ADVANCED PROBABILITY
Contents

1 Conditional Expectation 5
   1.1 Discrete Case 6
   1.2 Existence and Uniqueness 7
   1.3 Conditional Jensen’s Inequalities 11
   1.4 Product Measures and Fubini’s Theorem 13
   1.5 Examples of Conditional Expectation 14
   1.6 Notation for Example Sheet 1 14

2 Discrete Time Martingales 17
   2.1 Optional Stopping 18
   2.2 Hitting Probabilities for a Simple Symmetric Random Walk 20
   2.3 Martingale Convergence Theorem 21
   2.4 Uniform Integrability 22
   2.5 Backwards Martingales 25
   2.6 Applications of Martingales 25
       2.6.1 Martingale proof of the Radon-Nikodym theorem 26

3 Stochastic Processes in Continuous Time 27
4 Bibliography 29
Conditional Expectation

Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space. \(\Omega\) is a set, \(\mathcal{F}\) is a \(\sigma\)-algebra on \(\Omega\), and \(\mathbb{P}\) is a probability measure on \((\Omega, \mathcal{F})\).

**Definition 1.1.** \(\mathcal{F}\) is a \(\sigma\)-algebra on \(\Omega\) if it satisfies

- \(\emptyset, \Omega \in \mathcal{F}\)
- \(A \in \mathcal{F} \implies A^c \in \mathcal{F}\)
- \((A_n)_{n \geq 0}\) is a collection of sets in \(\mathcal{F}\) then \(\bigcup_n A_n \in \mathcal{F}\).

**Definition 1.2.** \(\mathbb{P}\) is a probability measure on \((\Omega, \mathcal{F})\) if

- \(\mathbb{P}: \mathcal{F} \to [0, 1]\) is a set function.
- \(\mathbb{P}(\emptyset) = 0, \mathbb{P}(\Omega) = 1)\)
- \((A_n)_{n \geq 0}\) is a collection of pairwise disjoint sets in \(\mathcal{F}\), then \(\mathbb{P}(\bigcup_n A_n) = \sum_n \mathbb{P}(A_n)\).

**Definition 1.3.** The Borel \(\sigma\)-algebra \(\mathcal{B}(\mathbb{R})\) is the \(\sigma\)-algebra generated by the open sets of \(\mathbb{R}\). Call \(\mathcal{O}\) the collection of open subsets of \(\mathbb{R}\), then

\[
\mathcal{B}(\mathbb{R}) = \cap\{\xi: \xi \text{ is a sigma algebra containing } \mathcal{O}\} \quad (1.1)
\]

**Definition 1.4.** \(A\) a collection of subsets of \(\Omega\), then we write \(\sigma(A) = \cap\{\xi: \xi \text{ a sigma algebra containing } A\}\)

**Definition 1.5.** \(X\) is a random variable on \((\Omega, \mathcal{F})\) if \(X: \Omega \to \mathbb{R}\) is a function with the property that \(X^{-1}(V) \in \mathcal{F}\) for all \(V\) open sets in \(\mathbb{R}\).
Exercise 1.6. If $X$ is a random variable then $\{B \subseteq \mathbb{R}, X^{-1}(B) \in \mathcal{F}\}$ is a \(\sigma\)-algebra and contains \(\mathcal{B}(\mathbb{R})\).

If \((X_i, i \in I)\) is a collection of random variables, then we write 
\(\sigma(X_i, i \in I) = \sigma(\{\omega \in \Omega : X_i(\omega) \in B, i \in I, B \in \mathcal{B}(\mathbb{R})\})\) and it is the smallest \(\sigma\)-algebra that makes all the \(X_i\)'s measurable.

Definition 1.7. First we define it for the positive simple random variables.

\[
\mathbb{E}\left(\sum_{i=1}^{n} c_i 1(A_i)\right) = \sum_{i=1}^{n} \mathbb{P}(A_i). \tag{1.2}
\]

with \(c_i\) positive constants, \((A_i) \in \mathcal{F}\).

We can extend this to any positive random variable \(X \geq 0\) by approximation \(X\) as the limit of piecewise constant functions.

For a general \(X\), we write \(X = X^+ - X^-\) with \(X^+ = \max(X, 0), X^- = \max(-X, 0)\).

If at least one of \(\mathbb{E}(X^+)\) or \(\mathbb{E}(X^-)\) is finite, then we define \(\mathbb{E}(X) = \mathbb{E}(X^+) + \mathbb{E}(X^-)\).

We call \(X\) integrable if \(\mathbb{E}(|X|) < \infty\).

Definition 1.8. Let \(A, B \in \mathcal{F}, \mathbb{P}(B) > 0\). Then

\[
\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}
\]

\[
\mathbb{E}[X|B] = \frac{\mathbb{E}(X1(B))}{\mathbb{P}(B)}
\]

Goal - we want to define \(\mathbb{E}(X|\mathcal{G})\) that is a random variable measurable with respect to the \(\sigma\)-algebra \(\mathcal{G}\).

1.1 Discrete Case

Suppose \(\mathcal{G}\) is a \(\sigma\)-algebra countably generated \((B_i)_{i \in \mathbb{N}}\) is a collection of pairwise disjoint sets in \(\mathcal{F}\) with \(\cup B_i = \Omega\). Let \(\mathcal{G} = \sigma(B_i, i \in \mathbb{N})\).

It is easy to check that \(\mathcal{G} = \{\cup_{j \in J} B_j, J \subseteq \mathbb{N}\}\).

Let \(X\) be an integrable random variable. Then

\[
X' = \mathbb{E}(X|\mathcal{G}) = \sum_{i \in \mathbb{N}} \mathbb{E}(X|B_i) 1(B_i)
\]
(i) $X'$ is $\mathcal{G}$-measurable (check).

(ii) 
\[ \mathbb{E}(|X'|) \leq \mathbb{E}(|X|) \]  
(1.3) 
and so $X'$ is integrable.

(iii) $\forall G \in \mathcal{G}$, then 
\[ \mathbb{E}(X 1_G) = \mathbb{E}(X' 1_G) \]  
(1.4) 
(check).

1.2 Existence and Uniqueness

Definition 1.9. $A \in \mathcal{F}$, $A$ happens almost surely (a.s.) if $\mathbb{P}(A) = 1$.

Theorem 1.10 (Monotone Convergence Theorem). If $X_n \geq 0$ is a sequence of random variables and $X_n \uparrow X$ as $n \to \infty$ a.s, then 
\[ \mathbb{E}(X_n) \uparrow \mathbb{E}(X) \]  
(1.5) 
almost surely as $n \to \infty$.

Theorem 1.11 (Dominated Convergence Theorem). If $(X_n)$ is a sequence of random variables such that $|X_n| \leq Y$ for $Y$ an integrable random variable, then if $X_n \overset{a.s.}{\to} X$ then $\mathbb{E}(X_n) \overset{a.s.}{\to} \mathbb{E}(X)$.

Definition 1.12. For $p \in [1, \infty)$, $f$ measurable functions, then 
\[ \|f\|_p = \mathbb{E}[|f|^p]^{\frac{1}{p}} \]  
(1.6) 
\[ \|f\|_\infty = \inf\{\lambda : |f| \leq \lambda a.e.\} \]  
(1.7) 

Definition 1.13.
\[ L^p = L^p(\Omega, \mathcal{F}, \mathbb{P}) = \{f : \|f\|_p < \infty\} \]

Formally, $L^p$ is the collection of equivalence classes where two functions are equivalent if they are equal a.e. We will just represent an element of $L^p$ by a function, but remember that equality is a.e.
Theorem 1.14. The space \((L^2, \| \cdot \|_2)\) is a Hilbert space with \(\langle U, V \rangle = E(UV)\).

Suppose \(\mathcal{H}\) is a closed subspace, then \(\forall f \in L^2\) there exists a unique \(g \in \mathcal{H}\) such that \(\|f - g\|_2 = \inf \{\|f - h\|_2, h \in \mathcal{H}\ and \langle f - g, h \rangle = 0\ for \ all \ h \in \mathcal{H}\} \).

We call \(g\) the orthogonal projection of \(f\) onto \(\mathcal{H}\).

Theorem 1.15. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be an underlying probability space, and let \(X\) be an integrable random variable, and let \(\mathcal{G} \subset \mathcal{F}\) sub-\(\sigma\)-algebra. Then there exists a random variable \(Y\) such that

(i) \(Y\) is \(\mathcal{G}\)-measurable

(ii) If \(A \in \mathcal{G}\),

\[E(X1(A)) = E(Y1(A))\]  \(1.8\)

and \(Y\) is integrable.

Moreover, if \(Y'\) also satisfies the above properties, then \(Y = Y'\) a.s.

Remark 1.16. \(Y\) is called a version of the conditional expectation of \(X\) given \(\mathcal{G}\) and we write \(\mathcal{G} = \sigma(Z)\) as \(Y = E(X|\mathcal{G})\).

Remark 1.17. (b) could be replaced by the following condition: for all \(Z\) \(\mathcal{G}\)-measurable, bounded random variables,

\[E(XZ) = E(YZ)\]  \(1.9\)

Proof. Uniqueness - let \(Y'\) satisfy (a) and (b). If we consider \(\{Y' - Y > 0\} = A, A\) is \(\mathcal{G}\) measurable. From (b),

\[E((Y' - Y)1(A)) = E(X1(A)) - E(X1(A)) = 0\]

and hence \(P(Y' - Y > 0)) = 0\) which implies that \(Y' \leq Y\) a.s. Similarly, \(Y' \geq Y\) a.s.

Existence - Complete the following three steps:

(i) \(X \in L^2(\Omega, \mathcal{F}, \mathbb{P})\) is a Hilbert space with \(\langle U, V \rangle = E(UV)\). The space \(L^2(\Omega, \mathcal{G}, \mathbb{P})\) is a closed subspace.

\[X_n \rightarrow X(L^2) \Rightarrow X_n \overset{p}{\rightarrow} X \Rightarrow \exists subseq X_{n_k} \overset{as}{\rightarrow} X \Rightarrow X' = \lim sup X_{n_k}\]  \(1.10\)
We can write
\[ L^2(\Omega, \mathcal{F}, \mathbb{P}) = L^2(\Omega, \mathcal{G}, \mathbb{P}) + L^2(\Omega, \mathcal{G}, \mathbb{P})^\perp \]
\[ X = Y + Z \]

Set \( Y = \mathbb{E}(X|\mathcal{G}) \), \( Y \) is \( \mathcal{G} \)-measurable, \( A \in \mathcal{G} \).

\[ \mathbb{E}(X \mathbb{1}(A)) = \mathbb{E}Y \mathbb{1}(A) + \mathbb{E}Z \mathbb{1}(A) = 0 \]

(ii) If \( X \geq 0 \) then \( Y \geq 0 \) a.s. Consider \( A = \{ Y < 0 \} \), then
\[ 0 \leq \mathbb{E}(X \mathbb{1}(A)) = \mathbb{E}(Y \mathbb{1}(A)) \leq 0 \quad (1.11) \]

Thus \( \mathbb{P}(A) = 0 \Rightarrow Y \geq 0 \) a.s.

Let \( X \geq 0 \), Set \( 0 \leq X_n = \max(X, n) \leq n \), so \( X_n \in L^2 \) for all \( n \).

Write \( Y_n = \mathbb{E}(X_n|\mathcal{G}) \), then \( Y_n \geq 0 \) a.s., \( Y_n \) is increasing a.s.. Set \( Y = \limsup Y_n \). So \( Y \) is \( \mathcal{G} \)-measurable. We will show \( Y = \mathbb{E}(X|\mathcal{G}) \) a.s. For all \( A \in \mathcal{G} \), we need to check \( \mathbb{E}(X \mathbb{1}(A)) = \mathbb{E}(Y \mathbb{1}(A)) \). We know that \( \mathbb{E}(X_n \mathbb{1}(A)) = \mathbb{E}(Y_n \mathbb{1}(A)) \), and \( Y_n \uparrow Y \) a.s. Thus, by monotone convergence theorem, \( \mathbb{E}(X \mathbb{1}(A)) = \mathbb{E}(Y \mathbb{1}(A)) \).

If \( X \) is integrable, setting \( A = \Omega \), we have \( Y \) is integrable.

(iii) \( X \) is a general random variable, not necessarily in \( L^2 \) or \( \geq 0 \). Then we have that \( X = X^+ + X^- \). We define \( \mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X^+|\mathcal{G}) - \mathbb{E}(X^-|\mathcal{G}) \). This satisfies (a), (b).

\[ \square \]

**Remark 1.18.** If \( X \geq 0 \), we can always define \( Y = \mathbb{E}(X|\mathcal{G}) \) a.s. The integrability condition of \( Y \) may not be satisfied.

**Definition 1.19.** Let \( \mathcal{G}_0, \mathcal{G}_1, \ldots \) be sub-\( \sigma \)-algebras of \( \mathcal{F} \). Then they are called independent if for all \( i, j \in \mathbb{N} \),
\[ \mathbb{P}(G_i \cap \cdots \cap G_j) = \prod_{i=1}^n \mathbb{P}(G_i) \quad (1.12) \]

**Theorem 1.20.** (i) If \( X \geq 0 \) then \( \mathbb{E}(X|\mathcal{G}) \geq 0 \)
(ii) $E(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(X)$ (A = $\Omega$)

(iii) $X$ is $\mathcal{G}$-measurable implies $\mathbb{E}(X|\mathcal{G}) = X$ a.s.

(iv) $X$ is independent of $\mathcal{G}$, then $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X)$.

**Theorem 1.21** (Fatou’s lemma). $X_n \geq 0$, then for all $n$,

\[ E(\liminf X_n) \leq \liminf E(X_n) \quad (1.13) \]

**Theorem 1.22** (Conditional Monotone Convergence). Let $X_n \geq 0$, $X_n \uparrow X$ a.s. Then

\[ \mathbb{E}(X_n|\mathcal{G}) \uparrow \mathbb{E}(X|\mathcal{G}) \text{ a.s.} \quad (1.14) \]

**Proof.** Set $Y_n = \mathbb{E}(X_n|\mathcal{G})$. Then $Y_n \geq 0$ and $Y_n$ is increasing. Set $Y = \limsup Y_n$. Then $Y$ is $\mathcal{G}$-measurable. \qed

**Theorem 1.23** (Conditional Fatou’s Lemma). $X_n \geq 0$, then

\[ E(\liminf X_n|\mathcal{G}) \leq \liminf E(X_n|\mathcal{G}) \text{ a.s.} \quad (1.15) \]

**Proof.** Let $X$ denote the limit inferior of the $X_n$. For every natural number $k$ define pointwise the random variable $Y_k = \inf_{n \geq k} X_n$. Then the sequence $Y_1, Y_2, \ldots$ is increasing and converges pointwise to $X$. For $k \leq n$, we have $Y_k \leq X_n$, so that

\[ E(Y_k|\mathcal{G}) \leq E(X_n|\mathcal{G}) \text{ a.s.} \quad (1.16) \]

by the monotonicity of conditional expectation, hence

\[ E(Y_k|\mathcal{G}) \leq \inf_{n \geq k} E(X_n|\mathcal{G}) \text{ a.s.} \quad (1.17) \]

because the countable union of the exceptional sets of probability zero is again a null set. Using the definition of $X$, its representation as pointwise limit of the $Y_k$, the monotone convergence theorem for conditional expectations, the last inequality, and the definition of the limit inferior, it follows that almost surely
\[
\mathbb{E}\left(\liminf_{n \to \infty} X_n|G\right) = \mathbb{E}(X|G) 
\]
(1.18)

\[
= \mathbb{E}\left(\lim_{k \to \infty} Y_k|G\right) 
\]
(1.19)

\[
= \lim_{k \to \infty} \mathbb{E}(Y_k|G) 
\]
(1.20)

\[
\leq \liminf_{k \to \infty} \mathbb{E}(X_n|G) 
\]
(1.21)

\[
= \liminf_{n \to \infty} \mathbb{E}(X_n|G) 
\]
(1.22)

\[ \square \]

**Theorem 1.24** (Conditional dominated convergence). *TODO*

### 1.3 Conditional Jensen’s Inequalities

Let $X$ be an integrable random variable such that $\phi(x)$ is integrable of $\phi$ is non-negative. Suppose $G \subset \mathcal{F}$ is a $\sigma$-algebra. Then

\[
\mathbb{E}(\phi(X)|G) \geq \phi(\mathbb{E}(X|G)) 
\]
(1.23)

almost surely. In particular, if $1 \leq p < \infty$, then

\[
\|\mathbb{E}(X|G)\|_p \leq \|X\|_p 
\]
(1.24)

**Proof.** Every convex function can be written as $\phi(x) = \sup_{i \in \mathbb{N}}(a_i x + b_i), a_i, b_i \in \mathbb{R}$. Then

\[
\mathbb{E}(\phi(X)|G) \geq a \mathbb{E}(X|G) + b_i
\]

\[
\mathbb{E}(\phi(X)|G) \geq \sup_{i \in \mathbb{N}}(a_i \mathbb{E}(X|G) + b_i)
\]

\[
= \phi(\mathbb{E}(X|G))
\]

The second part follows from

\[
\|\mathbb{E}(X|G)\|_p = \mathbb{E}(\|\mathbb{E}(X|G)\|_p) \leq \mathbb{E}(\|X\|_p) = \mathbb{E}(\|X\|_p) = \|X\|_p
\]
(1.25)

\[ \square \]

**Proposition 1.25** (Tower Property). Let $X \in L^1, \mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$ be
sub-$\sigma$-algebras. Then

$$\mathbb{E}(\mathbb{E}(X | G) | H) = \mathbb{E}(X | H) \quad (1.26)$$

almost surely.

Proof. Clearly $\mathbb{E}(X | H)$ is $H$-measurable. Let $A \in H$. Then

$$\mathbb{E}(\mathbb{E}(X | H) \mathbb{I}(A)) = \mathbb{E}(X \mathbb{I}(A)) = \mathbb{E}(\mathbb{E}(X | G) \mathbb{I}(A)) \quad (1.27)$$

$\square$

**Proposition 1.26.** Let $X \in L^1, G \subset \mathcal{F}$ be sub-$\sigma$-algebras. Suppose that $Y$ is bounded, $G$-measurable. Then

$$\mathbb{E}(XY | G) = Y \mathbb{E}(X | G) \quad (1.28)$$

almost surely.

Proof. Clearly $Y \mathbb{E}(X | G)$ is $G$-measurable. Let $A \in G$. Then

$$\mathbb{E}(Y \mathbb{E}(X | G) \mathbb{I}(A)) = \mathbb{E} \left( \mathbb{E}(X | G) \underbrace{Y \mathbb{I}(A)}_{\text{G-measurable, bounded}} \right) = \mathbb{E}(XY \mathbb{I}(A)) \quad (1.29)$$

$\square$

**Definition 1.27.** A collection $\mathcal{A}$ of subsets of $\Omega$ is called a $\pi$-system if for all $A, B \in \mathcal{A}$, then $A \cap B \in \mathcal{A}$.

**Proposition 1.28** (Uniqueness of extension). Suppose that $\xi$ is a $\sigma$-algebra on $E$. Let $\mu_1, \mu_2$ be two measures on $(E, \xi)$ that agree on a $\pi$-system generating $\xi$ and $\mu_1(E) = \mu_2(E) < \infty$. Then $\mu_1 = \mu_2$ everywhere on $\xi$.

**Theorem 1.29.** Let $X \in L^1, G, H \subset \mathcal{F}$ two sub-$\sigma$-algebras. If $\sigma(X, G)$ is independent of $H$, then

$$\mathbb{E}(X | \sigma(G, H)) = \mathbb{E}(X | G) \quad (1.30)$$

almost surely.
Proof. Take $A \in \mathcal{G}, B \in \mathcal{H}$.

\[
\mathbb{E}(\mathbb{E}(X|\mathcal{G}) \mathbb{I}(A) \mathbb{I}(B)) = \mathbb{P}(B) \mathbb{E}(\mathbb{E}(X|\mathcal{G}) \mathbb{I}(A)) \\
= \mathbb{P}(B) \mathbb{E}(X \mathbb{I}(A)) \\
= \mathbb{E}(X \mathbb{I}(A) \mathbb{I}(B)) \\
= \mathbb{E}(\mathbb{E}(X|\sigma(\mathcal{G}, \mathcal{H})) \mathbb{I}(A \cap B))
\]

Assume $X \geq 0$, the general case follows by writing $X = X^+ - X^-$. Now, letting $F \in \mathcal{F}$, we have that $\mu(F) = \mathbb{E}(\mathbb{E}(X|\mathcal{G}) \mathbb{I}(F))$, and if $\mu, \nu$ are two measures on $(\Omega, \mathcal{F})$, setting $A = \{A \cap B : A \in \mathcal{G}, B \in \mathcal{H}\}$. Then $A$ is a $\pi$-system.

$\mu, \nu$ are two measurables that agree on the $\pi$-system $A$ and $\mu(\Omega) = \mathbb{E}(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(X) = \nu(\Omega) < \infty$, since $X$ is integrable. Note that $A$ generates $\sigma(\mathcal{G}, \mathcal{H})$.

So, by the uniqueness of extension theorem, $\mu, \nu$ agree everywhere on $\sigma(\mathcal{G}, \mathcal{H})$. \qed

**Remark 1.30.** If we only had $X$ independent of $\mathcal{H}$ and $\mathcal{G}$ independent of $\mathcal{H}$, the conclusion can fail. For example, consider coin tosses $X, Y$ independent $0, 1$ with probability $\frac{1}{2}$, and $Z = \mathbb{I}(X = Y)$.

### 1.4 Product Measures and Fubini’s Theorem

**Definition 1.31.** A measure space $(E, \xi, \mu)$ is called $\sigma$-finite if there exists sets $(S_n)_n$ with $\cup S_n = E$ and $\mu(S_n) < \infty$ for all $n$.

Let $(E_1, \xi_1, \mu_1)$ and $(E_2, \xi_2, \mu_2)$ be two $\sigma$-finite measure spaces, with $A = \{A_1 \times A_2 : A_1 \in \xi_1, A_2 \in \xi_2\}$ a $\pi$-system of subsets of $E = E_1 \times E_2$. Define $\xi = \xi_1 \otimes \xi_2 = \sigma(A)$.

**Definition 1.32** (Product measure). Let $(E_1, \xi_1, \mu_1)$ and $(E_2, \xi_2, \mu_2)$ be two $\sigma$-finite measure spaces. Then there exists a unique measure $\mu$ on $(E, \xi)$ ($\mu = \mu_1 \otimes \mu_2$) satisfying

\[
\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)
\]

for all $A_1 \in \xi_1, A_2 \in \xi_2$. 


Theorem 1.33 (Fubini’s Theorem). Let \((E_1, \xi_1, \mu_1)\) and \((E_2, \xi_2, \mu_2)\) be \(\sigma\)-finite measure spaces. Let \(f \geq 0\), \(f\) is \(\xi\)-measurable. Then

\[
\mu(f) = \int_{E_1} \left( \int_{E_2} f(x_1, x_2) \mu_2(dx_2) \right) \mu_1(dx_1)
\]

If \(f\) is integrable, then \(x_2 \mapsto f(x_1, x_2)\) is \(u_2\)-integrable for \(u_1\)-almost all \(x\).

Moreover, \(x_1 \mapsto \int_{E_2} f(x_1, x_2) \mu_2(dx_2)\) is \(\mu_1\)-integrable and \(\mu(f)\) is given by (1.32).

1.5 Examples of Conditional Expectation

Definition 1.34. A random vector \((X_1, X_2, \ldots, X_n) \in \mathbb{R}^n\) is called a Gaussian random vector if and only if for all \(a_1, \ldots, a_n \in \mathbb{R}\),

\[
a_1 X_1 + \cdots + a_n X_n
\]

is a Gaussian random variable.

\((X_t)_{t \geq 0}\) is called a Gaussian process if for all \(0 \leq t_1 \leq t_2 \leq \cdots \leq t_n\), the vector \(X_{t_1}, \ldots, X_{t_n}\) is a Gaussian random vector.

Example 1.35 (Gaussian case). Let \((X, Y)\) be a Gaussian vector in \(\mathbb{R}^2\). We want to calculate

\[
\mathbb{E}(X|Y) = \mathbb{E}(X|\sigma(Y)) = X'
\]

where \(X' = f(Y)\) with \(f\) a Borel function. Let’s try \(f\) of a linear function \(X' = aY = b, a, b \in \mathbb{R}\) to be determined.

Note that \(\mathbb{E}(X) = \mathbb{E}(X')\) and \(E(X' - X)Y = 0 \Rightarrow Cov(X - X', Y) = 0\) by laws of conditional expectation. Then we have that

\[
a \mathbb{E}(Y) + b = \mathbb{E}(X) Cov(X, Y) = a \mathbb{V}(X)
\]

TODO - continue inference

1.6 Notation for Example Sheet 1

(i) \(G \vee H = \sigma(G, H)\).

(ii) Let \(X, Y\) be two random variables taking values in \(\mathbb{R}\) with joint density \(f_{X,Y}(x, y)\) and \(h: \mathbb{R} \to \mathbb{R}\) be a Borel function such that
Let \( h(X) \) be integrable. We want to calculate

\[
\mathbb{E}(h(X)|Y) = \mathbb{E}(h(X)|\sigma(Y)) \tag{1.36}
\]

Let \( g \) be bounded and measurable. Then

\[
\begin{align*}
\mathbb{E}(h(X)g(Y)) &= \int \int h(x)g(y)f_{X,Y}(x,y)dxdy \\
&= \int \int h(x)g(y)\frac{f_{X,Y}(x,y)}{f_Y(y)}f_Y(y)dxdy \\
&= \int \left( \int h(x)\frac{f_{X,Y}(x,y)}{f_Y(y)}dx \right)g(y)f_Y(y)dy
\end{align*} \tag{1.37}
\]

with \( 0/0 = 0 \)

Set \( \phi(y) = \int h(x)\frac{f_{X,Y}(x,y)}{f_Y(y)}dx \) if \( f_Y(y) > 0 \), and \( 0 \) otherwise. Then we have

\[
\mathbb{E}(h(X)|Y) = \phi(Y) \tag{1.40}
\]

almost surely, and

\[
\mathbb{E}(h(X)|Y) = \int h(x)\nu(Y,dx) \tag{1.41}
\]

with \( \nu(y,dx) = \frac{f_{X,Y}(x,y)}{f_Y(y)}\mathbb{1}(f_Y(y) > 0)dx = f_{X|Y}(x|y)dx \).

\( \nu(y,dx) \) is called the conditional distribution of \( X \) given \( Y = y \) and \( f_{X|Y}(x|y) \) is the conditional density of \( X \) given \( Y = y \).


Discrete Time Martingales

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(E, \xi)$ a measurable space. Usually $E = \mathbb{R}, \mathbb{R}^d, \mathbb{C}$. For us, $E = \mathbb{R}$. A sequence $X = (X_n)_{n \geq 0}$ of random variables taking values in $E$ is called a stochastic process.

A filtration is an increasing family $(\mathcal{F}_n)_{n \geq 0}$ of sub-$\sigma$-algebras of $\mathcal{F}_n$, so $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$.

Intuitively, $\mathcal{F}_n$ is the information available to us at time $n$. To every stochastic process $X$ we associate a filtration called the natural filtration

$$ (\mathcal{F}^X_n)_{n \geq 0}, \mathcal{F}^X_n = \sigma(X_k, k \leq n) \quad (2.1) $$

A stochastic process $X$ is called adapted to $(\mathcal{F}_n)_{n \geq 0}$ if $X_n$ is $\mathcal{F}_n$-measurable for all $n$.

A stochastic process $X$ is called integrable if $X_n$ is integrable for all $n$.

**Definition 2.1.** An adapted integrable process $(X_n)_{n \geq 0}$ taking values in $\mathbb{R}$ is called a

(i) martingale if $\mathbb{E}(X_n | \mathcal{F}_m) = X_m$ for all $n \geq m$.

(ii) super-martingale if $\mathbb{E}(X_n | \mathcal{F}_m) \leq X_m$ for all $n \geq m$.

(iii) sub-martingale if $\mathbb{E}(X_n | \mathcal{F}_m) \geq X_m$ for all $n \geq m$.

**Remark 2.2.** A (sub,super)-martingale with respect to a filtration $\mathcal{F}_n$ is also a (sub, super)-martingale with respect to the natural filtration of $X_n$ (by the tower property).
Example 2.3. Suppose \((\xi_i)\) are i.i.d random variables with \(E(\xi_i) = 0\). Set \(X_n = \sum_{i=1}^n \xi_i\). Then \((X_n)\) is a martingale.

Example 2.4. As above, but let \((\xi_i)\) be i.i.d with \(E(\xi_i) = 1\). Then \(X_n = \prod_{i=1}^n \xi_i\) is a martingale.

Definition 2.5. A random variables \(T : \Omega \to \mathbb{Z}_+ \cup \{\infty\}\) is called a stopping time if \(\{T \leq n\} \in \mathcal{F}_n\) for all \(n\). Equivalently, \(\{T = n\} \in \mathcal{F}_n\) for all \(n\).

Example 2.6. (i) Constant times are trivial stopping times.

(ii) \(A \in \mathcal{B}(\mathbb{R})\). Define \(T_A = \inf\{n \geq 0|X_n \in A\}\), with \(\inf\emptyset = \infty\). Then \(T_A\) is a stopping time.

Proposition 2.7. Let \(S, T, (T_n)\) be stopping times on the filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_n), \mathbb{P})\). Then \(S \wedge T, S \vee T, \inf_n T_n, \liminf_n T_n, \limsup_n T_n\) are stopping times.

Notation. \(T\) stopping time, then \(X_T(\omega) = X_{T(\omega)}(\omega)\). The stopped process \(X^T\) is defined by \(X^T_t = X_{T \wedge t}\).

\[\mathcal{F}_T = \{A \in \mathcal{F}|A \cap \{T \leq t\} \in \mathcal{F}_t, \forall t\}\]

Proposition 2.8. \((\Omega, \mathcal{F}, (\mathcal{F}_n), \mathbb{P}), X = (X_n)_{n \geq 0}\) is adapted.

(i) \(S \leq T\), stopping times, then \(\mathcal{F}_S \subseteq \mathcal{F}_T\)

(ii) \(X_T 1_{(T < \infty)}\) is \(\mathcal{F}_T\)-measurable.

(iii) \(T\) a stopping time, then \(X^T\) is adapted.

(iv) If \(X\) is integrable, then \(X^T\) is integrable.

Proof. Let \(A \in \mathbb{F}\). Need to show that \(\{X_T 1_{(T < \infty)} \in A\} \in \mathcal{F}_T\).

\[\{X_T 1_{(T < \infty)}\} \cap \{T \leq t\} = \bigcup_{s \leq t} \left( \left\{ T = s \right\} \cap \left\{ X_s \in A \right\} \right) \in \mathcal{F}_t \quad (2.2)\]

\[\square\]

2.1 Optional Stopping

Theorem 2.9. Let \(X\) be a martingale.
(i) If $T$ is a stopping time, then $X_T$ is also a martingale. In particular, 
\[ \mathbb{E}(X_{T \wedge t}) = \mathbb{E}(X_0) \text{ for all } t. \]

(ii)

(iii)

(iv)

Proof. By the tower property, it is sufficient to check
\[
\mathbb{E}(X_{T \wedge t} | \mathcal{F}_{t-1}) = \mathbb{E}(X_{T \wedge t} \mathbb{1}_{T \leq t} | \mathcal{F}_{t-1}) + \mathbb{E}(X_T \mathbb{1}_{T > t-1} | \mathcal{F}_{t-1})
\]
\[
= \sum_{s=0}^{t-1} \mathbb{1}(T = s) X_s + \mathbb{1}(T > t-1) X_{t-1} = X_{T \wedge (t-1)}
\]
Since it is a martingale, $\mathbb{E}(X_{T \wedge t}) = \mathbb{E}(X_0)$. \qed

**Theorem 2.10.** Let $X$ be a martingale.

(i) If $T$ is a stopping time, then $X_T$ is also a martingale, so in particular
\[ \mathbb{E}(X_{T \wedge t}) = \mathbb{E}(X_0) \] (2.3)

(ii) If $X \leq T$ are bounded stopping times, then $\mathbb{E}(X_T | \mathcal{F}_S) = X_S$ almost surely.

Proof. Let $S \leq T \leq n$. Then $X_T = (X_{T} - X_{T-1}) + (X_{T-1} - X_{T-2}) + \cdots + (X_{S+1} - X_S) + X_S + \sum_{k=0}^{n} (X_{k+1} - X_k) \mathbb{1}(S \leq k < T).

Let $A \in \mathcal{F}_s$. Then
\[
\mathbb{E}(X_T \mathbb{1}(A)) = \mathbb{E}(X_s \mathbb{1}(A)) + \sum_{k=0}^{n} \mathbb{E}((X_{k+1} - X_k) \mathbb{1}(S \leq k < T) \mathbb{1}(A))
\]
(2.4)
\[
= \mathbb{E}(X_s \mathbb{1}(A))
\] (2.5)
\qed

**Remark 2.11.** The optimal stopping theorem also holds for super/sub-martingales with the respective martingale inequalities in the statement.
Example 2.12. Suppose that \((\xi_i)\)
are random variables with
\[
P(\xi_i = 1) = P(\xi_i = -1) = \frac{1}{2}
\] (2.6)

Set \(X_0 = 0, X_n = \sum_{i=1}^{n} \xi_i\). This is a simply symmetric random walk on \(X_n\).

Let \(T = \inf\{n \geq 0 : X_n = 1\}\). Then \(\mathbb{P}T < \infty = 1\), but \(T\) is not bounded.

Proposition 2.13. If \(X\) is a positive supermartingale and \(T\) is a stopping time which is finite almost surely (\(P(T < \infty) = 1\)), then
\[
E(X_T) \leq E(X_0)
\] (2.7)

Proof. \[\Box\]

2.2 Hitting Probabilities for a Simple Symmetric Random Walk

Let \((\xi_i)\) be i.i.d \(\pm 1\) equally likely. Let \(X_0 = 0, X_n = \sum_{i=1}^{n} \xi_i\). For all \(x \in \mathbb{Z}\) let
\[
T_x = \inf\{n \geq 0 : X_n = x\}
\] (2.9)

which is a stopping time. We want to explore hitting probabilities \(P(T_a < T_b)\) for \(a, b > 0\). If \(E(T) < \infty\), then by (iv) in Theorem 2.10, \(E(X_T) = E(X_0) = 0\).

\[
E(X_T) = -aP(T_{-a} < T_b) + bP(T_b < T_{-a}) = 0
\] (2.10)

and thus obtain that
\[
P(T_{-a} < T_b) = \frac{b}{a+b}
\] (2.11)

Remains to check \(E(T) < \infty\). We have \(P(\xi_1 = 1, \xi_{a+b} = 1) = \frac{1}{2^{a+b}}\).
2.3 Martingale Convergence Theorem

**Theorem 2.14.** Let $X = (X_n)_{n \geq 0}$ be a (super-)-martingale bounded in $L^1$, that is, $\sup_{n \geq 0} \mathbb{E}(|X_n|) < \infty$. Then $X_n$ converges as $n \to \infty$ almost surely towards an a.s. finite limit $X \in L^1(\mathcal{F}_\infty)$ with $\mathcal{F}_\infty = \sigma(\mathcal{F}_n, n \geq 0)$. To prove it we will use Doob's trick which counts up-crossings of intervals with rational endpoints.

**Corollary 2.15.** Let $X$ be a positive supermartingale. Then it converges to an almost surely finite limit as $n \to \infty$.

**Proof.**

\[ \mathbb{E}(|X_n|) = \mathbb{E}(X_n) \leq \mathbb{E}(X_0) < \infty \quad (2.12) \]

□

**Proof.** Let $x = (x_n)_n$ be a sequence of real numbers, and let $a < b$ be two real numbers. Let $T_0(x) = 0$ and inductively for $k \geq 0$,

\[ S_{k+1}(x) = \inf\{n \geq T_k(x) : x_n \leq a\} \quad T_{k+1}(x) = \inf\{n \geq S_{k+1}(x) : x_n \geq b\} \quad (2.13) \]

with the usual convention that $\inf\emptyset = \infty$.

Define $N_n([a,b],x) = \sup\{k \geq 0 : T_k(x) \leq n\}$ - the number of up-crossings of the interval $[a,b]$ by the sequence $x$ by the time $n$. As $n \to \infty$, we have

\[ N_n([a,b],x) \uparrow N([a,b],x) = \sup\{k \geq 0 : T_k(x) < \infty\}, \quad (2.14) \]

the total number of up-crossings of the interval $[a,b]$. □

**Lemma 2.16.** A sequence of rationals $x = (x_n)_n$ converges in $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm \infty\}$ if and only if $N([a,b],x) < \infty$ for all rationals $a, b$.

**Proof.** Assume $x$ converges. Then if for some $a < b$ we had that $N([a,b],x) = \infty$, then $\liminf_n x_n \leq a < b \leq \limsup_n x_n$, which is a contradiction.

Then, suppose that $x$ does converge. Then $\liminf_n x_n > \limsup_n x_n$, and so taking $a, b$ rationals between these two numbers gives that $N([a,b],x) = \infty$ as required. □
**Theorem 2.17** (Doob’s up-crossing inequality). Let $X$ be a supermartingale and $a < b$ be two real numbers. Then for all $n \geq 0$,

\[(b - a)\mathbb{E}(N_n([a,b], X)) \leq \mathbb{E}((X_n - a)^-)\quad (2.15)\]

**Proof.** For all $k$,

\[X_{T_k} - X_{S_k} \geq b - a\quad (2.16)\]

\[\square\]

### 2.4 Uniform Integrability

**Theorem 2.18.** Suppose $X \in L^1$. Then the collection of random variables

\[\{\mathbb{E}(X|\mathcal{G})\}\quad (2.17)\]

for $\mathcal{G} \subseteq \mathcal{F}$ a sub-$\sigma$-algebra is uniformly integrable.

**Proof.** Since $X \in L^1$, for all $\epsilon > 0$ there exists $S > 0$ such that if $A \in \mathcal{F}$ and $\mathbb{P}(A) < \delta$, then $\mathbb{E}(|X|I(A)) \leq \epsilon$.

Set $Y = \mathbb{E}(X|\mathcal{G})$. Then $\mathbb{E}(|Y|) \leq \mathbb{E}(|X|)$. Choose $\lambda < \infty$ such that $\mathbb{E}(|X|) \leq \lambda \delta$. Then

\[\mathbb{P}(|Y| \geq \lambda) \leq \frac{\mathbb{E}(|Y|)}{\lambda} \leq \delta\quad (2.18)\]

by Markov’s inequality.

Then

\[\mathbb{E}(|Y|I(|Y| \geq \lambda)) \leq \mathbb{E}(|X|I|\mathcal{G}|I(|Y| \geq \lambda))\]

\[= \mathbb{E}(|X|I(|Y| \geq \lambda))\quad (2.19)\]

\[\leq \epsilon\quad (2.20)\]

\[\leq \epsilon\quad (2.21)\]

\[\square\]

**Definition 2.19.** A process $X = (X_n)_{n \geq 0}$ is called a uniformly integrable martingale if it is a martingale and the collection $(X_n)$ is uniformly integrable.

**Theorem 2.20.** Let $X$ be a martingale. Then the following are equivalent.
(i) X is a uniformly integrable martingale.

(ii) X converges almost surely and in $L^1$ to a limit $X_\infty$ as $n \to \infty$.

(iii) There exists a random variable $Z \in L^1$ such that $X_n = \mathbb{E}(Z|\mathcal{F}_n)$ almost surely for all $n \geq 0$.

**Theorem 2.21** (Chapter 13 of Williams). Let $X_n, X \in L^1$ for all $n \geq 0$ and suppose that $X_n \xrightarrow{a.s.} X$ as $n \to \infty$. Then $X_n$ converges to $X$ in $L^1$ if and only if $(X_n)$ is uniformly integrable.

**Proof.** We proceed as follows.

$(i) \Rightarrow (ii)$ Since $X$ is uniformly integrable, it is bounded in $L^1$ and by the martingale convergence theorem, we get that $X_n$ converges almost surely to a finite limit $X_\infty$. By the previous theorem, Theorem 2.21 gives $L^1$ convergence.

$(ii) \Rightarrow (iii)$ Set $Z = X_\infty$. We need to show that $X_n = \mathbb{E}(Z|\mathcal{F}_n)$ almost surely for all $n \geq 0$. For all $m \geq n$ by the martingale property we have

$$
\|X_n - \mathbb{E}(X_\infty|\mathcal{F}_n)\|_1 = \|\mathbb{E}(X_m - X_\infty|\mathcal{F}_n)\|_1 \leq \|X_m - X_\infty\|_1 \to 0 \quad (2.22)
$$

as $m \to \infty$.

$(iii) \Rightarrow (i)$ $\mathbb{E}(Z|\mathcal{F}_n)$ is a martingale by the tower property of conditional expectation. Uniform integrability follows from Theorem 2.18.

\[\square\]

**Remark 2.22.** If $X$ is UI then $X_\infty = \mathbb{E}(Z|\mathcal{F}_\infty)$ a.s where $F_\infty = \sigma(\mathcal{F}_n, n \geq 0)$.

**Remark 2.23.** If $X$ is a super/sub-martingale UI, then it converges almost surely and in $L^1$ to a finite limit $X_\infty$ with $\mathbb{E}(X_\infty|\mathcal{F}_n) (\geq) (\leq) X_n$ almost surely.

**Example 2.24.** Let $X_1, X_2, \ldots$ be IID random variables with $\mathbb{P}(X = 0) = \mathbb{P}(X = 2) = \frac{1}{2}$. Set $Y_n = X_1 \cdots X_n$. Then $Y_n$ is a martingale.

As $\mathbb{E}(Y_n) = 1$ for all $n$, we have $(Y_n)$ is bounded in $L^1$, and it converges almost surely to 0. But $\mathbb{E}(Y_n) = 1$ for all $n$, and hence it does not converge in $L^1$. 
If \( X \) is a UI martingale and \( T \) is a stopping time, then we can unambiguously define
\[
X_T = \sum_{n=0}^{\infty} X_n \mathbb{I}(T = n) + X_\infty \mathbb{I}(T = \infty) \tag{2.23}
\]

**Theorem 2.25** (Optional stopping for UI martingales). Let \( X \) be a UI martingale and let \( S, T \) be stopping times with \( S \leq T \). Then
\[
\mathbb{E}(X_T | \mathcal{F}_S) = X_S \tag{2.24}
\]
almost surely.

**Proof.** We first show that \( \mathbb{E}(X_\infty | \mathcal{F}_T) = X_T \) almost surely for any stopping time \( T \). First, check that \( X_T \in L^1 \). Since \( |X_n| \leq \mathbb{E}(|X_\infty| | \mathcal{F}_n) \), we have
\[
\mathbb{E}(|X_T|) = \sum_{n=0}^{\infty} \mathbb{E}(|X_n| \mathbb{I}(T = n) + \mathbb{E}(|X_\infty| \mathbb{I}(T = \infty))) \leq \sum_{n \in \mathbb{Z} \cup \{\infty\}} \mathbb{E}(|X_\infty| \mathbb{I}(T = n)) \tag{2.25}
\]
\[
= \mathbb{E}(|X_\infty|) \tag{2.26}
\]
Let \( B \in \mathcal{F}_T \). Then
\[
\mathbb{E}(\mathbb{I}(B) X_T) = \sum_{n \in \mathbb{Z} \cup \{\infty\}} \mathbb{E}(\mathbb{I}(B) \mathbb{I}(T = n) X_n) \tag{2.28}
\]
\[
= \sum_{n \in \mathbb{Z} \cup \{\infty\}} \mathbb{E}(\mathbb{I}(B) \mathbb{I}(T = n) X_\infty) \tag{2.29}
\]
\[
= \mathbb{E}(\mathbb{I}(B) X_\infty) \tag{2.30}
\]
where for the second equality we used that \( \mathbb{E}(X_\infty | \mathcal{F}_n) = X_n \) almost surely.

Clearly \( X_T \) is \( \mathcal{F}_T \)-measurable, and hence \( \mathbb{E}(X_\infty | \mathcal{F}_T) = X_T \) almost surely. Using the tower property of conditional expectation, we have
for stopping times $S \leq T$ (as $\mathcal{F}_S \subseteq \mathcal{F}_T$),

\[
\mathbb{E}(X_T | \mathcal{F}_S) = \mathbb{E}(\mathbb{E}(X_\infty | \mathcal{F}_T) | \mathcal{F}_S) = \mathbb{E}(X_\infty | \mathcal{F}_S) = X_S
\] (2.31)

almost surely.

2.5 Backwards Martingales

Let $\ldots \subseteq \mathcal{G}_{-2} \subseteq \mathcal{G}_{-1} \subseteq \mathcal{G}_0$ be a sequence of ....

2.6 Applications of Martingales

**Theorem 2.26** (Kolmogorov’s 0–1 law). Let $(X_i)_{i \geq 1}$ be a sequence of IID random variables. Let $\mathcal{F}_n = \sigma(X_k, k \geq n)$ and $\mathcal{F}_\infty = \cap_{n \geq 1} \mathcal{F}_n$. Then $\mathcal{F}_\infty$ is trivial - that is, every $A \in \mathcal{F}_\infty$ has probability $\mathbb{P}(A) \in \{0, 1\}$.

**Proof.** Let $\mathcal{G}_n = \sigma(X_k, k \leq n)$ and $A \in \mathcal{F}_\infty$. Since $\mathcal{G}_n$ is independent of $\mathcal{F}_n$, we have that

\[
\mathbb{E}(\mathbb{1}(A) | \mathcal{G}_n) = \mathbb{P}(A)
\] (2.34)

Theorem 2.26 (LN) gives that $\mathbb{P}(A) = \mathbb{E}(\mathbb{1}(A) | \mathcal{G}_n)$ converges to $\mathbb{E}(\mathbb{1}(A) | \mathcal{G}_\infty)$ as $n \to \infty$, where $\mathcal{G}_\infty = \sigma(\mathcal{G}_n, n \geq 0)$. Then we deduce that $\mathbb{E}(\mathbb{1}(A) | \mathcal{G}_n) = \mathbb{1}(A) = \mathbb{P}(A)$ as $\mathcal{F}_\infty \subseteq \mathcal{G}_\infty$. Therefore, $\mathbb{P}(A) = \ldots$

**Theorem 2.27** (Strong law of large numbers). Let $(X_i)_{i \geq 1}$ be a sequence of IID random variables in $L^1$ with $\mu = \mathbb{E}(X_i)$. Let $S_n = \sum_{i=1}^{n} X_i$ and $S_0 = 0$. Then $\frac{S_n}{n} \to \mu$ as $n \to \infty$ almost surely and in $L^1$.

**Proof.**

**Theorem 2.28** (Kakutani’s product martingale theorem). Let $(X_n)_{n \geq 0}$ be a sequence of independent non-negative random variables of mean 1. Let $M_0 = 1$, $M_n = \prod_{i=1}^{n} X_i$ for $n \in \mathbb{N}$. Then $(M_n)_{n \geq 0}$ is a non-negative martingale and $M_n \to M_\infty$ a.s. as $n \to \infty$ for some random variable $M_\infty$.

We set $a_n = \mathbb{E}(\sqrt{X_n})$ then $a_n \in (0, 1]$. Moreover,

(i) If $\prod_n a_n > 0$, then $M_n \to M_\infty$ in $L^1$ and $\mathbb{E}(M_\infty) = 1$,

(ii) If $\prod_n a_n = 0$, then $M_\infty = 0$ almost surely.
2.6.1 Martingale proof of the Radon-Nikodym theorem

Let $P, Q$ be two probability measures on the measurable space $\Omega, \mathcal{F}$. Assume that $\mathcal{F}$ is countably generated, that is, there exists a collection of sets $(F_n)_{n \in \mathbb{N}}$ such that $\mathcal{F} = \sigma(F_n, n \in \mathbb{N})$. Then the following are equivalent.

(i) $P(A) = 0 \Rightarrow Q(A)$ for all $A \in \mathcal{F}$. That is, $Q$ is absolutely continuous with respect to $P$ and write $Q \ll P$.

(ii) For all $\epsilon > 0$, there exists $\delta > 0$ such that $P(A) \leq \delta \Rightarrow Q(A) \leq \epsilon$.

(iii) There exists a non-negative random variable $X$ such that

$$Q(A) = \mathbb{E}_P(X\mathbb{1}(A)) \quad (2.35)$$

Proof. $(i) \rightarrow (ii)$. If $(ii)$ does not hold, then there exists $\epsilon > 0$ such that for all $n \geq 1$ there exists a set $A_n$ with $P(A_n) \leq \frac{1}{2^n}$ and $Q(A_n) \geq \epsilon$. By Borel-Cantelli, we get that $P(A_{n.i.o}) = 0$. Therefore from $(i)$ we get that $Q(A_{n.i.o}) = 0$. But

$$Q(A_{n.i.o}) = Q(\cap_{n \geq 1} A_k) = \lim_{n \to \infty} Q(\cup_{k \geq n} A_k) \geq \epsilon \quad (2.36)$$

which is a contradiction.

$(ii) \rightarrow (iii)$. Consider the filtration $\mathcal{F}_n = \sigma(F_k, k \leq n)$. Let

$$A_n = \{H_1 \cap \cdots \cap H_n | H_i = F_i \text{ or } F_i^c \} \quad (2.37)$$

then it is easy to see that $\mathcal{F}_n = \sigma(A_n)$. Note also that sets in $A_n$ are disjoint.

Proof. $(i) \rightarrow (ii)$. If $(ii)$ does not hold, then there exists $\epsilon > 0$ such that for all $n \geq 1$ there exists a set $A_n$ with $P(A_n) \leq \frac{1}{2^n}$ and $Q(A_n) \geq \epsilon$. By Borel-Cantelli, we get that $P(A_{n.i.o}) = 0$. Therefore from $(i)$ we get that $Q(A_{n.i.o}) = 0$. But

$$Q(A_{n.i.o}) = Q(\cap_{n \geq 1} A_k) = \lim_{n \to \infty} Q(\cup_{k \geq n} A_k) \geq \epsilon \quad (2.36)$$

which is a contradiction.

$(ii) \rightarrow (iii)$. Consider the filtration $\mathcal{F}_n = \sigma(F_k, k \leq n)$. Let

$$A_n = \{H_1 \cap \cdots \cap H_n | H_i = F_i \text{ or } F_i^c \} \quad (2.37)$$

then it is easy to see that $\mathcal{F}_n = \sigma(A_n)$. Note also that sets in $A_n$ are disjoint.
3

Stochastic Processes in Continuous Time

Our setting is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space with $t \in J \subseteq \mathbb{R}_+ = [0, \infty)$

**Definition 3.1.** A filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ is an increasing collection of $\sigma$-algebras $(\mathcal{F}_t)_{t \in J}$, satisfying $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $t \geq s$. A stochastic process in continuous time is an ordered collection of random variables on $\Omega$. 
4

Bibliography