Exercise 8.4 The answer of the two questions are straightforward for $m = 0$.

1. Let now $m \geq 1$ and $x, y \in E$, $x \neq y$. Note that $\mathbb{P}_x$ almost surely, $N_y = m$ is equivalent to $\sigma_m^{(y)} < \infty, \sigma_{m+1}^{(y)} = \infty$, which is equivalent to $\sigma_y^{(m)} < \infty, \sigma_y^{(m+1)} - \sigma_y^{(m)} = \infty$. Therefore, using that $\sigma_y^{(m)}$ is $\mathcal{F}_{\sigma_y^{(m)}}$-measurable and the strong Markov property, by conditioning on $\mathcal{F}_{\sigma_y^{(m)}}$, we get that

$$\begin{aligned}
\mathbb{P}_{P,x}\left(N_y = m\right) &= \mathbb{P}_{P,x}\left(\sigma_y^{(m)} < \infty, \sigma_y^{(m+1)} - \sigma_y^{(m)} = \infty\right) \\
&= \mathbb{E}_{P,x}\left[\mathbb{1}_{\{\sigma_y^{(m)} < \infty\}} \mathbb{P}_{P,x}\left(\sigma_0^{(y)} \circ \mathrm{S}^{\sigma_m^{(y)}} = \infty \| \, \mathcal{F}_{\sigma_m^{(y)}}\right)\right] \\
&= \mathbb{P}_{P,y}\left(\sigma_0^{(y)} = \infty\right) \mathbb{E}_{P,x}\left[\mathbb{1}_{\{\sigma_y^{(m)} < \infty\}}\right] \qquad (41)
\end{aligned}$$

It remains to study $\mathbb{P}_{P,x}(\mathbb{1}_{\{\sigma_y^{(m)} < \infty\}})$. But $\sigma_y^{(m)} < \infty$ is equivalent this time to $\sigma_y^{(m)} < \infty, \sigma_y^{(m)} - \sigma_y^{(m-1)} = \infty$, with the convention $\sigma_y^{(0)} = 0$. Also, using the same reasoning than before we have conditionning on

$$\begin{aligned}\mathbb{P}_{P,x}(\mathbb{1}_{\{\sigma_y^{(m)}<\infty\}}) &= \mathbb{P}_{P,x}\left(\sigma_y^{(m-1)} < \infty, \sigma_y^{(m)} - \sigma_y^{(m-1)} = \infty\right) \\ &= \mathbb{E}_{P,x}\left[\mathbb{1}_{\{\sigma_y^{(m-1)}<\infty\}} \mathbb{P}_{P,x}\left(\sigma_0^{(y)} \circ \mathrm{S}^{\sigma_{m-1}^{(y)}} = \infty \mid \mathcal{F}_{\sigma_{m-1}^{(y)}}\right)\right] \\ &= \mathbb{E}_{P,x}\left[\mathbb{1}_{\{\sigma_y^{(m)}<\infty\}} \mathbb{P}_{P,y}\left(\sigma_0^{(y)} < \infty\right)\right] .\end{aligned}$$

So, we get

$$\mathbb{P}_{P,x}(\mathbb{1}_{\{\sigma_y^{(m)}<\infty\}}) = \begin{cases}\mathbb{P}_{P,x}\left(\sigma_y^{(0)} < \infty\right) & \text{if } m = 1 \\ \mathbb{P}_{P,y}\left(\sigma_y^{(0)} < \infty\right) \mathbb{P}_{P,x}\left(\mathbb{1}_{\{\sigma_y^{(m-1)}<\infty\}}\right) & \text{otherwise}\end{cases} .$$

By an immediate induction, we have

$$\mathbb{P}_{P,x}\left(\sigma_y^{(0)} < \infty\right) = \mathbb{P}_{P,x}\left(\sigma_y^{(0)} < \infty\right)\left(\mathbb{P}_{P,y}\left(\sigma_y^{(0)} < \infty\right)\right)^{m-1} . \tag{42}$$

Combining (41) and (42) concludes the proof.

2. Let $m \geq 1$ and $x \in E$. Note that this time $\mathbb{P}_x$ almost surely, $N_x = m$ is equivalent to $\sigma_{m-1}^{(x)} < \infty, \sigma_m^{(x)} = \infty$, which is equivalent to $\sigma_x^{(m-1)} < \infty, \sigma_x^{(m)} - \sigma_x^{(m-1)} = \infty$. The proof is concluded following the same lines than the previous questions.

Solution of Exercise 8.5 1. A invariant probability measure for K on E is a two dimensional vector $\mu = [a_1, a_2]^T$ such that

$$[a_1, a_2]K = [a_1, a_2]^T \ , a_1 + a_2 = 1 .$$

Therefore, it suffices to show that this linear system has a unique solution for which $a_1 + a_2 = 1$. So, we need to solve

$$\begin{cases} a_1/4 + a_2/2 & = a_1 \\ 3a_1/4 + a_2/2 & = a_2 \\ a_1 + a_2 & = 1 \end{cases} \iff \begin{cases} a_1 & = 2/5 \\ a_2 & = 3/5 .\end{cases}$$

Therefore, K admits a unique invariant probability measure μ.

2. For all $i \in E$,

$$\mathbb{E}_{K,i}[F(\xi)] = i + \sum_{j=0}^{1} j\mathbb{P}_{K,i}(\xi_1 = j) .$$

So, we get

$$\mathbb{E}_{K,i}[F(\xi)] = i + (3/4)\delta_0(i) + (1/2)\delta_1(i) .$$

3. By the Markov property, we have for all $n \geq 1$ and $i \in E$

$$\mathbb{E}_{K,i}[F(\xi \circ \mathrm{S}^n) \mid \mathcal{F}_n^{\xi}] = \mathbb{E}_{K,\xi_n}[F(\xi)] = \xi_n + (3/4)\delta_0(\xi_n) + (1/2)\delta_1(\xi_n) .$$

4. We have for all $n \geq 1$ and $i \in E$ conditioning successively on $\mathcal{F}_k = \sigma(\xi_0, \cdots, \xi_k)$ for $k = (n-1), \cdots, 1$,

$$\begin{aligned}\mathbb{P}_{K,i}[\tau \geq n] = \mathbb{P}_{K,i}[\forall k \in \{1, \cdots, n\} \ , \ \xi_k = 0] &= \mathbb{E}_{K,i}[\mathbb{1}_{\{\forall k \in \{1,\cdots,(n-1)\} \ , \ \xi_k = 0\}} \mathbb{P}_{K,i}[\xi_n = 0 \mid \mathcal{F}_{n-1}]] \\ &= \mathbb{E}_{K,i}[\mathbb{1}_{\{\forall k \in \{1,\cdots,(n-1)\} \ , \ \xi_k = 0\}} \mathbb{P}_{K,i}[\xi_n = 0 \mid \xi_{n-1}]] \\ &= \mathbb{E}_{K,i}[\mathbb{1}_{\{\forall k \in \{1,\cdots,(n-1)\} \ , \ \xi_k = 0\}} \mathbb{P}_{K,\xi_{n-1}}[\xi_n = 0]] \\ &= (1/4)\mathbb{E}_{K,i}[\mathbb{1}_{\{\forall k \in \{1,\cdots,(n-1)\} \ , \ \xi_k = 0\}}] \\ &= \cdots \\ &= (1/4)^{n-1} .\end{aligned}$$

Therefore $\mathbb{P}_{K,i}[\tau < \infty] = 1$ and by the Markov property

$$\mathbb{E}_{K,i}[F(\xi \circ \mathrm{S}^\tau) \mid \mathcal{F}^\xi_\tau] = \mathbb{E}_{K,\xi_\tau}[F(\xi)] = 3/2 \, .$$

Solution of Exercise 8.6 1. We have for all $i \in E$,

$$\begin{aligned}\mathbb{E}_{K,i}[F(\xi)] &= \sum_{j,k=-1}^{1} (j-k)K(i,j)K(j,k) \\ &= (-1)K(i,-1)K(-1,0) - 2K(i,-1)K(-1,1) + K(i,0)K(0,-1) \\ &\qquad - 1K(i,0)K(0,-1) + 2K(i,1)K(1,-1) + K(i,1)K(1,0) \, .\end{aligned}$$

Therefore,

$$\mathbb{E}_{K,i}[F(\xi)] = \begin{cases} 3/16 & \text{if } i = -1 \\ -5/8 & \text{if } i = 0 \\ -1/12 & \text{if } i = 1 \end{cases}$$

2. By the Markov property, we have

$$\mathbb{E}_{K,i}[F(\xi \circ \mathrm{S}^n) \mid \mathcal{F}_n] = \mathbb{E}_{K,\xi_n}[F(\xi)] = \begin{cases} 3/16 & \text{if } \xi_n = -1 \\ -5/8 & \text{if } \xi_n = 0 \\ -1/12 & \text{if } \xi_n = 1 \end{cases}$$

3. By the Markov property, we have

$$\mathbb{E}_{K,i}[\mathbb{1}_{\{\tau<\infty\}} F(\xi \circ \mathrm{S}^n) \mid \mathcal{F}_\tau] = \mathbb{1}_{\{\tau<\infty\}} \mathbb{E}_{K,\xi_\tau}[F(\xi)] = -(1/12)\mathbb{1}_{\{\tau<\infty\}}$$

Solution of Exercise 8.7 1. By definition, for all $x \in E^{\mathbb{N}}$, we have $f(x) = 1$ if in the first three coordinates of x, there is exactly one 1, and $f(x) = 0$ otherwise. Therefore, for all $i \in E$, we have

$$\begin{aligned}\mathbb{E}_{K,i}[f(\xi)] &= \mathbb{P}_{K,i}(f(\xi) = 1) \\ &= \mathbb{P}_i\left(\xi_1 = 1, , \xi_2 \neq 1, \xi_3 \neq 1\right) + \mathbb{P}_i\left(\xi_1 \neq 1, , \xi_2 = 1, \xi_3 \neq 1\right) \\ &\qquad + \mathbb{P}_i\left(\xi_1 \neq 1, , \xi_2 \neq 1, \xi_3 = 1\right) \, .\end{aligned}$$

Also, using that for all $i_1, i_2, i_3 \in E$, $\mathbb{P}_i\left(\xi_1 = i_1, \xi_2 = i_2 1, \xi_3 = i_3\right) = K(i, i_1)K(i_1, i_2)K(i_2, i_3)$, we have

$$\begin{aligned}\mathbb{P}_i\left(\xi_1 = 1, , \xi_2 \neq 1, \xi_3 \neq 1\right) &= \sum_{i_2,i_3=2}^{3} K(i,1)K(1,i_2)K(i_2,i_3) \\ \mathbb{P}_i\left(\xi_1 \neq 1, , \xi_2 = 1, \xi_3 \neq 1\right) &= \sum_{i_1,i_3=2}^{3} K(i,i_1)K(i_1,1)K(1,i_3) \\ \mathbb{P}_i\left(\xi_1 \neq 1, , \xi_2 \neq 1, \xi_3 = 1\right) &= \sum_{i_1,i_2=2}^{3} K(i,i_1)K(i_1,i_2)K(i_2,1) \, .\end{aligned}$$

Using for example Python, we have

$$\mathbb{E}_{K,i}[f(\xi)] = \begin{cases} 3/4 & \text{if } i = 1 \\ 5/8 & \text{if } i = 2 \\ 5/8 & \text{if } i = 3 \, . \end{cases} \tag{43}$$

2. Using the Markov property, we have

$$\mathbb{E}_{K,i}[f(\xi \circ \mathrm{S}^n) \mid \mathcal{F}_n^\xi] = \mathbb{E}_{K,\xi_n}[f(\xi)] \ .$$

So,

$$\mathbb{E}_{K,i}[f(\xi)] = \begin{cases} 3/4 & \text{if } \xi_n = 1 \\ 5/8 & \text{if } \xi_n = 2 \\ 5/8 & \text{if } \xi_n = 3 \ . \end{cases}$$

3. Using the same reasoning as in Exercise 8.5, we have for all $i \in E$, $\mathbb{P}_{K,i}(\sigma < \infty) = 1$. We deduce from the strong Markov property that

$$\mathbb{E}_{K,i}[\mathbb{1}_{\{\tau<\infty\}} f(\xi \circ \mathrm{S}^n) \mid \mathcal{F}_\tau^\xi] = \mathbb{1}_{\{\tau<\infty\}} \mathbb{E}_{K,\xi_\tau}[f(\xi)] = \mathbb{1}_{\{\tau<\infty\}} 3/4 \ ,$$

where we have used (43) for the last inequality.

Solution of Exercise 6.1 Since X_n is a supermartingale, we get that

$$\forall n \in \mathbb{N}, \quad \mathbb{E}\left[X_{n+1} \middle| \mathcal{F}_n\right] \leq X_n,$$

so that $Y_n = X_n - \mathbb{E}\left[X_{n+1} \middle| \mathcal{F}_n\right]$ is a non-negative random variable with expectation equals to

$$\mathbb{E}\left[Y_n\right] = \mathbb{E}\left[X_n\right] - \mathbb{E}\left[\mathbb{E}\left[X_{n+1} \middle| \mathcal{F}_n\right]\right] = \mathbb{E}\left[X_n\right] - \mathbb{E}\left[X_{n+1}\right] = 0.$$

Therefore $Y_n = 0$ a.s., which means that $\mathbb{E}\left[X_{n+1} \middle| \mathcal{F}_n\right] = X_n$ a.s. and $(X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is a martingale.

Solution of Exercise 6.2 Define $\mathcal{F}_n = \sigma(\epsilon_0, \dots, \epsilon_n)$ so that $X_n = S_n^2 - T_n$ is adapted to $\mathcal{F}_n$.

$$\begin{aligned} X_{n+1} &= S_{n+1}^2 - T_{n+1} \\ &= (S_n + \epsilon_{n+1})^2 - T_n - \sigma_{n+1}^2 \\ &= S_n^2 - T_n + 2S_n\epsilon_{n+1} + \epsilon_{n+1}^2 - \sigma_{n+1}^2. \end{aligned}$$

Since $(\epsilon_n)_{n\in\mathbb{N}}$ is a sequence of independent random variables with mean zero and variances given by $\mathrm{Var}[\epsilon_n] = \sigma_n^2$, we have that

$$\begin{aligned} \mathbb{E}\left[S_n\epsilon_{n+1} \middle| \mathcal{F}_n\right] &= S_n\mathbb{E}\left[\epsilon_{n+1} \middle| \mathcal{F}_n\right] = S_n\mathbb{E}\left[\epsilon_{n+1}\right] = 0. \\ \mathbb{E}\left[\epsilon_{n+1}^2 \middle| \mathcal{F}_n\right] &= \mathbb{E}\left[\epsilon_{n+1}^2\right] = \sigma_{n+1}^2. \end{aligned}$$

Finally we get that $\mathbb{E}\left[X_{n+1} \middle| \mathcal{F}_n\right] = S_n^2 - T_n = X_n$ so that $(X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is a martingale.

Solution of Exercise 6.3

1. For $m \leq n$, since X_m is $\mathcal{F}_m$-mesurable, we have

$$\mathbb{E}\left[X_mY_n \middle| \mathcal{F}_m\right] = X_m\mathbb{E}\left[Y_n \middle| \mathcal{F}_m\right] = X_mY_m.$$

2. For $1 \leq k \leq n$, we have thanks to the previous question and by symmetry of X_k and Y_k,

$$\mathbb{E}\left[X_{k-1}Y_{k-1}\right] = \mathbb{E}\left[\mathbb{E}\left[X_{k-1}Y_k \middle| \mathcal{F}_{k-1}\right]\right] = \mathbb{E}\left[X_{k-1}Y_k\right] = \mathbb{E}\left[X_kY_{k-1}\right],$$

so we can develop the term in the sum,

$$\begin{aligned} \mathbb{E}\left[(X_k - X_{k-1})(Y_k - Y_{k-1})\right] &= \mathbb{E}\left[X_kY_k\right] - \mathbb{E}\left[X_kY_{k-1}\right] - \mathbb{E}\left[X_{k-1}Y_k\right] + \mathbb{E}\left[X_{k-1}Y_{k-1}\right] \\ &= \mathbb{E}\left[X_kY_k\right] - 2\mathbb{E}\left[X_{k-1}Y_{k-1}\right] + \mathbb{E}\left[X_{k-1}Y_{k-1}\right] \\ &= \mathbb{E}\left[X_kY_k\right] - \mathbb{E}\left[X_{k-1}Y_{k-1}\right]. \end{aligned}$$

So we get a telescoping sum

$$\sum_{k=1}^n \mathbb{E}\left[(X_k - X_{k-1})(Y_k - Y_{k-1})\right] = \sum_{k=1}^n \mathbb{E}\left[X_kY_k\right] - \mathbb{E}\left[X_{k-1}Y_{k-1}\right] = \mathbb{E}\left[X_nY_n\right] - \mathbb{E}\left[X_0Y_0\right].$$

Solution of Exercise 6.4

1. (Make a drawing and think of convexity) For $t > 0$ the function $g : x \mapsto \mathrm{e}^{tx}$ is convex and we can write x as the weighted sum of the points c and $-c$ as

$$x = \frac{x+c}{2c}c + \frac{c-x}{2c}(-c), \qquad \frac{x+c}{2c} + \frac{c-x}{2c} = 1$$

Using the convex inequality we get,

$$\begin{aligned} g(x) &= g\left(\frac{x+c}{2c}c + \frac{c-x}{2c}(-c)\right) \\ &\leq \frac{x+c}{2c}g(c) + \frac{c-x}{2c}g(-c) \\ &= \frac{g(c)+g(-c)}{2} + \frac{x}{2c}(g(c) - g(-c)) \\ \mathrm{e}^{tx} &\leq \frac{\mathrm{e}^{ct} + \mathrm{e}^{-ct}}{2} + \frac{\mathrm{e}^{ct} - \mathrm{e}^{-ct}}{2c}x. \end{aligned}$$

2. For a fixed $t > 0$ and $-c \leq x \leq c$, we have $\mathrm{e}^{tx} \leq f_c(x)$. Since the random variable Z takes value in $[-c, c]$, we have $\mathrm{e}^{tZ} \leq f_c(Z)$ with e^{tZ} and $f_c(Z)$ both in L^1, i.e.,

$$\mathrm{e}^{tZ} \leq \frac{\mathrm{e}^{ct} + \mathrm{e}^{-ct}}{2} + \frac{\mathrm{e}^{ct} - \mathrm{e}^{-ct}}{2c} Z.$$

   Taking the conditional expectation lead to

$$\mathbb{E}\left[\mathrm{e}^{tZ} \middle| \mathcal{F}\right] \leq \frac{\mathrm{e}^{ct} + \mathrm{e}^{-ct}}{2} + \frac{\mathrm{e}^{ct} - \mathrm{e}^{-ct}}{2c} \mathbb{E}\left[Z \middle| \mathcal{F}\right] = f_c(\mathbb{E}\left[Z \middle| \mathcal{F}\right]).$$

3. Denote $S_m = Y_1 + \ldots + Y_m$ and use Markov inequality for a fixed $t > 0$ and any $a \in \mathbb{R}$,

$$\mathbb{P}\left(S_m \geq a\right) = \mathbb{P}\left(\mathrm{e}^{tS_m} \geq \mathrm{e}^{ta}\right) \leq \mathrm{e}^{-ta}\mathbb{E}\left[\mathrm{e}^{tS_m}\right].$$

4. $Y_n = X_n - X_{n-1}$ is such that $|Y_n| \leq c_n$ and since $\left((X_n, \mathcal{F}_n)\right)_{n\in\mathbb{N}}$ is a martingale we have $\mathbb{E}\left[Y_n \middle| \mathcal{F}_{n-1}\right] = 0$, so we can apply the result of question 2. with $\mathcal{F} = \mathcal{F}_{n-1}$ and write

$$\mathbb{E}\left[\mathrm{e}^{tY_n} \middle| \mathcal{F}_{n-1}\right] \leq f_{c_n}(\mathbb{E}\left[Y_n \middle| \mathcal{F}_{n-1}\right]) = f_{c_n}(0).$$

   Notice that we need to compare the hyperbolic cosine with an exponential

$$f_{c_n}(0) = \cosh(c_n t) = \sum_{k=0}^{\infty} \frac{(c_n t)^{2k}}{(2k)!}, \qquad \mathrm{e}^{(c_n t)^2/2} = \sum_{k=0}^{\infty} \frac{(c_n t)^{2k}}{2^k (k!)}$$

   We can prove by induction on $n \in \mathbb{N}$ that $(2n)! \geq 2^n(n!)$ or simply notice that it is true for $n = 0$ and for $n \geq 1$,

$$(2n)! = (2n)(2n-1)\ldots(n+1)(n!) \geq 2^n(n!),$$

   so we finally get that $f_{c_n}(0) \leq \mathrm{e}^{(c_n t)^2/2}$.

5. We combine the results of the two previous questions, on the one hand we have a telescoping sum,

$$\mathbb{P}\left(X_m - X_0 \geq a\right) = \mathbb{P}\left(Y_1 + \ldots + Y_m \geq a\right) \leq \mathrm{e}^{-ta}\mathbb{E}\left[\mathrm{e}^{tY_1 + \ldots + tY_m}\right].$$

   On the other hand we have for all $1 \leq k \leq m$,

$$\mathbb{E}\left[\mathrm{e}^{tY_k} \middle| \mathcal{F}_{k-1}\right] \leq \mathrm{e}^{(c_k t)^2/2},$$

   so we need to recursively take the conditional expectation,

$$\begin{aligned}
\mathbb{E}\left[\mathrm{e}^{tS_m}\right] &= \mathbb{E}\left[\mathrm{e}^{tS_{m-1}}\mathrm{e}^{tY_m}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\mathrm{e}^{tS_{m-1}}\mathrm{e}^{tY_m} \middle| \mathcal{F}_{m-1}\right]\right] \\
&= \mathbb{E}\left[\mathrm{e}^{tS_{m-1}}\mathbb{E}\left[\mathrm{e}^{tY_m} \middle| \mathcal{F}_{m-1}\right]\right] \\
&\leq \mathrm{e}^{(c_n t)^2/2}\mathbb{E}\left[\mathrm{e}^{tS_{m-1}}\right] \\
&\leq \ldots \\
\mathbb{E}\left[\mathrm{e}^{tS_m}\right] &\leq \exp\left(\frac{t^2}{2}\sum_{k=1}^{m} c_k^2\right).
\end{aligned}$$

   Finally we obtain

$$\mathbb{P}\left(X_m - X_0 \geq a\right) \leq \exp\left(-at + \frac{t^2}{2}\sum_{k=1}^{m} c_k^2\right).$$

6. The term inside the exponential is a polynomial $P(t) = \frac{1}{2}\left(\sum_{k=1}^m c_k^2\right) t^2 - at$ whose minimal value is attained for $t_m = a/\sum_{k=1}^m c_k^2$ and is equal to

$$P(t_m) = -\frac{a^2}{2\sum_{k=1}^m c_k^2},$$

so we get the Azuma inequality,

$$\mathbb{P}\left(X_m - X_0 \geq a\right) \leq \exp\left(-\frac{a^2}{2\sum_{k=1}^m c_k^2}\right).$$

Solution of Exercise 6.5

1. Φ admits a left (and a right) derivative $\Phi'_<$ at any point $x \in (a,b)$, and, for all $y \in (a,b)$,

$$\Phi(y) \geq \Phi'_<(x)\,(y-x) + \Phi(x). \tag{44}$$

Since X is a L^1 random variable valued in $]a,\ b[$, we get $(y = X, x = \mathbb{E}\left[X\right])$

$$\Phi(X) \geq \Phi'_<(\mathbb{E}\left[X\right])\,(X - \mathbb{E}\left[X\right]) + \Phi(\mathbb{E}\left[X\right])$$

Taking the expectation leads to $\mathbb{E}\left[\Phi(X)\right] \geq \Phi(\mathbb{E}\left[X\right])$.

2. Taking $y = X, x = \mathbb{E}\left[X|\,\mathcal{G}\right]$ in (44) gives

$$\Phi(X) \geq \Phi(\mathbb{E}\left[X|\,\mathcal{G}\right]) + \Phi'_<(\mathbb{E}\left[X|\,\mathcal{G}\right])\,(X - \mathbb{E}\left[X|\,\mathcal{G}\right]).$$

For any given $n \geq 1$, define $A_n = \{|\Phi(\mathbb{E}[X \mid \mathcal{G}])| \vee |\Phi'(\mathbb{E}[X \mid \mathcal{G}])| \leq n\} \in \mathcal{G}$, we have

$$\begin{aligned}
\mathbb{E}\left[\mathbb{1}_{A_n}\Phi(X)|\,\mathcal{G}\right] &\geq \mathbb{E}\left[\mathbb{1}_{A_n}\left(\Phi(\mathbb{E}\left[X|\,\mathcal{G}\right]) + \Phi'_<(\mathbb{E}\left[X|\,\mathcal{G}\right])\,(X - \mathbb{E}\left[X|\,\mathcal{G}\right])\right)|\,\mathcal{G}\right]\\
\mathbb{1}_{A_n}\mathbb{E}\left[\Phi(X)|\,\mathcal{G}\right] &\geq \mathbb{E}\left[\mathbb{1}_{A_n}\Phi(\mathbb{E}\left[X|\,\mathcal{G}\right])|\,\mathcal{G}\right] + \mathbb{E}\left[\mathbb{1}_{A_n}\Phi'_<(\mathbb{E}\left[X|\,\mathcal{G}\right])\,(X - \mathbb{E}\left[X|\,\mathcal{G}\right])|\,\mathcal{G}\right]\\
\mathbb{1}_{A_n}\mathbb{E}\left[\Phi(X)|\,\mathcal{G}\right] &\geq \mathbb{1}_{A_n}\Phi(\mathbb{E}\left[X|\,\mathcal{G}\right]) + \mathbb{1}_{A_n}\Phi'_<(\mathbb{E}\left[X|\,\mathcal{G}\right])\,\underbrace{\mathbb{E}\left[X - \mathbb{E}\left[X|\,\mathcal{G}\right]|\,\mathcal{G}\right]}_{=0}
\end{aligned}$$

and finally

$$\mathbb{1}_{A_n}\mathbb{E}\left[\Phi(X)|\,\mathcal{G}\right] \geq \mathbb{1}_{A_n}\Phi(\mathbb{E}\left[X|\,\mathcal{G}\right]).$$

Using that $\bigcup\limits_{n\in\mathbb{N}} A_n = \Omega$ we can conclude

$$\Phi(\mathbb{E}[X \mid \mathcal{G}]) \leq \mathbb{E}[\Phi(X) \mid \mathcal{G}] \quad \mathbb{P}\text{-a.s.}$$

3. (faster) Since Φ is a convex function we can write it as a supremum over affine functions that lie below its graph: $\Phi(x) = \sup_{h_\alpha \in A} h_\alpha(x)$ with $A = \{h_\alpha \text{ affine s.t. } h_\alpha(\alpha) = \Phi(\alpha) : \alpha \in \mathbb{R}\}$. On the one hand, for any $\alpha \in \mathbb{R}, h_\alpha$ lies below the graph of Φ so $\sup_{h_\alpha \in A} h_\alpha(x) \leq \Phi(x)$. On the other hand $\sup_{h_\alpha \in A} h_\alpha(x) \geq h_x(x) = \Phi(x)$ because h_x goes through $(x, \Phi(x))$. Let $h(x) = \lambda x + \mu$ lies below Φ, we have

$$\mathbb{E}\left[\Phi(X)|\,\mathcal{G}\right] \geq \mathbb{E}\left[\lambda X + \mu|\,\mathcal{G}\right] = h(\mathbb{E}\left[X|\,\mathcal{G}\right]).$$

This is true for all affine function h so taking the supremum leads to

$$\mathbb{E}\left[\Phi(X)|\,\mathcal{G}\right] \geq \sup_{h_\alpha \in A} h_\alpha(\mathbb{E}\left[X|\,\mathcal{G}\right]) = \Phi(\mathbb{E}\left[X|\,\mathcal{G}\right]).$$

4. Let $(X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ be a martingale valued in an open interval $I \subset \mathbb{R}$ and $f : I \to \mathbb{R}$ be a convex function such that $(f(X_n))_{n\in\mathbb{N}}$ is L^1. Using the previous question and the fact that $(X_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is a martingale, we get that

$$\mathbb{E}\left[f(X_{n+1}|\,\mathcal{F}_n\right] \geq f(\mathbb{E}\left[X_{n+1}|\,\mathcal{F}_n\right]) = f(X_n),$$

which exactly means that $(f(X_n), \mathcal{F}_n)_{n\in\mathbb{N}}$ is a submartingale.

Solution of Exercise 6.6

1. The predictable quadratic variation $\langle S\rangle_n = A_n$ of S is defined by $A_0 = 0$ and

$$\forall n \geq 1, \Delta A_n = \mathbb{E}\left[\Delta(S^2)_n \middle| \mathcal{F}_{n-1}\right]$$

We have

$$\begin{aligned}\Delta(S^2)_n &= S_n^2 - S_{n-1}^2\\ &= (S_{n-1}+Y_n)^2 - S_{n-1}^2\\ \Delta(S^2)_n &= Y_n^2 + 2S_{n-1}Y_n.\end{aligned}$$

$Y_n \in \{-1,+1\}$ so $Y_n^2 = 1$ and the independence gives $\mathbb{E}\left[S_{n-1}Y_n \middle| \mathcal{F}_{n-1}\right] = 0$, so finally

$$\mathbb{E}\left[\Delta(S^2)_n \middle| \mathcal{F}_{n-1}\right] = 1, \qquad A_n = \langle S\rangle_n = n$$

2. $M_n = \sum_{k=1}^n \operatorname{sgn}(S_{k-1})Y_k$ is adapted to $\mathcal{F}_n$ and L^1. Its increment is $(\Delta M)_n = \operatorname{sgn}(S_{n-1})Y_n$ and we have

$\mathbb{E}\left[(\Delta M)_n \middle| \mathcal{F}_{n-1}\right] = \mathbb{E}\left[(\operatorname{sgn}(S_{n-1})Y_n \middle| \mathcal{F}_{n-1}\right] = \operatorname{sgn}(S_{n-1})\mathbb{E}\left[Y_n \middle| \mathcal{F}_{n-1}\right] = \operatorname{sgn}(S_{n-1})\mathbb{E}\left[Y_n\right] = 0.$

so the process $((\Delta M_n, \mathcal{F}_n))_{n\in\mathbb{N}}$ is a martingale difference and $((M_n, \mathcal{F}_n))_{n\in\mathbb{N}}$ is a martingale. The predictable quadratic variation is defined by

$$\langle M\rangle_n = \sum_{j=1}^n \mathbb{E}\left[(\Delta M_j)^2 \middle| \mathcal{F}_{j-1}\right]$$

We have

$$\Delta\langle M\rangle_n = \mathbb{E}\left[(\Delta M_n)^2 \middle| \mathcal{F}_{n-1}\right] = \mathbb{E}\left[(\operatorname{sgn}(S_{n-1})Y_n)^2 \middle| \mathcal{F}_{n-1}\right] = \operatorname{sgn}(S_{n-1})^2 = \mathbb{1}_{\{S_{n-1}\neq 0\}}$$

$$\langle M\rangle_n = \sum_{j=1}^n \mathbb{1}_{\{S_{j-1}\neq 0\}}$$

3. $((S_n, \mathcal{F}_n))_{n\in\mathbb{N}}$ is a martingale and $x \mapsto |x|$ is convex so $((|S_n|, \mathcal{F}_n))_{n\in\mathbb{N}}$ is a submartingale and it admits a Doob decomposition

$$|S_n| = L_n + B_n$$

where $((L_n, \mathcal{F}_n))_{n\in\mathbb{N}}$ is a martingale and $(B_n)_{n\in\mathbb{N}}$ is a non-increasing predictable process,

$$\Delta B_n = \mathbb{E}\left[\Delta(|S|)_n \middle| \mathcal{F}_{n-1}\right]$$

We can compute

$$\Delta(|S|)_n = |S_n| - |S_{n-1}| = |S_{n-1}+Y_n| - |S_{n-1}| = \begin{cases} 1 & \text{if } S_{n-1} = 0\\ Y_n & \text{if } S_{n-1} \geq 1\\ -Y_n & \text{if } S_{n-1} \leq -1\end{cases}$$

So we finally get

$$\Delta B_n = \mathbb{E}\left[\Delta(|S|)_n \middle| \mathcal{F}_{n-1}\right] = \mathbb{1}_{\{S_{n-1}=0\}}, \qquad B_n = \sum_{j=1}^n \mathbb{1}_{\{S_{j-1}=0\}}$$

$$(\Delta M)_n = \operatorname{sgn}(S_{n-1})Y_n = \begin{cases} 0 & \text{if } S_{n-1} = 0\\ Y_n & \text{if } S_{n-1} \geq 1\\ -Y_n & \text{if } S_{n-1} \leq -1\end{cases} = \begin{cases} 0 & \text{if } S_{n-1} = 0\\ \Delta(|S|)_n & \text{else}\end{cases}$$

Therefore $(\Delta M)_n$ is $\sigma(|S_1|, \dots, |S_n|)$-mesurable, $M_0 = 0$ is $\sigma(|S_1|)$-mesurable and $M_n = \sum_{k=1}^n (\Delta M)_k$ is $\sigma(|S_1|, \dots, |S_n|)$-mesurable.

Solution of Exercise 6.7

1. We consider the martingale $M_n = \sum_{k=1}^n (X_k - \mathbb{E}[X_k]) = \sum_{k=1}^n X_k - n\mathbb{E}[X_0]$. Indeed we have $\mathbb{E}[M_{n+1}|\mathcal{F}_n] = \mathbb{E}[M_n + X_{n+1} - \mathbb{E}[X_0]|\mathcal{F}_n] = M_n$. First consider the case where the stopping time τ is L^∞. Thanks to the optional stopping theorem, $\mathbb{E}[M_\tau] = \mathbb{E}[M_0] = 0$.

   Assume now that τ is any L^1 stopping time and that all the $X_k \geq 0$. For any $n \in \mathbb{N}$, $\tau \wedge n$ is a L^∞ stopping time so that $\mathbb{E}\left[\sum_{k=1}^{\tau\wedge n} X_k\right] = \mathbb{E}[X_0]\mathbb{E}[\tau \wedge n]$. By monotone convergence theorem since $\lim_{n\to\infty} \mathbb{E}[\tau \wedge n] = \mathbb{E}[\tau]$, we have

$$\lim_{n\to\infty} \mathbb{E}\left[\sum_{k=1}^{\tau\wedge n} X_k\right] = \mathbb{E}\left[\sum_{k=1}^{\tau} X_k\right] = \mathbb{E}[X_0]\,\mathbb{E}[\tau].$$

   Finally for the general case, we can first write $\mathbb{E}\left[\sum_{k=1}^{\tau\wedge n} X_k\right] = \mathbb{E}[X_0]\mathbb{E}[\tau \wedge n]$, then consider the absolute value to apply the dominated convergence theorem. We have $|\sum_{k=1}^{\tau\wedge n} X_k| \leq \sum_{k=1}^{\tau} |X_k|$ and this last term has finite expectation since we can apply the result to the non negative random variables $|X_k| \geq 0$ and write $\mathbb{E}[\sum_{k=1}^{\tau} |X_k|] = \mathbb{E}[|X_0|]\mathbb{E}[\tau]$ and by the dominated convergence theorem we can conclude

$$\mathbb{E}\left[\sum_{k=1}^{\tau} X_k\right] = \mathbb{E}[X_0]\,\mathbb{E}[\tau].$$

2. (If the stopping time is adapted to the natural σ-algebra $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$) First recall the expression of the expectation as an infinite sum, using Fubini-Tonelli,

$$\mathbb{E}[Y] = \mathbb{E}\left[\sum_{k=1}^{\infty} \mathbb{1}_{\{Y\geq k\}}\right] = \sum_{k=1}^{\infty} \mathbb{P}(Y \geq k)$$

   Since $\mathbb{1}_{\{\tau\geq k\}} = 1 - \mathbb{1}_{\{\tau\leq k-1\}}$ depends only on $X_1, \dots, X_{k-1}$, we have that X_k is independent from $\mathbb{1}_{\{\tau\geq k\}}$ so that $\mathbb{E}\left[|X_k \mathbb{1}_{\{\tau\geq k\}}|\right] = \mathbb{E}[|X_k|]\,\mathbb{P}(\tau \geq k)$ and the series $\sum \mathbb{P}(\tau \geq k)$ converges if and only if the stopping time τ is integrable. Finally we can exchange summation and expecation to write

$$\mathbb{E}[S_\tau] = \mathbb{E}\left[\sum_{k=1}^{\tau} X_k\right] = \mathbb{E}\left[\sum_{k=1}^{\infty} X_k \mathbb{1}_{\{\tau\geq k\}}\right] = \sum_{k=1}^{\infty} \mathbb{E}\left[X_k \mathbb{1}_{\{\tau\geq k\}}\right] = \sum_{k=1}^{\infty} \mathbb{E}[X_k]\,\mathbb{P}(\tau \geq k) = \mathbb{E}[X_1]\,\mathbb{E}[\tau].$$

Solution of Exercise 6.8

1. Since $(V_n)_{n\geq 1}$ is a non-negative process, we have

$$\mathbb{E}\left[\sum_{n=1}^{\infty} V_n|\Delta X_n|\right] = \sum_{n=1}^{\infty} \mathbb{E}[V_n|\Delta X_n|]$$

   For any given $n \geq 0$, V_n is non-decreasing and is $\mathcal{F}_{n-1}$ mesurable so we can write

$$\mathbb{E}[V_n|\Delta X_n|] = \lim_{s\to+\infty} \mathbb{E}[V_{n\wedge s}|\Delta X_n|] = \lim_{s\to+\infty} \mathbb{E}[\mathbb{E}[V_{n\wedge s}|\Delta X_n||\mathcal{F}_{n-1}]] \leq \lim_{s\to+\infty} M\mathbb{E}[V_{n\wedge s}] = M\mathbb{E}[V_n]$$

   By sum, we have that

$$\mathbb{E}\left[\sum_{n=1}^{\infty} V_n|X_n - X_{n-1}|\right] \leq M \sum_{n=1}^{\infty} \mathbb{E}[V_n].$$

2. Define $V_n = \mathbb{1}_{\{\nu \geq n\}}$ which is non negative and $\mathcal{F}_{n-1}$ mesurable, because $\{\nu \geq n\} = \{\nu < n-1\}^c$, and apply the result of the previous question,

$$\mathbb{E}\left[\sum_{n=1}^{\infty} \mathbb{1}_{\{\nu \geq n\}} |X_n - X_{n-1}|\right] \leq M \sum_{n=1}^{\infty} \mathbb{E}\left[\mathbb{1}_{\{\nu \geq n\}}\right] = M\mathbb{E}\left[\nu\right] < +\infty.$$

We have

$$\sum_{n\geq 1} \mathbb{1}_{\{\nu \geq n\}}(X_n - X_{n-1}) = \sum_{n=1}^{\nu}(X_n - X_{n-1}) = X_\nu - X_0$$

Therefore

$$\mathbb{E}\left[|X_\nu - X_0|\right] \leq \mathbb{E}\left[\sum_{n=1}^{\infty} \mathbb{1}_{\{\nu \geq n\}} |X_n - X_{n-1}|\right] \leq M\mathbb{E}\left[\nu\right] < +\infty.$$

so can conclude that X_ν is L^1.

3. For any integrable stopping time ν we have

$$X_\nu = X_0 + \sum_{n\geq 1} \mathbb{1}_{\{\nu \geq n\}}(X_n - X_{n-1}).$$

Since ν is an integrable stopping time we also have that $\nu \wedge p$ is an integrable stopping time so

$$\begin{aligned}
X_{\nu\wedge p} &= X_0 + \sum_{n\geq 1} \mathbb{1}_{\{\nu\wedge p \geq n\}}(X_n - X_{n-1}) \\
|X_{\nu\wedge p}| &\leq |X_0| + \sum_{n\geq 1} \mathbb{1}_{\{\nu\wedge p \geq n\}}|X_n - X_{n-1}| \\
&\leq |X_0| + \sum_{n\geq 1} \mathbb{1}_{\{\nu \geq n\}}|X_n - X_{n-1}|
\end{aligned}$$

The term on the right hand side is integrable and independent of p so we can conclude using dominated convergence theorem, and write the convergence in L^1

$$X_{\nu\wedge p} \underset{p\to+\infty}{\to} X_\nu.$$

4. $\nu_1 \leq \nu_2$ with $\nu_2 \in L^1$ so we have by Doob's inequality,

$$\forall p \in \mathbb{N}, \mathbb{E}\left[X_{\nu_2\wedge p} | \mathcal{F}_{\nu_1\wedge p}\right] \leq X_{\nu_1\wedge p}.$$

For any given $A \in \mathcal{F}_{\nu_1}$, we have $A \cap \{\nu_1 \leq k\} \in \mathcal{F}_{\nu_1\wedge k}$ because $\mathcal{F}_{\nu_1\wedge k} = \mathcal{F}_{\nu_1} \cap \mathcal{F}_k$ so we can write

$$\mathbb{E}\left[X_{\nu_2\wedge k}\mathbb{1}_{A\cap\{\nu_1\leq k\}}\right] \leq \mathbb{E}\left[X_{\nu_1\wedge k}\mathbb{1}_{A\cap\{\nu_1\leq k\}}\right]$$

We have that $X_{\nu_2\wedge k}\mathbb{1}_{A\cap\{\nu_1\leq k\}} \to X_{\nu_2}\mathbb{1}_A$ and $X_{\nu_1\wedge k}\mathbb{1}_{A\cap\{\nu_1\leq k\}} \to X_{\nu_1}\mathbb{1}_A$. Moreover we have the convergence in L^1 so we can take the limit when k goes to infinity,

$$\forall A \in \mathcal{F}_{\nu_1}, \mathbb{E}\left[X_{\nu_2}\mathbb{1}_A\right] \leq \mathbb{E}\left[X_{\nu_1}\mathbb{1}_A\right],$$

which exactly means that

$$\mathbb{E}\left[X_{\nu_2} | \mathcal{F}_{\nu_1}\right] \leq X_{\nu_1}$$

Solution of Exercise 6.9

1. $\{\nu_i = \infty\}$ means that the sum S_n^i is never bigger than the integer $a \in \mathbb{N}$ so that at some point, all the random variables X_n^i are equal to 0,

$$\{\nu_i = \infty\} = \{Card(k : X_k^i = 1) < a\} \subset \{\exists n \in \mathbb{N}, \forall k \geq n, X_k^i = 0\} \subset \bigcup_{n \in \mathbb{N}} \bigcap_{k \geq n} \{X_k^i = 0\}.$$

   For any given $n \in \mathbb{N}$,

$$\mathbb{P}\left(\bigcap_{k \geq n} \{X_k^i = 0\}\right) = \lim_{m \to +\infty} \mathbb{P}\left(\bigcap_{n \leq k \leq m} \{X_k^i = 0\}\right) = \lim_{m \to +\infty} \left(\frac{1}{2}\right)^{m-n} = 0.$$

   Every countable union of negligible events is negligible and finally,

$$\mathbb{P}(\{\nu_i < \infty\}) = 1.$$

2. $M_n^i = 2S_n^i - n$ and $M_n^{i,j} = M_n^i M_n^j - n\delta_{i,j}$ are adapted to $\mathcal{F}_n = \sigma(X_k^i :\ i \in \{1,2\},\ k \leq n)$, and we have

$$\begin{aligned} M_{n+1}^i &= 2S_{n+1}^i - (n+1) = M_n^i + 2X_{n+1}^i - 1 \\ \mathbb{E}\left[M_{n+1}^i \middle| \mathcal{F}_n\right] &= M_n^i + 2\mathbb{E}\left[X_{n+1}^i\right] - 1 = M_n^i. \end{aligned}$$

$$\begin{aligned} M_{n+1}^{i,j} &= M_{n+1}^i M_{n+1}^j - (n+1)\delta_{i,j} \\ &= (M_n^i + 2X_{n+1}^i - 1)(M_n^j + 2X_{n+1}^j - 1) - n\delta_{i,j} - \delta_{i,j} \\ &= M_n^{i,j} + M_n^i(2X_{n+1}^j - 1) + M_n^j(2X_{n+1}^i - 1) + (2X_{n+1}^i - 1)(2X_{n+1}^j - 1) - \delta_{i,j} \end{aligned}$$

   Taking the conditional expectation we have

$$\mathbb{E}\left[M_n^i(2X_{n+1}^j - 1) \middle| \mathcal{F}_n\right] = 0, \quad \mathbb{E}\left[M_n^j(2X_{n+1}^i - 1) \middle| \mathcal{F}_n\right] = 0.$$
$$\mathbb{E}\left[(2X_{n+1}^i - 1)(2X_{n+1}^j - 1) \middle| \mathcal{F}_n\right] = \delta_{i,j}.$$

   Finally,

$$\mathbb{E}\left[M_{n+1}^{i,j} \middle| \mathcal{F}_n\right] = M_n^{i,j}.$$

3. For any given $n \in \mathbb{N}$, $\nu \wedge n$ is L^∞ stopping time so $\mathbb{E}\left[M_{\nu \wedge n}^i\right] = \mathbb{E}\left[M_0^i\right] = 0$ which means

$$\mathbb{E}[\nu \wedge n] = 2\mathbb{E}\left[S_{\nu \wedge n}^i\right].$$

   We have by definition $\mathbb{E}[\nu_1] = \mathbb{E}[\nu_2] = 2a$ and $\mathbb{E}\left[S_{\nu \wedge n}^i\right] \leq \mathbb{E}[a] = a$ so

$$\forall n \in \mathbb{N}, \mathbb{E}[\nu \wedge n] \leq 2a.$$

   By monotone convergence theorem,

$$\lim_{n \to \infty} \mathbb{E}[\nu \wedge n] = \mathbb{E}\left[\lim_{n \to \infty} \nu \wedge n\right] = \mathbb{E}[\nu],$$

   and finally

$$\mathbb{E}[\nu] \leq 2a.$$

4. For any given $n \in \mathbb{N}$, $\nu \wedge n$ is L^∞ stopping time so $\mathbb{E}\left[M_{\nu \wedge n}^{i,j}\right] = \mathbb{E}\left[M_0^{i,j}\right] = 0$ which means

$$\mathbb{E}\left[(2S_{\nu \wedge n}^i - \nu \wedge n)(2S_{\nu \wedge n}^j - \nu \wedge n)\right] = \mathbb{E}[\delta_{i,j}\nu \wedge n].$$

$$
\begin{aligned}
\mathbb{E}\left[(\nu\wedge n)^2\right] &= \mathbb{E}\left[\delta_{i,j}\nu\wedge n\right] - 4\mathbb{E}\left[S^i_{\nu\wedge n}S^j_{\nu\wedge n}\right] + 2\mathbb{E}\left[(S^i_{\nu\wedge n}+S^j_{\nu\wedge n})(\nu\wedge n)\right]\\
&\leq \mathbb{E}\left[\nu\right]\delta_{i,j} + 2a\mathbb{E}\left[\nu\right]\\
\mathbb{E}\left[(\nu\wedge n)^2\right] &\leq 2a(1+2a).
\end{aligned}
$$

Therefore,

$$\mathbb{E}\left[\nu^2\right] = \lim_{n\to\infty}\mathbb{E}\left[(\nu\wedge n)^2\right] < \infty.$$

$$
\begin{aligned}
(2S^i_{\nu\wedge n}-\nu\wedge n)(2S^j_{\nu\wedge n}-\nu\wedge n) &\leq \nu^2+4a^2\\
\mathbb{E}\left[(2S^i_{\nu\wedge n}-\nu\wedge n)(2S^j_{\nu\wedge n}-\nu\wedge n)\right] &\leq \mathbb{E}\left[\nu^2+4a^2\right] < \infty.
\end{aligned}
$$

By dominated convergence theorem,

$$\mathbb{E}\left[(2S^i_\nu-\nu)(2S^j_\nu-\nu)\right] = \mathbb{E}\left[\nu\delta_{i,j}\right]$$

and finally

$$\mathbb{E}\left[M^{i,j}_\nu\right] = 0.$$

5.

$$
\begin{aligned}
M^{1,1}_n - 2M^{1,2}_n + M^{2,2}_n &= (2S^1_n-n)^2 - 2(2S^1_n-n)(2S^2_n-n) + (2S^2_n-n)^2 - 2n\\
&= \left[(2S^1_n-n)-(2S^2_n-n)\right]^2 - 2n\\
&= 4\left(S^1_n-S^2_n\right)^2 - 2n.
\end{aligned}
$$

Since we have a martingale, the expectation at ν is equal to 0 and we get

$$
\begin{aligned}
4\mathbb{E}\left[\left(S^1_\nu-S^2_\nu\right)^2\right] &= 2\mathbb{E}\left[\nu\right] \leq 4a\\
\mathbb{E}\left[\left(S^1_\nu-S^2_\nu\right)^2\right] &\leq a
\end{aligned}
$$

We can conclude by Jensen's inequality,

$$\mathbb{E}\left[|S^1_\nu-S^2_\nu|\right] \leq \sqrt{\mathbb{E}\left[\left(S^1_\nu-S^2_\nu\right)^2\right]} \leq \sqrt{a}.$$

Solution of Exercise 6.10 We use the same idea as for the Wald first inequality and write

$$
\begin{aligned}
\mathrm{Var}\left(S_\tau\right) = \mathrm{Var}\left(\sum_{k=1}^{\tau}X_k\right) &= \mathbb{E}\left[\left(\sum_{k=1}^{\tau}X_k\right)^2\right] = \mathbb{E}\left[\left(\sum_{k=1}^{\infty}X_k\mathbb{1}_{\{\tau\geq k\}}\right)^2\right]\\
&= \mathbb{E}\left[\sum_{k=1}^{\infty}X_k^2\mathbb{1}_{\{\tau\geq k\}} + 2\sum_{1=q<p}^{\infty}X_pX_q\mathbb{1}_{\{\tau\geq p\}}\right]
\end{aligned}
$$

Using the independence we get that,

$$
\begin{aligned}
\mathrm{Var}\left(S_\tau\right) &= \sum_{k=1}^{\infty}\mathbb{E}\left[X_k^2\mathbb{1}_{\{\tau\geq k\}}\right] + 2\sum_{p=1}^{\infty}\sum_{q=1}^{p-1}\mathbb{E}\left[X_pX_q\mathbb{1}_{\{\tau\geq p\}}\right]\\
&= \sum_{k=1}^{\infty}\mathbb{E}\left[X_k^2\right]\mathbb{P}\left(\tau\geq k\right) + 2\underbrace{\sum_{p=1}^{\infty}\sum_{q=1}^{p-1}\mathbb{E}\left[X_p\right]\mathbb{E}\left[X_q\mathbb{1}_{\{\tau\geq p\}}\right]}_{=0}\\
\mathrm{Var}\left(S_\tau\right) &= \mathbb{E}\left[\tau\right]\mathbb{E}\left[X_1^2\right] = \mathbb{E}\left[\tau\right]\sigma^2.
\end{aligned}
$$

Solution of Exercise 6.11

1. Define $\mathcal{F}_n = \mathcal{F}_n^U = \begin{cases} (\emptyset, \Omega) & \text{if } n = 0 \\ \sigma(U_1, \ldots, U_n) & \text{if } n \geq 1 \end{cases}$ so that $Z_n = (q/p)^{S_n}$ is adapted to $\mathcal{F}_n$.

$$\begin{aligned}\mathbb{E}\left[Z_{n+1} \middle| \mathcal{F}_n\right] &= \mathbb{E}\left[Z_n \left(\frac{q}{p}\right)^{U_{n+1}} \middle| \mathcal{F}_n\right] = Z_n \mathbb{E}\left[\left(\frac{q}{p}\right)^{U_{n+1}} \middle| \mathcal{F}_n\right] \\ &= Z_n \mathbb{E}\left[\left(\frac{q}{p}\right)^{U_{n+1}}\right] \\ &= Z_n \left(\left(\frac{q}{p}\right)^1 p + \left(\frac{q}{p}\right)^{-1} q\right) \\ \mathbb{E}\left[Z_{n+1} \middle| \mathcal{F}_n\right] &= Z_n\end{aligned}$$

$(Z_n, \mathcal{F}_n)_{n \in \mathbb{N}}$ is a martingale.

2. Using maximal inequality, we get for any $a \in \mathbb{R}$,

$$\begin{aligned} a\mathbb{P}\left(\sup_{n\geq 0} Z_n \geq a\right) &\leq \mathbb{E}\left[Z_0\right] = 1 \\ \mathbb{P}\left(\sup_{n\geq 0} \left(\frac{q}{p}\right)^{S_n} \geq a\right) &\leq \frac{1}{a} \\ \mathbb{P}\left(\sup_{n\geq 0} S_n \ln\left(\frac{q}{p}\right) \geq \ln(a)\right) &\leq \frac{1}{a}. \end{aligned}$$

If $q > p$, define $k = \frac{\ln(a)}{\ln(q/p)}$ so that $a = \left(\frac{q}{p}\right)^k$,

$$\mathbb{P}\left(\sup_{n\geq 0} S_n \geq k\right) \leq \left(\frac{p}{q}\right)^k.$$

$\sup_{n\geq 0} S_n$ is non-negative so we finally get by sum,

$$\mathbb{E}\left[\sup_{n\geq 0} S_n\right] = \sum_{k=1}^{\infty} \mathbb{P}\left(\sup_{n\geq 0} S_n \geq k\right) \leq \sum_{k=1}^{\infty} \left(\frac{p}{q}\right)^k = \frac{p}{q-p}.$$

Solution of Exercise 6.12

1. By the law of large numbers, we have $\frac{1}{n}\sum_{i=1}^n X_i \to m < 0$, so $S_n \to_{n\to+\infty} -\infty$ and $\sup_{n\geq 0} S_n < +\infty$, $\quad \mathbb{P}\left(W < +\infty\right) = 1$.

2. $\mathbb{E}\left[\mathrm{e}^{\lambda S_{n+1}} \middle| \mathcal{F}_n\right] = \mathbb{E}\left[\mathrm{e}^{\lambda S_n} \mathrm{e}^{\lambda X_{n+1}} \middle| \mathcal{F}_n\right] = \mathrm{e}^{\lambda S_n} \mathbb{E}\left[\mathrm{e}^{\lambda X_{n+1}}\right] = \mathrm{e}^{\lambda S_n} \mathrm{e}^{\lambda m + \lambda^2 \sigma^2/2}$.

3. $(\exp(\lambda_0 S_n))_{n\in\mathbb{N}}$ is a martingale $\Leftrightarrow \lambda_0 m + \lambda_0^2 \sigma^2/2 = 0 \Leftrightarrow \lambda_0 = \frac{-2m}{\sigma^2}$.

4. $(\exp(\lambda_0 S_n))_{n\in\mathbb{N}}$ is a non-negative martingale so we have by maximal inequality that for any $a > 1$,

$$a\mathbb{P}\left(\sup_{n\in\mathbb{N}} \mathrm{e}^{\lambda_0 S_n} > a\right) \leq \mathbb{E}\left[\mathrm{e}^{\lambda_0 S_0}\right] = 1,$$

hence,

$$\forall a > 1, \quad \mathbb{P}\left(\mathrm{e}^{\lambda_0 W} > a\right) \leq \frac{1}{a}.$$

Define $t = \ln(a)/\lambda_0$ so that $a = \mathrm{e}^{\lambda_0 t}$,

$$\begin{aligned}\forall a > 1, \quad \mathbb{P}\left(\lambda_0 W > \ln(a)\right) &\leq \frac{1}{a}\\ \mathbb{P}\left(W > \ln(a)/\lambda_0\right) &\leq \frac{1}{a}\\ \mathbb{P}\left(W > t\right) &\leq \mathrm{e}^{-\lambda_0 t}.\end{aligned}$$

5. We have

$$\mathrm{e}^{\lambda W} - \mathrm{e}^{\lambda 0} = \int_0^W \lambda \mathrm{e}^{\lambda s} \mathrm{d}s,$$

which gives

$$\mathrm{e}^{\lambda W} = 1 + \lambda \int_0^{+\infty} \mathbb{1}_{[0,W]}(s) \mathrm{e}^{\lambda t} \mathrm{d}t.$$

Using Fubini theorem,

$$\mathbb{E}\left[\mathrm{e}^{\lambda W}\right] = 1 + \lambda \int_0^{+\infty} \mathbb{P}\left(W > t\right) \mathrm{e}^{\lambda t} \mathrm{d}t.$$

For any $0 < \lambda < \lambda_0$, since $\mathbb{P}\left(W > t\right) \leq \mathrm{e}^{-\lambda_0 t}$, we have

$$\mathbb{E}\left[\mathrm{e}^{\lambda W}\right] \leq 1 + \int_0^{+\infty} \lambda \mathrm{e}^{-\lambda_0 t} \mathrm{e}^{\lambda t} \mathrm{d}t = 1 + \frac{\lambda}{\lambda_0 - \lambda} < +\infty.$$

Solution of Exercise 6.13

1. Since A_n is predictable we have for all $k \leq n, \{A_{k+1} > a^2\} \in \mathcal{F}_k$. We can write $\{\tau_a \leq n\} = \cup_{k \leq n}\{A_{k+1} > a^2\} \in \mathcal{F}_n$ and τ_a is stopping time.

2. Since $(M_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is a martingale and $x \mapsto x^2$ is convex, we have that $\left(M_n^2, \mathcal{F}_n\right)_{n\in\mathbb{N}}$ is a non-negative submartingale and so is $\left(M_{n\wedge\tau_a}^2, \mathcal{F}_n\right)_{n\in\mathbb{N}}$. We can apply maximal inequality $a^2 \mathbb{P}\left(\max_{0\leq k\leq n} |M_{k\wedge\tau_a}| > a\right) = a^2 \mathbb{P}\left(\max_{0\leq k\leq n} M_{k\wedge\tau_a}^2 > a^2\right) \leq \mathbb{E}\left[M_{n\wedge\tau_a}^2\right] = \mathbb{E}\left[A_{n\wedge\tau_a}\right]$,

$$\mathbb{P}\left(\max_{0\leq k\leq n} |M_{k\wedge\tau_a}| > a\right) \leq a^{-2}\mathbb{E}\left[A_{n\wedge\tau_a}\right].$$

We have $\{\max_{0\leq k\leq n} |M_{k\wedge\tau_a}| > a\} = \{\exists k \leq n : |M_{k\wedge\tau_a}| > a\} = \cup_{0\leq k\leq n}\{|M_{k\wedge\tau_a}| > a\}$ so that when n goes to $+\infty$ we get the supremum over $n \in \mathbb{N}$. For the term on the right hand side, we need to consider whether A_∞ is finite or not.

$$\begin{aligned}\mathbb{E}\left[A_{n\wedge\tau_a}\right] &= \mathbb{E}\left[A_{n\wedge\tau_a}\mathbb{1}_{\{\tau_a=\infty\}}\right] + \mathbb{E}\left[A_{n\wedge\tau_a}\mathbb{1}_{\{\tau_a<\infty\}}\right]\\ &= \underbrace{\mathbb{E}\left[A_{n\wedge\tau_a}\mathbb{1}_{\{A_\infty\leq a^2\}}\right]}_{\to\mathbb{E}\left[(A_\infty\wedge a^2)\mathbb{1}_{\{\tau_a=\infty\}}\right]} + \underbrace{\mathbb{E}\left[A_{n\wedge\tau_a}\mathbb{1}_{\{A_\infty>a^2\}}\right]}_{\to\mathbb{E}\left[(A_\infty\wedge a^2)\mathbb{1}_{\{\tau_a<\infty\}}\right]}\end{aligned}$$

3. We have

$$\mathbb{P}\left\{\sup_{n\geq 0}|M_n| > a\right\} = \mathbb{P}\left\{\sup_{n\geq 0}|M_n| > a, \tau_a < \infty\right\} + \mathbb{P}\left\{\sup_{n\geq 0}|M_n| > a, \tau_a = \infty\right\}$$

On $\{\tau_a = \infty\}$, we have $|M_n| = |M_{n\wedge\tau_a}|$ so $\mathbb{P}\left\{\sup_{n\geq 0}|M_n| > a, \tau_a = \infty\right\} = \mathbb{P}\left\{\sup_{n\geq 0}|M_{n\wedge\tau_a}| > a\right\}$ Besides, $\{\tau_a < \infty\} \subset \{A_\infty > a^2\}$ so

$$\mathbb{P}\left\{\sup_{n\geq 0}|M_n| > a, \tau_a < \infty\right\} \leq \mathbb{P}\left\{\tau_a < \infty\right\} \leq \mathbb{P}\left\{A_\infty > a^2\right\}$$

Finally

$$\mathbb{P}\left\{\sup_{n\geq 0}|M_n| > a\right\} \leq \mathbb{P}\left\{A_\infty > a^2\right\} + \mathbb{P}\left\{\sup_{n\in\mathbb{N}}|M_{n\wedge\tau_a}| > a\right\}.$$

4. Since X is a non-negative random variable, we have by Fubini that for all λ,

$$\int_0^\lambda \mathbb{P}(X>t)\,dt = \int_0^\lambda \mathbb{E}\left[\mathbb{1}_{\{X>t\}}\right]dt = \mathbb{E}\left[\int_0^\lambda \mathbb{1}_{\{X>t\}}dt\right] = \mathbb{E}\left[\int_0^{X\wedge\lambda} 1dt\right] = \mathbb{E}\left[X\wedge\lambda\right],$$

$$\int_0^{+\infty} \frac{\mathbb{E}\left[X\wedge a^2\right]}{a^2}da = \mathbb{E}\left[\int_0^{\sqrt{X}} 1da + X\int_{\sqrt{X}}^{+\infty}\frac{1}{a^2}da\right] = \mathbb{E}\left[\sqrt{X}+\frac{X}{\sqrt{X}}\right] = \mathbb{E}\left[2\sqrt{X}\right].$$

5. We have by integrating with respect to a between 0 and $+\infty$,

$$\begin{aligned}
\mathbb{P}\left(\sup_{n\geq 0}|M_n|>a\right) &\leq \mathbb{P}\left(A_\infty>a^2\right)+\mathbb{P}\left(\sup_{n\in\mathbb{N}}|M_{n\wedge\tau_a}|>a\right)\\
&\leq \mathbb{P}\left(\sqrt{A_\infty}>a\right)+a^{-2}\mathbb{E}\left[A_\infty\wedge a^2\right]\\
\int_0^{+\infty}\mathbb{P}\left(\sup_{n\geq 0}|M_n|>a\right)da &\leq \int_0^{+\infty}\mathbb{P}\left(\sqrt{A_\infty}>a\right)da+\int_0^{+\infty}a^{-2}\mathbb{E}\left[A_\infty\wedge a^2\right]da\\
\mathbb{E}\left[\sup_{n\geq 0}|M_n|\right] &\leq \mathbb{E}\left[\sqrt{A_\infty}\right]+2\mathbb{E}\left[\sqrt{A_\infty}\right]\\
\mathbb{E}\left[\sup_{n\geq 0}|M_n|\right] &\leq 3\mathbb{E}\left[\sqrt{A_\infty}\right].
\end{aligned}$$

6. Denote $\mathbb{E}\left[Y_1^2\right]=\sigma^2$. Using the same idea as for Wald identity, we can write using optional stopping theorem that $\mathbb{E}\left[S_{\tau\wedge n}\right]=0$ and we want to let n goes to $+\infty$ so we uniformly bound this term with an integrable process and apply dominated convergence theorem. We have $A_n=\langle S\rangle_n=n\sigma^2$ so that $A_\tau=\tau\sigma^2$ and also

$$\mathbb{E}\left[\sup_{n\geq 0}|S_{n\wedge\tau}|\right]\leq 3\mathbb{E}\left[\sqrt{A_\tau}\right]=3\sigma\sqrt{\tau}<+\infty$$

Therefore $\mathbb{E}\left[S_\tau\right]=\mathbb{E}\left[\lim_{n\to\infty}S_{\tau\wedge n}\right]=\lim_{n\to\infty}\mathbb{E}\left[S_{\tau\wedge n}\right]=0$.

Solution of Exercise 6.14

1.
2.
3.
4.

Solution of Exercise 6.15

1. $\forall n\geq 1, \{\tau=n\}=\{U_1=0,\ldots,U_{n-1}=0,U_n=+1\}\in\mathcal{F}_n$ so the random variable τ is a $\mathcal{F}$-stopping time. It follows a geometric distribution and is finite almost surely: $\forall n\geq 1, \mathbb{P}(\{\tau=n\})=(1/2)^n, \quad \mathbb{P}(\{\tau<+\infty\})=\sum_{k=1}^{+\infty}(1/2)^k=1$. At time n, if no Tails was drawn, then the player lost all his bets and his fund is equal to $Y_n=-S_1-S_2-\ldots-S_n=-\sum_{k=1}^n 2^k=-2^{n+1}+2$. When the first Tails comes out, at time $n=\tau$, the player gets $S_n=2^n$ and his fund is equal to $Y_n=Y_{n-1}+2^n=2$. The strategy of the classical martingale is that if we play a long enough amount of time (i.e. until we reach the random time τ), we can always repay the debts and reach a positive fund of 2 euros.

2. We have: $\forall n\geq 1, S_n=2^n\mathbb{1}_{\{\tau>n-1\}}$. Since τ is a $\mathcal{F}$-stopping time, the random variable S_n is $\mathcal{F}_{n-1}$-mesurable and the process (S_n) is $\mathcal{F}$-predictable.

3. For $n \leq \tau$, we have: $Y_n - Y_{n-1} = (2U_n - 1)2^n$, so:$\forall n \geq 1, Y_n = \sum_{k=1}^{n\wedge\tau} 2^k(2U_k - 1)$. The process defined by $M_n = \sum_{k=1}^{n} 2^k(2U_k - 1) = M_{n-1} + 2^n(2U_n - 1)$ is a $\mathcal{F}$-martingale since $\mathbb{E}\left[M_n | \mathcal{F}_{n-1}\right] = M_{n-1} + 2^n(2\mathbb{E}\left[U_n\right] - 1) = M_{n-1}$ because U_n is independent from $\mathcal{F}_{n-1}$ and has expectation equal to 1/2. The random process (Y_n) is therefore a $\mathcal{F}$-martingale, equals to the martingale M_n stopped at the $\mathcal{F}$-stopping time, finite a.s. τ : $\forall n \geq 1, Y_n = M_{n\wedge\tau}$.

4. We have: $\forall n \geq 1, Y_n^2 = M_{n\wedge\tau}^2$ and the r.v. M_n^2 is L^1 (bounded). The predictable quadratic variation $\langle Y\rangle_n$ of Y_n is equal to the predictable process of the submartingale M_n^2 stopped at time τ. We have $\langle M\rangle_n = \sum_{k=1}^{n} \mathbb{E}\left[(M_k - M_{k-1})^2 \middle| \mathcal{F}_{k-1}\right] = \sum_{k=1}^{n} 4^k = (4/3)(4^n - 1)$, and therefore $\langle Y\rangle_n = \langle M\rangle_{n\wedge\tau} = (4/3)(4^{n\wedge\tau} - 1)$.

$$\begin{aligned}\mathbb{E}\left[\langle Y\rangle_n\right] &= \frac{4}{3}\left(\sum_{k=1}^{n-1} 4^k \mathbb{P}\left(\tau = k\right) + 4^n \mathbb{P}\left(\tau \geq n\right) - 1\right) \\ &= \frac{4}{3}\left(\sum_{k=1}^{n-1} 4^k (1/2)^k + 4^n (1/2)^{n-1} - 1\right) \\ &= 4(2^n - 1).\end{aligned}$$

Thus, $\mathbb{E}\left[\langle Y\rangle_n\right]$ goes to $+\infty$ when $n \to +\infty$. In virtue of Doob decomposition, we have $\mathbb{E}\left[Y_n^2\right] = \mathbb{E}\left[\langle Y\rangle_n\right]$ so that the sequence (Y_n) does not converge in L^2.

5. We have $Y_0 = 0$ and for all $n \geq 1, Y_n = 2\mathbb{1}_{\{n\geq\tau\}} + (2 - 2^{n+1})\mathbb{1}_{\{n<\tau\}} = 2 - 2^{n+1}\mathbb{1}_{\{n<\tau\}}$. So:$\mathbb{P}\left(Y_n = 2\right) = \mathbb{P}\left(n \geq \tau\right) = \sum_{k=1}^{n}(1/2)^k = 1 - (1/2)^n$ and $\mathbb{P}\left(Y_n = 2 - 2^{n+1}\right) = (1/2)^n$. Since τ is finite a.s., Y_n goes to 2 with probability 1 when n goes to $+\infty$.

6. If (Y_n) converges in L^1 then it is necessarily towards its limit a.s. which is 2. But, since we have a martingale, we have $\mathbb{E}\left[Y_n\right] = \mathbb{E}\left[Y_0\right] = 0 \neq 2$, so there is no convergence in L^1.

7. We set $L = 2^k$ for a fixed $k \geq 1$. Define $\tilde{\tau} = \inf\{n \in \mathbb{N} : Y_n - 2^{n+1} < -L\}$. The process (Y_n) being adapted to $\mathcal{F}$, the random variable $\tilde{\tau}$ is a $\mathcal{F}$-stopping time: it is the reaching time of (Y_n) into $\{\dots, -L + 2^{n+1} - 2, -L + 2^{n+1} - 1\}$. The game stops at time $N = \tau \wedge \tilde{\tau} : Z_n = Y_{n\wedge N}$. Notice that we can also write $N = \tau \wedge (k-1)$: $k - 1$ is the largest integer l such that $-S_1 - \ldots - S_l - 2^{l+1} < -L$ (i.e. the longest series of autorised succesive fails). As a minimum of two stopping times, the random variable N is a $\mathcal{F}$-stopping time and we have for $n \in \{1, \dots, k-2\}$,

$$\mathbb{P}\left(N = n\right) = \mathbb{P}\left(\tau = n\right) = (1/2)^n, \quad \mathbb{P}\left(N = k-1\right) = \mathbb{P}\left(\tau \geq k-1\right) = (1/2)^{k-1}.$$

The sequence (Z_n) is obtained by stopping the $\mathcal{F}$-martingale (Y_n) at the $\mathcal{F}$-stopping time N, so it is also a $\mathcal{F}$-martingale.

We have $Z_n = M_{n\wedge N}$ for all $n \geq 1$. Since $N \leq \tau$ is finite a.s., Z_n converges a.s. towards $M_N = M_{\tau\wedge(k-1)}$. The sequence (Z_n) is bounded : it takes its values in $\{-L, \dots, 2\}$, so there is convergence in L^1 thanks to the dominated convergence theorem.

Solution of Exercise 6.16

1. By induction, we show that for all $n \geq 1$, X_n is $\mathcal{F}_n$-mesurable and square integrable. We have: $\forall n \geq 1, \mathbb{E}\left[X_n | \mathcal{F}_{n-1}\right] = 2X_{n-1}\mathbb{E}\left[U_n\right] = X_{n-1}$ a.s.

2. $\langle X\rangle_n = \sum_{k=1}^{n} \mathbb{E}\left[(X_k - X_{k-1})^2 \middle| \mathcal{F}_{k-1}\right] = \sum_{k=1}^{n} X_{k-1}^2 \mathbb{E}\left[(2U_k - 1)^2\right] = (1/3)\sum_{k=1}^{n} X_{k-1}^2$. Moreover, $X_n = 2^n U_1 \times \ldots \times U_n$, so: $\mathbb{E}\left[X_n^2\right] = 2^{2n}\left(\mathbb{E}\left[U_1^2\right]\right)^n = (4/3)^n$. Therefore: $\mathbb{E}\left[\langle X\rangle_n\right] = (1/3)\sum_{k=1}^{n}(4/3)^{k-1} = (4/3)^n - 1$. This term goes to $+\infty$ as n goes to $+\infty$. The sequence (X_n) is not bounded in L^2 so it does not converge in L^2.

3. We have $X_n \geq 0$ a.s. and $\mathbb{E}\left[X_n\right] = 2^n\left(\mathbb{E}\left[U_1\right]\right)^n = 1$. Hence, $\sup_n \mathbb{E}\left[X_n\right] < +\infty$ so that the sequence (X_n) converges a.s. towards an integrable random variable X_∞.

4. $\forall n \geq 1, Y_n = \ln(X_n) = n\ln(2) + \sum_{k=1}^n \ln(U_k)$. By the law of large numbers applied to the i.i.d sequence of integrable r.v. (U_k), we have $Y_n/n \to \ln(2) + \int_0^1 \ln(u)\mathrm{d}u = -1 + \ln(2)$. Thus, $X_n \sim \exp\left(-(1-\ln(2)n\right) \to 0$ a.s. when $n \to +\infty$.

5. By contradiction, if (X_n) converges in L^1 then on the one hand it is necessarily towards its limit a.s. which is equal to 0 but we have on the other hand that $\mathbb{E}\left[X_n\right] = +1 \neq 0$. Therefore, (X_n) does not converge in L^1.

Solution of Exercise 6.17

1. By induction, we show that for all $n \geq 1$, X_n is $\mathcal{F}_n$-mesurable and square integrable. We have: $\forall n \geq 1, \mathbb{E}\left[X_{n+1} \middle| \mathcal{F}_n\right] = (1-\lambda)X_n + \lambda X_n \mathbb{E}\left[\xi_n\right] = X_n$ a.s.

2. We have $\mathbb{E}\left[X_n\right] = \mathbb{E}\left[X_0\right] = 1$.

3. We have $X_n \geq 0$ a.s. and $\sup_n \mathbb{E}\left[X_n\right] < +\infty$ so that the sequence (X_n) converges a.s. towards an integrable random variable X_∞.

4. We have $\mathbb{E}\left[\xi_{n+1}\right] = 1$ and $\mathbb{E}\left[\xi_{n+1}^2\right] = 2$, so that

$$\mathbb{E}\left[X_{n+1}^2\right] = (1-\lambda)^2\mathbb{E}\left[X_n^2\right] + 2\lambda(1-\lambda)\mathbb{E}\left[X_n^2\right] + 2\lambda^2\mathbb{E}\left[X_n^2\right] = (\lambda^2+1)\mathbb{E}\left[X_n^2\right].$$

   Therefore $\mathbb{E}\left[X_n^2\right] = (\lambda^2+1)^n$.

5. For all $\lambda > 0$, the sequence (X_n) is not bounded in L^2 so it does not converge in L^2.

6. $\forall n \geq 1, \langle X\rangle_n = \sum_{k=1}^n \mathbb{E}\left[(X_k - X_{k-1})^2 \middle| \mathcal{F}_{k-1}\right] = \sum_{k=1}^n \lambda^2 X_{k-1}^2 \mathbb{E}\left[(\xi_k - 1)^2\right] = \sum_{k=1}^n \lambda^2 X_{k-1}^2$.

7. $\mathbb{P}\left(X_n = 2^n\right) = (1/2)^n = 1 - \mathbb{P}\left(X_n = 0\right)$. We deduce that X_n converges towards 0 in probability, so that its limit a.s. is also equals to 0. We have $X_n \geq 0$ a.s. and $\mathbb{E}\left[X_n\right] = 1$ for all n. The sequence (X_n) cannot converge in L^1 otherwise the limit would necessarily be equal to 0. The random variables X_n are not uniformly integrable since this would imply the convergence in L^1.

Solution of Exercise 6.18

1. First notice that for all $n \geq 0$, the random variables F_n^+ and F_n^- are $\mathcal{F}_n$-mesurable, non negative and bounded by the integrable random variable $|F_n|$. The sequences (F_n^+) and (F_n^-) are two non negative $\mathcal{F}$-adapted sequences of L^1. Besides, $\mathbb{E}\left[F_{n+1}^+ \middle| \mathcal{F}_n\right] \geq \max\{\mathbb{E}\left[F_{n+1} \middle| \mathcal{F}_n\right], 0\} = \max\{F_n, 0\} = F_n^+$, so (F_n^+) is a submartingale. Similarly we show that (F_n^-) is also a submartingale.

2. We have $A_n - B_n = F_n - M_n + N_n$. The random process $(A_n - B_n)$ is a $\mathcal{F}$-martingale, it is also $\mathcal{F}$-predictable. It is therefore constant a.s., equals to $A_0 - B_0 = 0$. Indeed, for all $n \in \mathbb{N}$,

$$A_n - B_n = \mathbb{E}\left[A_{n+1} - B_{n+1} \middle| \mathcal{F}_n\right] = A_{n+1} - B_{n+1}.$$

3. The sequence (A_n) is the predictable process of a submartingale so it is non-decreasing a.s. It is non negative so it converges with probability 1 in $[0, +\infty]$. We set $A_\infty = \lim_{n\to\infty} A_n$. By non-decreasing limit, we get $\mathbb{E}\left[A_\infty\right] = \lim_{n\to\infty} \mathbb{E}\left[A_n\right]$. For all $n \in \mathbb{N}$,

$$\mathbb{E}\left[A_n\right] \leq \sup_k \mathbb{E}\left[|F_k|\right] - \mathbb{E}\left[M_0\right] < +\infty.$$

   We have that $\mathbb{E}\left[A_\infty\right] < \infty$, the r.v. A_∞ is finite a.s. because integrable and $\mathbb{E}\left[|A_\infty - A_n|\right] = \mathbb{E}\left[A_\infty\right] - \mathbb{E}\left[A_n\right] \to 0$ when $n \to +\infty$.

4. The process $(\mathbb{E}\left[A_n \middle| \mathcal{F}_n\right])$ is a regular martingale so that $(F_n^\oplus)_{n\in\mathbb{N}}$ and $(F_n^\ominus)_{n\in\mathbb{N}}$ are two martingales as sums of $\mathcal{F}$-martingales, and $F_n^\oplus = F_n^+ + \mathbb{E}\left[A_\infty - A_n \middle| \mathcal{F}_n\right] \geq 0$ (because $A_n \leq A_\infty$). Similarly we have $F_n^\ominus \geq 0$.

5. We have $F_n^{\oplus} - F_n^{\ominus} = M_n - N_n = F_n^{+} - F_n^{-} = F_n$ because $A_n = B_n$.

Solution of Exercise 6.19

Solution of Exercise 6.20

1. $S_0 = 0$ and $S_n = S_{n-1} + \epsilon_n$ so that $\Delta\langle S\rangle_n = \mathbb{E}\left[(\Delta S_n)^2 \middle| \mathcal{F}_{n-1}\right] = \mathbb{E}\left[\epsilon_n^2 \middle| \mathcal{F}_{n-1}\right] = \epsilon_n^2 = 1$ and $\langle S\rangle_n = n$. Therefore (Doob) $M_n = S_n^2 - n$ is a martingale.

2. Since $((M_n, \mathcal{F}_n))_{n\in\mathbb{N}}$ is a martingale and $T = \inf\{n \geq 0, S_n \notin]-a, a[\}$ is a stopping time, we have $\mathbb{E}\left[M_{T\wedge n}\right] = \mathbb{E}\left[S_{T\wedge n}^2 - (T\wedge n)\right] = \mathbb{E}\left[M_0\right] = 0$ so that $\mathbb{E}\left[S_{T\wedge n}^2\right] = \mathbb{E}\left[T\wedge n\right]$. We have $S_{T\wedge n}^2 \leq S_T^2 = a^2$ so by dominated convergence $\mathbb{E}\left[S_{T\wedge n}^2\right] \to \mathbb{E}\left[S_T^2\right]$. Since $\mathbb{E}\left[T\wedge n\right] = \mathbb{E}\left[S_{T\wedge n}^2\right] \leq a^2$, we have by monotone convergence that $\mathbb{E}\left[T\wedge n\right] \to \mathbb{E}\left[T\right]$ and finally $\mathbb{E}\left[S_T^2\right] = \mathbb{E}\left[T\right] = a^2$.

3. We have for all $n \geq 1$,

$$\begin{aligned} Y_n &= S_n^4 - 6nS_n^2 + bn^2 + cn \\ Y_{n-1} &= S_{n-1}^4 - 6(n-1)S_{n-1}^2 + b(n-1)^2 + c(n-1) \\ \Delta Y_n &= \Delta S_n^4 - 6nS_n^2 + 6(n-1)S_{n-1}^2 + (2n-1)n + c. \end{aligned}$$

Moreover $M_n^2 = S_n^4 - 2nS_n^2 + n^2$ so that $\mathbb{E}\left[\Delta(M_n^2) \middle| \mathcal{F}_{n-1}\right] = 4S_{n-1}^2, \mathbb{E}\left[\Delta(S_n)^4 \middle| \mathcal{F}_{n-1}\right] = 6S_{n-1}^2 + 1$ and $\mathbb{E}\left[\Delta Y_n \middle| \mathcal{F}_{n-1}\right] = (2b-6)n + (1+c-b)$ must be equal to 0. We obtain $b = 3$ and $c = 2$.

4. We have $Y_n = S_n^4 - 6nS_n^2 + 3n^2 + 2n$ and $\mathbb{E}\left[Y_{T\wedge n}\right] = \mathbb{E}\left[Y_0\right] = 0$ so that

$$3\mathbb{E}\left[(T\wedge n)^2\right] + 2\mathbb{E}\left[T\wedge n\right] = 6\mathbb{E}\left[(T\wedge n)S_{T\wedge n}^2\right] - \mathbb{E}\left[S_{T\wedge n}^4\right].$$

For all $k \leq T, |S_k| \leq a$ so that $|S_{T\wedge n}|^4 \leq a^4$ and by dominated convergence we get,

$$\lim_{n\to\infty} \mathbb{E}\left[S_{T\wedge n}^4\right] = \mathbb{E}\left[\lim_{n\to\infty} S_{T\wedge n}^4\right] = \mathbb{E}\left[S_T^4\right] = a^4.$$

$|(T\wedge n)S_{T\wedge n}^2| \leq Ta^2$ and $\mathbb{E}\left[T\right] < +\infty$ so by dominated convergence, we get

$$\lim_{n\to\infty} \mathbb{E}\left[(T\wedge n)S_{T\wedge n}^2\right] = \mathbb{E}\left[\lim_{n\to\infty}(T\wedge n)S_{T\wedge n}^2\right] = \mathbb{E}\left[TS_T^2\right] = a^2\mathbb{E}\left[T\right] = a^4.$$

Finally we obtain $3\mathbb{E}\left[T^2\right] + 2a^2 = 6a^4 - a^4 = 5a^4$ and $\mathbb{E}\left[T^2\right] = (a^2/3)(5a^2 - 2)$.

Solution of Exercise 6.21

1. Note that $(Z_k^2 \wedge m)X_k^2 \in L^1$ for all k and $m \in \mathbb{R}_+$ and $\mathbb{E}\left[(Z_k^2 \wedge m)X_k^2 \middle| \mathcal{F}_{k-1}\right] = (Z_k^2 \wedge m)V_k$. Thus $\mathbb{E}\left[(Z_k^2 \wedge m)X_k^2\right] = \mathbb{E}\left[(Z_k^2 \wedge m)V_k\right]$ and by monotone convergence, $\mathbb{E}\left[Z_k^2 X_k^2\right] = \lim_{m\to\infty} \mathbb{E}\left[(Z_k^2 \wedge m)V_k\right]$. Since, for all n, $\mathbb{E}[A_n] < +\infty$, we have that $Z_k^2 V_k$ is L^1 for all k and we thus conclude that $Z_k^2 X_k^2 \in L^1$ for all k. We then easily get, since Z_n is $\mathcal{F}_{n-1}$-mesurable and $\mathbb{E}\left[X_n \middle| \mathcal{F}_{n-1}\right] = 0$, that M_n is a square-integrable martingale.

2. The predictable quadratic variation is $\langle M\rangle_n = \sum_{k=1}^n \mathbb{E}\left[(\Delta M_k)^2 \middle| \mathcal{F}_{k-1}\right]$. For $1 \leq k \leq n$, $\mathbb{E}\left[(\Delta M_k)^2 \middle| \mathcal{F}_{k-1}\right] = \mathbb{E}\left[Z_k^2 X_k^2 \middle| \mathcal{F}_{k-1}\right] = Z_k^2 \mathbb{E}\left[X_k^2 \middle| \mathcal{F}_{k-1}\right] = Z_k^2 V_k$, so $\langle M\rangle_n = A_n$ and therefore $Y_n = M_n^2 - A_n$ is a martingale.

3. Thanks to the optional stopping theorem, we have $\mathbb{E}\left[M_\nu\right] = \mathbb{E}\left[M_0\right] = 0$ and $\mathbb{E}\left[Y_\nu\right] = \mathbb{E}\left[Y_0\right] = 0$, i.e., $\mathbb{E}\left[M_\nu^2\right] = \mathbb{E}\left[A_\nu\right]$.

4. Denote by $B_n = M_{\nu\wedge n}$. By the optional stopping theorem, $B_n^2 = M_{\nu\wedge n}^2$ has the Doob decomposition by stopping the components of the Doob decomposition of M_n^2, namely, its martingale part is $Y_{\nu\wedge n}$ and its predictable part is $A_{\nu\wedge n}$. with $\mathbb{E}[Y_{\nu\wedge n}] = \mathbb{E}[Y_0] = 0$ and $\mathbb{E}[A_{\nu\wedge n}] \leq \mathbb{E}[A_\nu] < +\infty$ which implies $\mathbb{E}[B_n^2] < +\infty$. Therefore, (B_n) converges a.s. and in L^2 to a random variable M_∞. On $\{\nu = +\infty\}$ we have $M_{\nu\wedge n} = M_n \to M_\infty$.

(a) Using the Doob decomposition of B_n^2, we get that (B_n) is bounded in L^2 if and only if $(A_{\nu\wedge n})$ is bounded in L^1. Since $(A_{\nu\wedge n})$ is non-negative and non-decreasing a.s., the monotone convergence theorem gives that $\mathbb{E}\left[|A_{\nu\wedge n}|\right] = \mathbb{E}\left[A_{\nu\wedge n}\right]$ converges to $\mathbb{E}\left[A_\nu\right]$ as $n \to \infty$ in a non-decreasing way. Since $\mathbb{E}\left[A_\nu\right]$ is finite we get that (B_n) is bounded in L^2. Thus it converges a.s. and in L^2. On $\{\nu = +\infty\}$ we get that $M_{\nu\wedge n} = M_n \to M_\infty$ a.s.

(b) From the previous question, we have $M_{\nu\wedge n} \to M_\nu$ as $n \to \infty$ on $\{\nu = \infty\}$ a.s. Obviously we also have $M_{\nu\wedge n} \to M_\nu$ as $n \to \infty$ on $\{\nu < \infty\}$. So we get that $M_{\nu\wedge n} \to M_\nu$ as $n \to \infty$ a.s. On the other hand since $M_{\nu\wedge n}$ also converges in L^2, we get that $M_{\nu\wedge n} \to M_\nu$ as $n \to \infty$ in L^2. By Question 3, we know that $\mathbb{E}\left[M_{\nu\wedge n}\right] = 0$ and $\mathbb{E}\left[M_{\nu\wedge n}^2\right] = \mathbb{E}\left[A_{n\wedge\nu}\right]$ for all n. Convergence in L^2 implies the convergence of the two first moments to those of the limits so we get that $\mathbb{E}\left[M_\nu\right] = 0$ and $\mathbb{E}\left[M_\nu^2\right] = \lim_{n\to\infty} \mathbb{E}\left[A_{n\wedge\nu}\right] = \mathbb{E}\left[A_\nu\right]$.

(c) For all $n \geq 1$, the maximal inequality for non-negative submartingales give that, for all $\rho > 0$,

$$\mathbb{P}\left(\max_{1\leq k\leq n} M_{k\wedge\nu}^2 > \rho^2\right) \leq \frac{1}{\rho^2}\mathbb{E}\left[M_{n\wedge\nu}^2\right] \ .$$

Since

$$\bigcup_n \left\{\max_{1\leq k\leq n} M_{k\wedge\nu}^2 > \rho^2\right\} = \left\{\sup_{n\leq\nu} |M_n| > \rho\right\} \ ,$$

and the displayed union is that of an increasing sequence of sets, we have

$$\mathbb{P}\left(\sup_{n\leq\nu} |M_n| > \rho\right) = \lim_{n\to\infty} \mathbb{P}\left(\max_{1\leq k\leq n} M_{k\wedge\nu}^2 > \rho^2\right) \leq \frac{1}{\rho^2} \lim_{n\to\infty} \mathbb{E}\left[M_{n\wedge\nu}^2\right] = \frac{1}{\rho^2}\mathbb{E}\left[A_\nu\right] \ ,$$

where we used the previous question in the last convergence. The conclusion follows from this inequality and applying

$$\mathbb{P}\left(\sup_{n\leq\nu} |M_n| \geq \rho\right) = \lim_{\rho'\uparrow\rho} \mathbb{P}\left(\sup_{n\leq\nu} |M_n| > \rho'\right) \ .$$

Solution of Exercise 6.22

1. Since by assumption $\mathcal{F}_n^X$ is independent of $\mathcal{T}$, we get that $\mathbb{E}\left[\mathbb{1}_A \middle| \mathcal{F}_n^X\right] = \mathbb{P}(A)$.

2. $Y_n = \mathbb{E}\left[\mathbb{1}_A \middle| \mathcal{F}_n^X\right]$ defines a closed martingale. So it is regular and using Theorem 6.5.10(b) we have

$$Y_\infty = \mathbb{E}\left[\mathbb{1}_A \middle| \bigvee_n \mathcal{F}_n^X\right] \qquad \mathbb{P}\text{-a.s.}$$

   By the previous question, we have $Y_\infty = \mathbb{P}(A)$ $\mathbb{P}$-a.s. and since $A \in \bigvee_{n\in\mathbb{N}} \mathcal{F}_n^X$, we get

$$\mathbb{P}(A) = \mathbb{1}_A \qquad \mathbb{P}\text{-a.s.}$$

   This implies that $\mathbb{P}(A)$ takes value 0 or 1.