Mathematical Foundations for Finance

Exercise sheet 1

This sheet contains material which is ***fundamental*** *for this course and* ***assumed to be known****. An exception to this are questions 5 and 6 which are technically more involved as well as the coding question 7 which is not part of the examinable material. We consider questions 5, 6 and 7 as BONUS problems. Results and facts from measure-theoretic probability theory will be often used in the lectures and exercise classes. To get an idea of the prerequesites, read the Appendix of your lecture notes (pages 129 to 136). The material* ***should be familiar*** *and in case it is not, you are expected to catch up as quicly as possible. The book Probability Essentials by Jean Jacod and Philip Protter contains all results needed and can be downloaded for free from Springer (within the ETH network or using VPN). A possible alternative to the above textbook are the ETH lecture notes for the standard course on Probability Theory, accessible from the MFF course website.*
Please upload your solutions until Wednesday, 29/09/2021, 12:00 using the link on the course website.

Exercise 1.1 Let $(\Omega, \mathcal{F}, P)$ be a probability space with $\Omega := \{UU, UD, DD, DU\}$, $\mathcal{F} := 2^\Omega$ and P defined by $P[\omega] := 1/4$ for all $\omega \in \Omega$. Let $Y_1, Y_2 : \Omega \to \mathbb{R}$ be two random variables with $Y_1(UU) = Y_1(UD) := 2$, $Y_1(DD) = Y_1(DU) := 1/2$, $Y_2(UU) = Y_2(DU) := 2$ and $Y_2(DD) = Y_2(UD) := 1/2$. Define the process $X = (X_k)_{k=0,1,2}$ by

$$\begin{aligned} X_0(\omega) &= 8 \quad \text{for all } \omega \in \Omega, \\ X_k(\omega) &= X_0(\omega) \prod_{i=1}^{k} Y_i(\omega) \quad \text{for } k = 1, 2. \end{aligned}$$

(a) Write down explicitly the sequences of σ-fields $\mathbb{F} = (\mathcal{F}_k)_{k=0,1,2}$ and $\mathbb{G} = (\mathcal{G}_k)_{k=0,1,2}$ defined by $\mathcal{F}_k := \sigma(X_i, 0 \le i \le k)$ and $\mathcal{G}_k := \sigma(X_k)$, $k = 0, 1, 2$.

(b) Show that $Z : \Omega \to \mathbb{R}$ defined by $Z(\omega) := 2X_1(\omega) + 1$ is $\mathcal{G}_1$-measurable.

(c) Are $\mathbb{F}$ and $\mathbb{G}$ filtrations on $(\Omega, \mathcal{F})$? Why or why not?

(d) Is X adapted to $\mathbb{F}$ or $\mathbb{G}$ (in case any of the former is a filtration on $(\Omega, \mathcal{F})$)?

(e) Try to give financial interpretations for X and $\mathbb{F}$.

Solution 1.1

(a) We have that $\mathcal{F}_0 = \mathcal{G}_0$ by definition with $\mathcal{F}_0 = \{\{X_0 = 8\}, \{X_0 \neq 8\}\} = \{\emptyset, \Omega\}$. Since $\mathcal{F}_0 = \mathcal{G}_0$ are contained in any other σ-field (by the definition of σ-field), we also have that $\mathcal{F}_1 = \sigma(X_0, X_1) = \sigma(X_1) = \mathcal{G}_1$. Furthermore,

$$\begin{aligned} \mathcal{G}_1 = \sigma(X_1) &= \big\{\{X_1 = 16\}, \{X_1 = 4\}, \{X_1 = 16\} \cup \{X_1 = 4\}, \{X_1 \neq 16, X_1 \neq 4\}\big\} \\ &= \big\{\{Y_1 = 2\}, \{Y_1 = 1/2\}, \{Y_1 = 2\} \cup \{Y_1 = 1/2\}, \{Y_1 \neq 2, Y_1 \neq 1/2\}\big\} \\ &= \big\{\{Y_1 = 2\}, \{Y_1 = 1/2\}, \Omega, \emptyset\big\} \\ &= \big\{\{UU, UD\}, \{DD, DU\}, \Omega, \emptyset\big\}, \end{aligned}$$

and

$$\begin{aligned}\mathcal{G}_2 = \sigma(X_2) &= \big\{\{X_2 = 32\}, \{X_2 = 8\}, \{X_2 = 2\}, \{X_2 = 32\} \cup \{X_2 = 8\},\\ &\quad \{X_2 = 32\} \cup \{X_2 = 2\}, \{X_2 = 8\} \cup \{X_2 = 2\}, \Omega, \emptyset\big\}\\ &= \big\{\{Y_1 = 2, Y_2 = 2\}, \{Y_1 = 1/2, Y_2 = 2\} \cup \{Y_1 = 2, Y_2 = 1/2\},\\ &\quad \{Y_1 = 1/2, Y_2 = 1/2\},\\ &\quad \{Y_1 = 2, Y_2 = 2\} \cup \{Y_1 = 1/2, Y_2 = 2\} \cup \{Y_1 = 2, Y_2 = 1/2\},\\ &\quad \{Y_1 = 2, Y_2 = 2\} \cup \{Y_1 = 1/2, Y_2 = 1/2\},\\ &\quad \{Y_1 = 1/2, Y_2 = 2\} \cup \{Y_1 = 2, Y_2 = 1/2\} \cup \{Y_1 = 1/2, Y_2 = 1/2\}, \Omega, \emptyset\big\},\\ &= \big\{\{UU\}, \{DU, UD\}, \{DD\}, \{UU, DU, UD\}, \{UU, DD\}, \{DU, UD, DD\},\\ &\quad \Omega, \emptyset\big\},\end{aligned}$$

$$\mathcal{F}_2 = \sigma(X_1, X_2) = \ \mathcal{F}.$$

(b) Since $g(x) = 2x + 1$ is a continuous function, we immediately get that $Z = g(X_1)$ is $\sigma(X_1)$-measurable. One could also argue by writing out $\sigma(Z)$ explicitly and showing that $\sigma(Z) \subseteq \sigma(X_1)$. We have that

$$\begin{aligned}\sigma(Z) &= \big\{\{Z = 33\}, \{Z = 9\}, \{Z = 33\} \cup \{Z = 9\}, \{Z \neq 33\, Z \neq 9\}\big\}\\ &= \big\{\{Z = 33\}, \{Z = 9\}, \Omega, \emptyset\big\}\\ &= \big\{\{UU, UD\}, \{DD, DU\}, \Omega, \emptyset\big\} = \sigma(X_1).\end{aligned}$$

Yet another approach would be to show that all sets of the form $\{Z \leq c\}$ for $c \in \mathbb{R}$ lie in $\sigma(X_1)$. This works because X_1 is clearly measurable with respect to $\sigma(X_1)$ (by definition). Indeed,

$$\{Z \leq c\} = \{2X_1 + 1 \leq c\} = \left\{X_1 \leq \frac{c-1}{2}\right\} \in \sigma(X_1),$$

since X_1 is $\sigma(X_1)$-measurable.

(c) We know that a filtration $\mathbb{H} = (\mathcal{H}_k)_{k=1,\dots,T}$, $T \in \mathbb{N}$ on a measurable space $(\Omega, \mathcal{F})$ is an increasing family of σ-fields $\mathcal{H}_k \subseteq \mathcal{F}$ in the sense that $\mathcal{H}_k \subseteq \mathcal{H}_n$ for $k \leq n$. $\mathbb{F}$ clearly forms a filtration because $\mathcal{F}_0 = \{\emptyset, \Omega\} \subseteq \mathcal{F}_1$ since $\mathcal{F}_0$ is contained in any σ-field and $\mathcal{F}_1 \subseteq \mathcal{F}_2$ because $\sigma(X_1) \subseteq \sigma(X_1, X_2)$. On the other hand, $\mathbb{G}$ does not form a filtration since $\{DD, DU\} \in \mathcal{G}_1$ but at the same time $\{DD, DU\} \notin \mathcal{G}_2$. Therefore $\mathcal{G}_1 \nsubseteq \mathcal{G}_2$.

(d) Since $\mathbb{G}$ does not form a filtration, we are only interested in whether X is adapted to $\mathbb{F}$. In order to decide this, we need to check that X_k is $\mathcal{F}_k$-measurable for $k = 0, 1, 2$. This is, however, trivial since we have constructed $\mathcal{F}_1$ (by definition of σ-field generated by a random variable) as the smallest σ-field such that X_1 is $\mathcal{F}_1$-measurable, and $\mathcal{F}_2$ as the smallest σ-field such that X_1 and X_2 are $\mathcal{F}_2$-measurable. For this reason, $\{\sigma(X_i, 1 \leq i \leq k)\}_k$ is also called the *canonical filtration* for $(X_k)_k$.

(e) The process X can be interpreted as the price of a stock that is only allowed to move up or down by factors of 2 and 1/2 respectively in each period.

The filtration $\mathbb{F}$ can be thought of as the *cumulative* information that the stock price evolution provides us with over time.

Exercise 1.2 Let $(\Omega, \mathcal{F}, P)$ be a probability space and $X : \Omega \to \mathbb{R}$ a random variable with $X \geq 0$ P-a.s. Prove that $E[X] = 0$ implies that $X = 0$ P-a.s.
Hint: Find a way to use the monotone convergence theorem.

Solution 1.2 Since $X \geq 0$ P-a.s., it is sufficient to show that $P[X > 0] = 0$. Markov's inequality together with our assumption that $E[X] = 0$ imply that

$$P\left[X \geq \frac{1}{n}\right] \leq nE[X] = 0 \quad \text{for all } n \in \mathbb{N}. \tag{1}$$

But since

$$\left\{X \geq \frac{1}{n}\right\} \subseteq \left\{X \geq \frac{1}{n+1}\right\} \quad \text{for all } n \in \mathbb{N}, \tag{2}$$

we obtain that

$$\mathbb{1}_{\{X \geq \frac{1}{n}\}} \leq \mathbb{1}_{\{X \geq \frac{1}{n+1}\}} \quad P\text{-a.s. for all } n \in \mathbb{N}.$$

So $\mathbb{1}_{\{X \geq \frac{1}{n}\}}$, $n \in \mathbb{N}$ form a sequence of almost surely increasing functions and therefore

$$\begin{aligned}\lim_{n\to\infty} P\left[X \geq \frac{1}{n}\right] &= \lim_{n\to\infty} E\left[\mathbb{1}_{\{X \geq \frac{1}{n}\}}\right] = E\left[\lim_{n\to\infty} \mathbb{1}_{\{X \geq \frac{1}{n}\}}\right] = E\left[\lim_{n\to\infty} \mathbb{1}_{\left\{\cup_{k=1}^{n}\{X \geq \frac{1}{n}\}\right\}}\right] \\ &= E\left[\mathbb{1}_{\left\{\lim_{n\to\infty} \cup_{k=1}^{n}\{X \geq \frac{1}{n}\}\right\}}\right] = E\left[\mathbb{1}_{\{X>0\}}\right] = P[X > 0],\end{aligned}$$

where the second equality follows from the monotone convergence theorem and the third equality follows from (2). Combining the above with (1), we conclude that

$$P[X > 0] = \lim_{n\to\infty} P\left[X \geq \frac{1}{n}\right] = 0.$$

Exercise 1.3 Let $(\Omega, \mathcal{F}, P)$ be a probability space, X an integrable random variable and $\mathcal{G} \subseteq \mathcal{F}$ a σ-field. Then the P-a.s. unique random variable Z such that

- Z is $\mathcal{G}$-measurable and integrable,
- $E[X\mathbb{1}_A] = E[Z\mathbb{1}_A]$ for all $A \in \mathcal{G}$,

is called *conditional expectation of X given $\mathcal{G}$* and is denoted by $E[X \,|\, \mathcal{G}]$. (This is the formal definition of conditional expectation of X given $\mathcal{G}$; see Section 8.2 in the lecture notes.)

(a) Use the definition above to show that if X is $\mathcal{G}$-measurable, then $E[X \,|\, \mathcal{G}] = X$ P-a.s.

(b) Use the definition of conditional expectation to show that $E[E[X \,|\, \mathcal{G}]] = E[X]$.

(c) Use the definition of conditional expectation to show that if $P[A] \in \{0, 1\}$ for all $A \in \mathcal{G}$, i.e. if $\mathcal{G}$ is P-trivial, then $E[X \,|\, \mathcal{G}] = E[X]$ P-a.s.

(d) Consider another integrable random variable Y on $(\Omega, \mathcal{F}, P)$, and two constants $a, b \in \mathbb{R}$. Use the definition of conditional expectation to show that $E[aX + bY \,|\, \mathcal{G}] = aE[X \,|\, \mathcal{G}] + bE[Y \,|\, \mathcal{G}]$ P-a.s.

(e) Suppose that $\mathcal{G}$ is generated by a finite partition of Ω, i.e. there exists a collection $(A_i)_{i=1}^n$ of sets $A_i \in \mathcal{F}$ such that $\bigcup_{i=1}^n A_i = \Omega$, $A_i \cap A_j = \emptyset$ for $i \neq j$ and $\mathcal{G} = \sigma(A_1, \dots, A_n)$. Additionally, assume that $P[A_i] > 0$ for all $i = 1, \dots, n$. Use the definition of conditional expectation to show that

$$E[X \,|\, \mathcal{G}] = \sum_{i=1}^{n} E[X \,|\, A_i] \mathbb{1}_{A_i} \qquad P\text{-a.s.}$$

Hint 1: Recall that $E[X \,|\, A_i] = E[X\mathbb{1}_{A_i}]/P[A_i]$ and try to write X as a sum of random variables each of which only takes non-zero values on a single A_i.
Hint 2: Check that any set $A \in \mathcal{G}$ is of the form $\cup_{j\in J} A_j$ for some $J \subseteq \{1, \dots, n\}$.

Solution 1.3

(a) X is $\mathcal{G}$-measurable and integrable by assumption, so the first requirement in the definition of conditional expectation is satisfied for $Z = X$. Moreover, we clearly have that $E[X\mathbb{1}_A] = E[X\mathbb{1}_A]$ for all $A \in \mathcal{G}$, hence $E[X \mid \mathcal{G}] = X$ P-a.s.

(b) In the definition of conditional expectation set $A = \Omega$. We have that $E[E[X \mid \mathcal{G}]] = E[E[X \mid \mathcal{G}]\mathbb{1}_\Omega] = E[X\mathbb{1}_\Omega] = E[X]$.

(c) Since $|E[X]| \le E[|X|]$ by Jensen's inequality and $E[|X|] < \infty$ since X is integrable by assumption, we have that $E[X]$ is integrable as well. $E[X]$ is also trivially $\mathcal{G}$-measurable since it is a constant random variable. Moreover, in this setting, $A \in \mathcal{G}$ only if $P[A] = 0$ or $P[A] = 1$. Noting that

$$\begin{aligned} E[X\mathbb{1}_A] &= 0 = E[E[X]\mathbb{1}_A] && \forall\, A \in \mathcal{G} \text{ such that } P[A] = 0,\\ E[X\mathbb{1}_A] &= E[X] = E[E[X]\mathbb{1}_A] && \forall\, A \in \mathcal{G} \text{ such that } P[A] = 1, \end{aligned}$$

we obtain $E[X \mid \mathcal{G}] = E[X]$ P-a.s.

(d) By the definition of conditional expectation, we have that $E[X \mid \mathcal{G}]$ and $E[Y \mid \mathcal{G}]$ are $\mathcal{G}$-measurable and integrable, hence the same holds for $aE[X \mid \mathcal{G}] + bE[Y \mid \mathcal{G}]$. Choosing any $A \in \mathcal{G}$ we can compute that

$$\begin{aligned} E\left[\big(aE[X \mid \mathcal{G}] + bE[Y \mid \mathcal{G}]\big)\mathbb{1}_A\right] &= aE[E[X \mid \mathcal{G}]\mathbb{1}_A] + bE[E[Y \mid \mathcal{G}]\mathbb{1}_A]\\ &= aE[X\mathbb{1}_A] + bE[Y\mathbb{1}_A] = E\left[\big(aX + bY\big)\mathbb{1}_A\right], \end{aligned}$$

where the first equality uses the linearity of (classical) expectation and the second uses the definition of $E[X \mid \mathcal{G}]$ and $E[Y \mid \mathcal{G}]$. This shows that $E[aX + bY \mid \mathcal{G}] = aE[X \mid \mathcal{G}] + bE[Y \mid \mathcal{G}]$ P-a.s.

(e) First recall that $E[X \mid A_i] = E[X\mathbb{1}_{A_i}]/P[A_i]$. Using that $X = \sum_{i=1}^n X\mathbb{1}_{A_i}$, we get by (d) that

$$E[X \mid \mathcal{G}] = \sum_{i=1}^n E[X\mathbb{1}_{A_i} \mid \mathcal{G}] \quad P\text{-a.s.},$$

and hence we only have to show that $E[X\mathbb{1}_{A_i} \mid \mathcal{G}] = \frac{E[X\mathbb{1}_{A_i}]}{P[A_i]}\mathbb{1}_{A_i}$ P-a.s. for each $i \in \{1, \dots, n\}$. Since $A_i \in \mathcal{G}$ and $E[X \mid A_i] = E[X\mathbb{1}_{A_i}]/P[A_i] \in \mathbb{R}$, we already know that $E[X \mid A_i]\mathbb{1}_{A_i}$ is $\mathcal{G}$-measurable and integrable. One can verify that the family of sets $A = \bigcup_{j\in J} A_j$ for $J \in 2^{\{1,\dots,n\}}$ (the power set of $\{1, \dots, n\}$) forms a σ-field. Let's denote this σ-field by $\tilde{\mathcal{G}}$. Since we clearly have that $A_i \in \tilde{\mathcal{G}}$ for all $i \in \{1, \dots, n\}$, we get that $\tilde{\mathcal{G}} \supseteq \mathcal{G}$, which for any $A \in \mathcal{G}$ implies that $A = \bigcup_{j\in J} A_j$ for some $J \subseteq \{1, \dots, n\}$. For any such $A \in \mathcal{G}$ we have that

$$\mathbb{1}_{A_i}\mathbb{1}_A = \begin{cases} \mathbb{1}_{A_i} & \text{if } i \in J,\\ 0, & \text{else.} \end{cases}$$

Hence we can then compute

$$E\left[\left(\frac{E[X\mathbb{1}_{A_i}]}{P[A_i]}\mathbb{1}_{A_i}\right)\mathbb{1}_A\right] = \begin{cases} E[X\mathbb{1}_{A_i}]\frac{P[A_i]}{P[A_i]} = E[X\mathbb{1}_{A_i}] & \text{if } i \in J,\\ 0 & \text{else.} \end{cases}$$

On the other hand, we have that

$$E[X\mathbb{1}_{A_i}\mathbb{1}_A] = \begin{cases} E[X\mathbb{1}_{A_i}] & \text{if } i \in J,\\ 0 & \text{else.} \end{cases}$$

This shows that $E[X\mathbb{1}_{A_i} \mid \mathcal{G}] = \frac{E[X\mathbb{1}_{A_i}]}{P[A_i]}\mathbb{1}_{A_i}$ P-a.s. and concludes the proof.

Exercise 1.4 Let (U, V) be a two dimensional random vector. Suppose that the distribution of (U, V) admits a density function $f_{U,V}$ with respect to the Lebesgue measure. Let

$$f_U(u) = \int_{\mathbb{R}} f_{U,V}(u, v)dv$$

denote the marginal distribution of U and define the conditional density of V given U by

$$f_{V|U}(v|u) := \frac{f_{U,V}(u, v)}{f_U(u)}.$$

Consider a Borel measurable function $h : \mathbb{R} \to \mathbb{R}$ such that $h(V) \in L^1$ and define

$$g(u) := \int h(v)f_{V|U}(v|u)dv.$$

Show that $E[h(V) \,|\, U] = g(U)$.

Solution 1.4 We need to check the three properties of the conditional expectation.

- $g(U) \in L^1$ since by Jensen's inequality,

$$E[|g(U))|] = E[|E[h(V) \,|\, U]\,|] \le E[E[|h(V)| \,|\, U]] = E[|h(V)|] < \infty$$

  since $h(V) \in L^1$.

- $g(U)$ is $\sigma(U)$-measurable. This can be shown using the standard "measure theoretic induction" argument. We first note that $g(\cdot) : \mathbb{R} \to \mathbb{R}$ is $\mathcal{B}(\mathbb{R})$-measurable for $h = \mathbb{1}_B$ for any $B \in \mathcal{B}(\mathbb{R})$ because the pointwise limit of Riemann sums is measurable. By linearity it also holds for simple functions, i.e functions for the form $f = \sum_{k=1}^n a_k \mathbb{1}_{B_k}$. It therefore also holds for positive measurable functions since every such functions is the increasing limit of a sequence of simple functions and hence we can apply Monotone convergence Theorem to exchange the limit and the integral. Finally by decompositing any measurable integrable function h into its negative and positive parts, we obtain the desired result. This shows that $g(\cdot) : \mathbb{R} \to \mathbb{R}$ is $\mathcal{B}(\mathbb{R})$-measurable and therefore $g \circ U : \Omega \to \mathbb{R}$ is $\sigma(U)$-measurable as the composition of measurable functions.

- Let $A \in \sigma(U)$. Then we have $A = \{\omega \in \Omega \quad s.t \quad U(\omega) \in B\}$ for some Borel set $B \in \mathcal{B}(\mathbb{R})$. Moreover we have

$$\begin{aligned}
E[h(V)\mathbb{1}_A] &= \int\int h(v)\mathbb{1}_{u\in B} f_{U,V}(u, v)dudv \\
&= \int\int h(v)\mathbb{1}_{u\in B} f_{V|U}(v|u)f_U(u)dudv \\
&= \int g(u)\mathbb{1}_{u\in B}du \\
&= E[g(U)\mathbb{1}_A]
\end{aligned}$$

  where in the third equality we used Fubini's Theorem and the definition of g.

Exercise 1.5 *The goal of this exercise is to show existence and uniqueness of the conditional expectation under the restrictive assumption of square integrability.*

Let $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-sigma-algebra of $\mathcal{F}$. Show that X admits a $\mathbb{P}$-almost surely unique conditional expectation, i.e that there exits a $\mathbb{P}$-almost surely unique random variable Y satisfying the following properties:

- Y is $\mathcal{G}$-measurable and integrable,
- $E[X\mathbb{1}_A] = E[Y\mathbb{1}_A]$ for all $A \in \mathcal{G}$.

The exercise involves Hilbert spaces and in particular an important result concerning orthogonal projections on convex closed subsets of a Hilbert space. For the sake of completeness we recall the definition of a Hilbert space and state this important result that you will need to use in question b).

A Hilbert space is a vector space H with an inner product $\langle f, g\rangle$ such that the norm defined by

$$||f|| := \sqrt{\langle f, f\rangle}$$

turns H into a complete metric space. In particular it can be shown that $L^2(\Omega, \mathcal{F}, \mathbb{P})$ is a Hilbert space with the inner product defined by

$$\langle X, Y\rangle := E[XY].$$

Note that this integral is well defined by the Cauchy-Schwartz inequality and indeed defines an inner product on $L^2(\Omega, \mathcal{F}, \mathbb{P})$ (good bonus exercise). Equipped with the norm induced by this inner product, it is not too difficult to show that L^2 is complete and hence a Hilbert space. One of the nicest properties of Hilbert spaces is that one can generalize the notion of orthogonal projections from finite dimensional euclidiean vector spaces to (potentially) infinite dimensional Hilbert spaces as suggested by the following theorem: Let Γ be a closed convex subset of a Hilbert space H. Then for any point $x \in H$, there exists a unique point $\pi_\Gamma(x) \in \Gamma$ such that

$$||x - \pi_\Gamma(x)|| = \inf_{g\in\Gamma} ||x - g||.$$

Moreover $x - \pi_\Gamma(x) \in \Gamma^\perp$, where $\Gamma^\perp$ denotes the orthogonal complement of Γ, i.e.

$$\Gamma^\perp = \{x \in H s.t. \langle x, g\rangle = 0 \quad \forall g \in \Gamma\}.$$

(a) Show that if two random variables Y_1 and Y_2 satisfy the above properties, then we have $\mathbb{P}(Y_1 = Y_2) = 1$. This shows the $\mathbb{P}$-almost uniqueness.

Remains to show the existence. To do this we will apply the orthogonal projection theorem on the closed convex subset $L^2(\Omega, \mathcal{G}, \mathbb{P})$ of the Hilbert space $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

(b) Show that $L^2(\Omega, \mathcal{G}, \mathbb{P})$ is convex in $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

(c) Show that $L^2(\Omega, \mathcal{G}, \mathbb{P})$ is closed in $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

(d) Deduce the existence of the conditional expectation of X given $\mathcal{G}$.

(e) Give a geometric interpretation of the above construction.

Solution 1.5

(a) Suppose that both Y_1 and Y_2 satisfy the properties of a conditional expectation. Then, clearly the event $A = \{Y_1 > Y_2\} \in \mathcal{G}$ and therefore by the second property of conditional expectation, we have

$$E[(Y_1 - Y_2)\mathbb{1}_A] = E[Y_1\mathbb{1}_A] - E[Y_2\mathbb{1}_A] = E[X\mathbb{1}_A] - E[X\mathbb{1}_A] = 0.$$

But by definition of A, we have $(Y_1 - Y_2)\mathbb{1}_A \geq 0 \quad \mathbb{P}$-a.s, and so by the result of question 3, $(Y_1 - Y_2)\mathbb{1}_A = 0 \quad \mathbb{P}$-a.s, hence $\mathbb{P}[A] = 0$ (because by definition of A, $(Y_1 - Y_2) \neq 0$. This proves that $Y_1 \leq Y_2 \quad \mathbb{P}$-a.s. By changing the role of Y_1 and Y_2, we get the desired result.

Alternative proof. Consider the event $A_k = \{Y_1 \geq Y_2 + \frac{1}{k}\}$ and note that $A_k \in \mathcal{G}$ since both Y_1 and Y_2 are $\mathcal{G}$-measurables. Let's apply the second condition defining a conditional expectation with $A = A_k$:

$$\begin{aligned} E[Y_2 \mathbb{1}_{A_k}] &= E[Y_1 \mathbb{1}_{A_k}] \\ &\geq E[Y_2 \mathbb{1}_{A_k}] + \frac{1}{k}\mathbb{P}[A_k] \end{aligned}$$

where in the last inequality we used the definition of A_k. Hence

$$\frac{1}{k}\mathbb{P}[A_k] \leq 0$$

and thus

$$\mathbb{P}[Y_1 \geq Y_2 + \frac{1}{k}] = 0.$$

This being true for any k, we get using Monotone Convergence Theorem $\mathbb{P}[Y_1 > Y_2] = 0$. The arguments are very similar to those in question 1.2 and hence it is a *very good* exercise to write down the details on your own. By changing the role of Y_1 and Y_2, we also get $\mathbb{P}[Y_2 > Y_1] = 0$ and hence $\mathbb{P}[Y_1 = Y_2] = 1$.

(b) Let Y_1 and Y_2 be two elements of $L^2(\Omega, \mathcal{G}, \mathbb{P})$ and let $\lambda \in [0, 1]$. We need to show that the convex combination $\lambda Y_1 + (1 - \lambda)Y_2$ is in $L^2(\Omega, \mathcal{G}, \mathbb{P})$. Both the square integrability and the measurability are trivial.

(c) Let $(Y_n)_{n\geq 0}$ a sequence of elements of $L^2(\Omega, \mathcal{G}, \mathbb{P})$ converging in L^2 to a random variable Y in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. We need to show that $Y \in L^2(\Omega, \mathcal{G}, \mathbb{P})$, i.e that Y is actually $\mathcal{G}$-measurable. Recall that convergence in L^2 implies convergence in probability which in turns implies the existence of a subsequence converging almost surely, hence there exists a subsequence $(n_k)_k$ such that $(Y_{n_k})_k$ converges almost surely to Y. By assumption the random variables Y_{n_k} are $\mathcal{G}$-measurable for all k and hence by a general result from measure therory, their almost sure limit is also $\mathcal{G}$-measurable. This shows the required closedness.

(d) Since $L^2(\Omega, \mathcal{G}, \mathbb{P})$ is a convex, closed subspace of the Hilbert space $L^2(\Omega, \mathcal{F}, \mathbb{P})$, there exists a projection

$$\pi : L^2(\Omega, \mathcal{F}, \mathbb{P}) \to L^2(\Omega, \mathcal{G}, \mathbb{P}).$$

We define the conditional expectation $E[X \mid \mathcal{G}]$ as

$$E[X \mid \mathcal{G}] := \pi(X)$$

and show that it satisfies the three properties appearing in the definition of conditional expectation. Measurability is clear by construction. Integrability follows by Hölder's inequality which implies the inclusion of L^2 in L^1. It thus only remains to show that

$$E[X \mathbb{1}_A] = E[E[X \mid \mathcal{G}] \mathbb{1}_A]$$

for all $A \in \mathcal{G}$. However this is a direct consequence of the definition of orthogonal projection as

$$E[X \mathbb{1}_A] - E[E[X \mid \mathcal{G}] \mathbb{1}_A] = E[(X - E[X \mid \mathcal{G}])\mathbb{1}_A]$$

is the scalar product of $(X - E[X \mid \mathcal{G}]) \in L^2(\Omega, \mathcal{G}, \mathbb{P})^\perp$ and $\mathbb{1}_A \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ and hence

$$E[(X - E[X \mid \mathcal{G}])\mathbb{1}_A] = 0.$$

(e) The conditional expectation for L^2 integrable random variables is constructed as the projection operator on $L^2(\Omega, \mathcal{G}, \mathbb{P})$. The conditional expectation $E[X \mid \mathcal{G}]$ is thus the best estimate of X given the information given by the sigma-algebra $\mathcal{G}$.

Exercise 1.6 *This is a **bonus** question on the Capital Asset Pricing Model (CAPM) and linear regression. This question is **not** part of the examinable material; it's purpose is just to motivate you to improve your programming skills. Programming skills are becoming more and more important especially for those of you wanting to work in the industry. The goal of the exercise is to compute the beta of a stock using linear regression.* The CAPM is a model that prices securities by examining the relationship between expected returns and risk. More precisely the model states that the return of a risky asset is given by

$$E[R_i] = r + \beta_i(E[R_m] - r) \tag{3}$$

where $E[R_i]$ and $E[R_m]$ are the expected returns of asset i and of the market respectively and r is the risk-free rate of interest. One can show that β_i has a closed from solution given by

$$\beta_i = \frac{\mathrm{Cov}(R_i, R_m)}{\mathrm{Var}(R_m)}$$

The Beta of an asset is thus a measure of the sensitivity of its returns relative to a market benchmark (usually a market index). We could in principle calculate the Beta of a given asset using the above formula, however in practice it is easier to think of the Beta as the slope of the linear regression (3) and estimate it using Ordinary Least Squares method.

In this exercise we will estimate the Beta of Facebook relative to the S&P500 market index using historical data from 02/11/2014 to 19/11/2017. The data can be downloaded from the course web page ('FB.csv' and 'SP500.csv'). Your task is to complete **either** the below Python code **or** the R code in order to perform an Ordinary Least Squares Regression with Statsmodel. If your code works well, you should get a Beta value close to 0.58 which is the Beta value of Facebook quoted on Yahoo Finance on the 19/11/2017.

```
###### Python Code ########
###########################

# import libraries
import pandas as pd
import statsmodels.api as sm

'''
Download monthly prices of Facebook and S&P 500 index from 2014 to 2017
CSV file downloaded from Yahoo File
start period: 02/11/2014
end period: 30/11/2017
period format: DD/MM/YEAR
'''

# Step 1: Use pandas read_csv method to load the two csv files downloaded
    from the course web page

fb = # TO DO #
sp_500 = # TO DO #

# joining the closing prices of the two datasets
monthly_prices = pd.concat([fb['Close'], sp_500['Close']], axis=1)
monthly_prices.columns = ['FB', 'SP500']

# check the head of the dataframe
print(monthly_prices.head())

# calculate monthly returns
monthly_returns = # TO DO #
```

```
clean_monthly_returns = monthly_returns.dropna(axis=0)

# split dependent and independent variable
X = # TO DO #
y = # TO DO #

# Add a constant to the independent value
X1 = sm.add_constant(X)

# make regression model
model = # TO DO #

# fit model and print results
results = # TO DO #
print(results.summary())
```

```
###### R Code ########
#####################

# CAPM linear regression
# read in the csv files
fb <- # TO DO #
sp_500 <- # TO DO #

# joining the closing prices of the two datasets
monthly_prices = cbind(fb['Close'], sp_500['Close'])
colnames(monthly_prices) = c('FB', 'SP500')

# check the head of the dataframe
head(monthly_prices)

# calculate monthly returns
# Denote n the number of time periods:
n <- nrow(monthly_prices)
monthly_returns <- ((monthly_prices[2:n, ] - monthly_prices[1:(n-1), ])/
    monthly_prices[1:(n-1), ])
# note that this is already the clean_monthly_returns from the Python
    code

# split dependent and independent variable
X = # TO DO #
y = # TO DO #

# the lm function automatically ads the intercept term
fit <- # TO DO #
summary(fit) # for interpretation see below
# Interpretation:
#   FB_returns = 0.020270 + 0.575091*SP500_returns

# visualize the results
coeff <- coefficients(fit)
plot(X,y, xlab = 'SP500 monthly returns', ylab = 'Facebook monthly
    returns')
abline(a=coeff[1], b=coeff[2], col=2, lwd=2)
```

Make sure you understand the summary of the linear regression and that in particular you can find the corresponding Beta value.

Solution 1.6

```
# import libraries
import pandas as pd
import statsmodels.api as sm

'''
Download monthly prices of Facebook and S&P 500 index from 2014 to 2017
CSV file downloaded from Yahoo File
start period: 02/11/2014
end period: 30/11/2017
period format: DD/MM/YEAR
'''
fb = pd.read_csv('FB.csv', parse_dates=True, index_col='Date',)
sp_500 = pd.read_csv('SP500.csv', parse_dates=True, index_col='Date')

# joining the closing prices of the two datasets
monthly_prices = pd.concat([fb['Close'], sp_500['Close']], axis=1)
monthly_prices.columns = ['FB', 'SP500']

# check the head of the dataframe
print(monthly_prices.head())

# calculate monthly returns
monthly_returns = monthly_prices.pct_change(1)
clean_monthly_returns = monthly_returns.dropna(axis=0)

# split dependent and independent variable
X = clean_monthly_returns['SP500']
y = clean_monthly_returns['FB']

# Add a constant to the independent value
X1 = sm.add_constant(X)

# make regression model
model = sm.OLS(y, X1)

# fit model and print results
results = model.fit()
print(results.summary())
```

For the ones preferring R, the solution should look like

```
###### R Code ########
#####################

# CAPM linear regression
# read in the csv files
fb <- read.csv('FB.csv', header = TRUE)
sp_500 <- read.csv('SP500.csv', header = TRUE)

# joining the closing prices of the two datasets
monthly_prices = cbind(fb['Close'], sp_500['Close'])
colnames(monthly_prices) = c('FB', 'SP500')

# check the head of the dataframe
head(monthly_prices)
```

```r
# calculate monthly returns
# Denote n the number of time periods:
n <- nrow(monthly_prices)
monthly_returns <- ((monthly_prices[2:n, ] - monthly_prices[1:(n-1), ])/
    monthly_prices[1:(n-1), ])
# note that this is already the clean_monthly_returns from the Python
    code

# split dependent and independent variable
X = as.matrix(monthly_returns['SP500'])
y = as.matrix(monthly_returns['FB'])

# the lm function automatically ads the intercept term
fit <- lm(y~X)
summary(fit) # for interpretation see below
# Interpretation:
#   FB_returns = 0.020270 + 0.575091*SP500_returns

# visualize the results
coeff <- coefficients(fit)
plot(X,y, xlab = 'SP500 monthly returns', ylab = 'Facebook monthly
    returns')
abline(a=coeff[1], b=coeff[2], col=2, lwd=2)
```

A more detailed solution is available on the course website explained how each quantity appearing in the R sumary output is calculated. Although not part of this course, basic knowledge on linear models is crucial for each quant.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                     FB   R-squared:                       0.101
Model:                            OLS   Adj. R-squared:                  0.074
Method:                 Least Squares   F-statistic:                     3.816
Date:                Tue, 19 Feb 2019   Prob (F-statistic):             0.0590
Time:                        10:49:51   Log-Likelihood:                 57.383
No. Observations:                  36   AIC:                            -110.8
Df Residuals:                      34   BIC:                            -107.6
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0203      0.009      2.330      0.026       0.003       0.038
^GSPC          0.5751      0.294      1.953      0.059      -0.023       1.173
==============================================================================
Omnibus:                        0.948   Durbin-Watson:                   2.208
Prob(Omnibus):                  0.623   Jarque-Bera (JB):                0.232
Skew:                          -0.074   Prob(JB):                        0.891
Kurtosis:                       3.364   Cond. No.                         34.9
==============================================================================
```

Figure 1: Summary of the OLS fit (Python): our regression model gives a Beta value of 0.5751 which is very close to the quoted Beta of 0.58

```
Call:
lm(formula = y ~ X)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.13589 -0.02702 -0.00284  0.02871  0.10217 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)  
(Intercept) 0.020270   0.008698   2.330   0.0259 *
X           0.575091   0.294397   1.953   0.0590 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.05057 on 34 degrees of freedom
Multiple R-squared:  0.1009,	Adjusted R-squared:  0.07447 
F-statistic: 3.816 on 1 and 34 DF,  p-value: 0.05903
```

Figure 2: Summary of the OLS fit (R): our regression model gives a Beta value of 0.5751 which is very close to the quoted Beta of 0.58

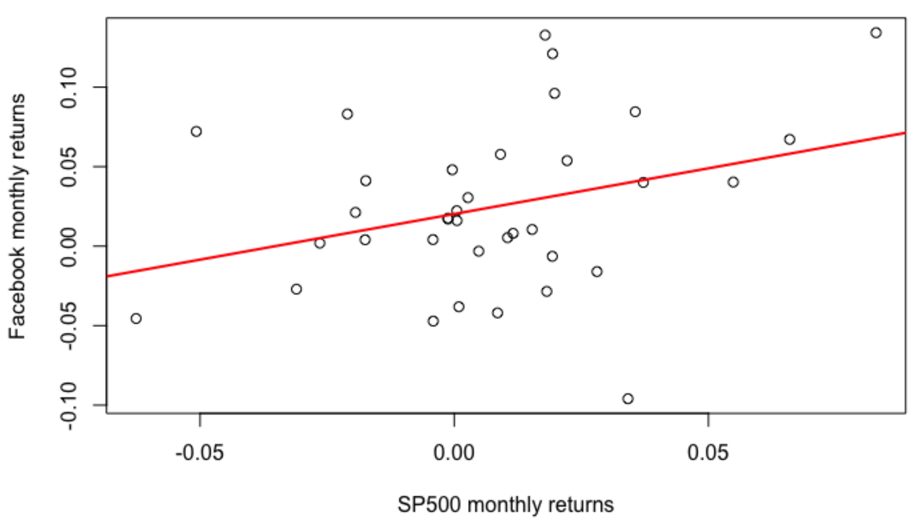

Figure 3: Plot of the fitted model