Mathematics II: Handout

Klaus Pötzelberger
Vienna University of Economics and Business
Institute for Statistics and Mathematics
E-mail: klaus.poetzelberger@wu.ac.at

Contents

Chapter 1

Martingales

1.1 Definition and Examples

Let a probability space $(\Omega, \mathcal{F}, P)$ be given. A *filtration* $(\mathcal{F}_n)_{n=0}^\infty$ is an increasing sequence $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$ of sub σ-algebras of $\mathcal{F}$. A *stochastic process* $(X_n)_{n=0}^\infty$ is a sequence of random variables $X_n : (\Omega, \mathcal{F}) \to (\mathbb{R}^m, \mathcal{B}^m)$. A stochastic process $(X_n)_{n=0}^\infty$ is *adapted* (not anticipating) (to the filtration $(\mathcal{F}_n)_{n=0}^\infty$), if X_n is $\mathcal{F}_n$-measurable for all n. It is *predictable*, if X_n is $\mathcal{F}_{n-1}$-measurable for all $n \geq 1$ and X_0 is $\mathcal{F}_0$-measurable.

Example 1.1 *Let $(X_n)_{n=0}^\infty$ be given. Define $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$. $(\mathcal{F}_n)_{n=0}^\infty$ is the filtration generated by $(X_n)_{n=0}^\infty$, also called the history of $(X_n)_{n=0}^\infty$.*

Example 1.2 *Let $(X_n)_{n=0}^\infty$ be adapted to the filtration $(\mathcal{F}_n)_{n=0}^\infty$. Then a process (H_n) with $H_n = f_n(X_0, \ldots, X_{n-1})$ is predictable. For instance, if (H_n) is non-random (deterministic), $H_n = f(n)$, it is predictable. Or $(H_n) = (X_{n-1})$ is.*

Example 1.3 *A game with stochastic outcome is played repeatedly. Let $(X_n)_{n=0}^\infty$ be the outcomes and (H_n) the stakes a gambler chooses before the n-th game. $(X_n)_{n=0}^\infty$ and the sequence of accumulated gains are adapted to the history of $(X_n)_{n=0}^\infty$, (H_n) is predictable.*

Example 1.4 *Let (S_n) be the price of a financial asset, H_n the number of assets a trader holds in the interval $(n-1, n]$ and $(\mathcal{F}_n)_{n=0}^\infty$ the history of (S_n). The wealth of the trader at time n is*

$$V_n = V_0 + \sum_{k=1}^{n} H_k (S_k - S_{k-1}).$$

(V_n) is adapted $(\mathcal{F}_n)_{n=0}^\infty$, the trading strategy (H_n) is predictable.

Definition 1.5 *A stochastic process $(X_n)_{n=0}^\infty$ is a* martingale *(or a $(\mathcal{F}_n)_{n=0}^\infty$ martingale), if*

1. $\mathbb{E}(|X_n|) < \infty$ *for all $n \geq 0$,*

2. $(X_n)_{n=0}^{\infty}$ *is adapted (to* $(\mathcal{F}_n)_{n=0}^{\infty}$*),*

3. $\mathbb{E}(X_n \mid \mathcal{F}_{n-1}) = X_{n-1}$ *for all* $n \geq 1$*.*

Theorem 1.6 *Let* $(X_n)_{n=0}^{\infty}$ *be adapted to* $(\mathcal{F}_n)_{n=0}^{\infty}$ *and integrable (i.e. for all* n*,* $\mathbb{E}(|X_n|) < \infty$*). Then the following statements are equivalent:*

1. $(X_n)_{n=0}^{\infty}$ *is a* $(\mathcal{F}_n)_{n=0}^{\infty}$ *martingale.*

2. $\mathbb{E}(X_n \mid \mathcal{F}_m) = X_m$ *for all* $m < n$*.*

3. $\mathbb{E}(X_n - X_{n-1} \mid \mathcal{F}_{n-1}) = 0$ *for all* $n \geq 1$*.*

Furthermore, if $(X_n)_{n=0}^{\infty}$ *is a martingale, then* $\mathbb{E}(X_n) = \mathbb{E}(X_m)$ *for all* $m < n$*.*

Proof. Note that since X_{n-1} is $\mathcal{F}_{n-1}$-measurable,

$$\mathbb{E}(X_n \mid \mathcal{F}_{n-1}) = \mathbb{E}(X_n - X_{n-1} + X_{n-1} \mid \mathcal{F}_{n-1}) = \mathbb{E}(X_n - X_{n-1} \mid \mathcal{F}_{n-1}) + X_{n-1},$$

which proves the equivalence of 1. and 3. Statement 2. implies 1. It follows from 1. by iteration, for instance,

$$X_{n-2} = \mathbb{E}(X_{n-1} \mid \mathcal{F}_{n-2}) = \mathbb{E}(\mathbb{E}(X_n \mid \mathcal{F}_{n-1}) \mid \mathcal{F}_{n-2}) = \mathbb{E}(X_n \mid \mathcal{F}_{n-2}).$$

Finally, for $m < n$,

$$\mathbb{E}(X_m) = \mathbb{E}(\mathbb{E}(X_n \mid \mathcal{F}_m)) = \mathbb{E}(X_n).$$

□

If we do not refer to the filtration, we assume that it is the filtration generated by the stochastic process, i.e. $\mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n)$. In this case, (X_n) is automatically adapted.

Example 1.7 *(Random Walk). Let* (Z_n) *be a sequence of independent and integrable random variables. Let* $X_n = Z_0 + Z_1 + \cdots + Z_n$*. If* $(Z_n)_{n=1}^{\infty}$ *are identically distributed, the process is called a* random walk*. If* $\mathbb{E}(Z_n) = 0$ *for all* $n \geq 1$ *it is a martingale.*

Special cases are the simple random walk. Here $Z_n \in \{-1, 1\}$*, i.e. the random walk jumps up or down with jump size equal to* 1 *in absolute value. It is symmetric, if* $P(Z_n = 1) = P(Z_n = -1) = 1/2$*. The symmetric simple random walk is a martingale.*

Example 1.8 *(Geometric Random Walk). Let* (Z_n) *be a sequence of independent and integrable random variables. Let* $X_n = Z_0 Z_1 \cdots Z_n$*. If* $(Z_n)_{n=1}^{\infty}$ *are identically distributed, the process is called a* geometric random walk*. If* $\mathbb{E}(Z_n) = 1$ *for all* $n \geq 1$ *it is a martingale.*

A special case is the (geometric) binomial process. It is the process underlying the model of Cox, Ross and Rubinstein (CRR-model) in financial mathematics. $Z_n \in \{U, D\}$*, where* $D < U$ *are* *constants. CRR-process jumps up or down. In the first case it is multiplied by U, in the second by D. Let p denote the probability of an up and $1-p$ of a down. (X_n) is a martingale, if and only if*

$$1 = \mathbb{E}(Z_n) = Up + D(1-p),$$

which hold if and only if $D \leq 1 \leq U$ and

$$p = \frac{1-D}{U-D}.$$

Example 1.9 *(Martingale Transform). Let $(\mathcal{F}_n)_{n=0}^{\infty}$ be a filtration and $(X_n)_{n=0}^{\infty}$ a martingale. Furthermore, let $(H_n)_{n=0}^{\infty}$ be predictable. Assume either that both H_n and X_n are square-integrable or that H_n is bounded. Let $Y_0 = H_0 X_0$ and for $n \geq 1$,*

$$Y_n = H_0 X_0 + \sum_{k=0}^{n} H_k (X_k - X_{k-1}). \tag{1.1}$$

(Y_n) is a martingale:

$$\mathbb{E}(Y_n \mid \mathcal{F}_{n-1}) = \mathbb{E}(Y_{n-1} + H_n(X_n - X_{n-1}) \mid \mathcal{F}_{n-1}) = Y_{n-1} + H_n \mathbb{E}(X_n - X_{n-1} \mid \mathcal{F}_{n-1}) = Y_{n-1} + H_n\, 0 = Y_{n-1}.$$

The martingale transform of a martingale (X_n) and a predictable process (H_n) is denoted by $H \bullet X$.

1.2 Doob's Decomposition

Definition 1.10 *Let $(X_n)_{n=0}^{\infty}$ be a stochastic process such that*

1. *$\mathbb{E}(|X_n|) < \infty$ for all $n \geq 0$,*
2. *$(X_n)_{n=0}^{\infty}$ is adapted (to $(\mathcal{F}_n)_{n=0}^{\infty}$).*

It is called a supermartingale *if for all $n \geq 1$,*

3. $\mathbb{E}(X_n \mid \mathcal{F}_{n-1}) \leq X_{n-1}$.

It is called a submartingale *if for all $n \geq 1$,*

3.' $\mathbb{E}(X_n \mid \mathcal{F}_{n-1}) \geq X_{n-1}$.

Proposition 1.11 *Let $(X_n)_{n=0}^{\infty}$ be a martingale and let $f : \mathbb{R} \to \mathbb{R}$ be measurable and convex such that for all n the random variable $f(X_n)$ is integrable. Let $Y_n = f(X_n)$. Then (Y_n) is a submartingale.*

Proof. Jensen's inequality for conditional expectations says that $\mathbb{E}(f(X_n) \mid \mathcal{F}_{n-1}) \geq f(\mathbb{E}(X_n \mid \mathcal{F}_{n-1}))$. Since (X_n) is a martingale, the r.h.s. is $f(X_{n-1}) = Y_{n-1}$ and the l.h.s. is $\mathbb{E}(Y_n \mid \mathcal{F}_{n-1})$. □

Theorem 1.12 *(Doob Decomposition). Let $(\mathcal{F}_n)_{n=0}^{\infty}$ be a filtration and $(X_n)_{n=0}^{\infty}$ an adapted integrable process. There exists a martingale (M_n) and a predictable process (A_n) with $A_0 = 0$ such that $X_n = M_n + A_n$. If $(X_n)_{n=0}^{\infty}$ is a submartingale, then the process is nondecreasing (i.e. $A_n \leq A_{n+1}$ for all n). The decomposition is unique.*

Proof. Let $M_0 = X_0$ and $A_0 = 0$. Suppose such a decomposition is possible. Then

$$\begin{aligned} X_k &= M_k + A_k, \\ \mathbb{E}(X_k \mid \mathcal{F}_{k-1}) &= M_{k-1} + A_k, \\ X_{k-1} &= M_{k-1} + A_{k-1}, \end{aligned}$$

which implies

$$A_k - A_{k-1} = \mathbb{E}(X_k \mid \mathcal{F}_{k-1}) - X_{k-1}. \tag{1.2}$$

Let us define

$$A_n = \sum_{k=1}^{n} (\mathbb{E}(X_k \mid \mathcal{F}_{k-1}) - X_{k-1})$$

and $M_n = X_n - A_n$. Clearly, (A_n) is predictable. Furthermore,

$$\begin{aligned} \mathbb{E}(M_n \mid \mathcal{F}_{n-1}) &= \mathbb{E}(X_n - A_n \mid \mathcal{F}_{n-1}) \\ &= \mathbb{E}(X_n \mid \mathcal{F}_{n-1}) - A_n \\ &= \mathbb{E}(X_n \mid \mathcal{F}_{n-1}) - A_{n-1} - \mathbb{E}(X_n \mid \mathcal{F}_{n-1}) + X_{n-1} \\ &= M_{n-1}. \end{aligned}$$

Finally, (1.2) implies $A_k \geq A_{k-1}$ iff $\mathbb{E}(X_k \mid \mathcal{F}_{k-1}) \geq X_{k-1}$ □

Remark 1.13 *1. The predictable process (A_n) is called the* compensator *of (X_n).*
2. Let $(X_n)_{n=0}^{\infty}$ be a square-integrable martingale. Proposition 1.11 implies that (X_n^2) is a submartingale. Hence, there exists a predictable process (A_n), the compensator of (X_n^2), such that $X_n^2 - A_n$ is a martingale. (A_n) is called the quadratic variation *of (X_n).*

Example 1.14 *Let $(X_n)_{n=0}^{\infty}$ be a square integrable martingale. Then the quadratic variation is $A_n = \sum_{k=1}^{n} \mathbb{E}((X_k - X_{k-1})^2 \mid \mathcal{F}_{k-1})$.*

Example 1.15 *Let $(X_n)_{n=0}^{\infty}$ be a square integrable random walk that is a martingale and let σ^2 the variance of the innovations $Z_n = X_n - X_{n-1}$. Then the quadratic variation is $A_n = n\sigma^2$.*

Example 1.16 *Let $(X_n)_{n=0}^{\infty}$ be a square integrable martingale with quadratic variation (A_n). Let (H_n) be predictable and square integrable and let (Y_n) be the martingale transform of (X_n) and (H_n). Then the quadratic variation of (Y_n) is $B_n = \sum_{k=1}^{n} H_k^2 \mathbb{E}((X_k - X_{k-1})^2 \mid \mathcal{F}_{k-1})$, i.e. $B = H^2 \bullet A$.*

1.3 Stopping Times

Let $\bar{\mathbb{N}} = \mathbb{N}_0 \cup \{\infty\}$.

Definition 1.17 *A random variable $T : \Omega \to \bar{\mathbb{N}}$ is called a* stopping time *if for all n, the event $\{T \leq n\}$ is $\mathcal{F}_n$-measurable.*

Note that T is a stopping time if and only if $\{T = n\} \in \mathcal{F}_n$ for all n.

Example 1.18 *1. Let the stochastic process $(X_n)_{n=0}^{\infty}$ be adapted to the filtration $(\mathcal{F}_n)_{n=0}^{\infty}$. Let b be a constant and define $T = \min\{n \mid X_n \geq b\}$.*

Since

$$\{T \leq n\} = \{X_0 \geq b\} \cup \{X_1 \geq b\} \cup \cdots \cup \{X_n \geq b\}$$

and the events $\{X_k \geq b\} \in \mathcal{F}_k \subseteq \mathcal{F}_n$, we have $\{T \leq n\} \in \mathcal{F}_n$ and T is a stopping time.

2. The time $T = \min\{n \mid X_{n+2} \geq b\}$ is not a stopping time.

3. Let $m < n$ and $A \in \mathcal{F}_m$. Define $T(\omega) = m$ if $\omega \in A$ and $T(\omega) = n$ if $\omega \notin A$. T is a stopping time.

A stopping time T is finite, if $P(T = \infty) = 0$. It is bounded, if there is a constant N s.t. $P(T \leq N) = 1$. For a finite stopping time T we define X_T as

$$X_T(\omega) = X_{T(\omega)}(\omega) = \sum_{n=0}^{\infty} X_n(\omega) I_{\{T(\omega)=n\}}. \tag{1.3}$$

Proposition 1.19 *Let T be a bounded stopping time and $(X_n)_{n=0}^{\infty}$ a martingale. Then $\mathbb{E}(X_T) = \mathbb{E}(X_0)$.*

Proof. Let T be bounded by N. First note that X_T may be written as

$$X_T = X_0 + \sum_{n=1}^{N} I_{\{T \geq n\}}(X_n - X_{n-1}) = X_0 + \sum_{n=1}^{N} H_n(X_n - X_{n-1}),$$

with $H_n = I_{\{T \geq n\}} = 1 - I_{\{T \leq n-1\}}$ $\mathcal{F}_{n-1}$-measurable. That is, if (Y_n) denotes the martingale transform with coefficients (H_n) and driven by (X_n), then $X_T = Y_N$. Since (Y_n) is a martingale, $\mathbb{E}(X_T) = \mathbb{E}(X_0)$ follows. $\square$

Definition 1.20 *Let T be a stopping time. The stopping time σ-algebra $\mathcal{F}_T$ is*

$$\mathcal{F}_T = \{A \in \mathcal{F} \mid A \cap \{T \leq n\} \in \mathcal{F}_n \text{ for all } n\}.$$

$\mathcal{F}_n$ is the collection of events that are measurable up to time n, analogously $\mathcal{F}_T$ is the collection of events that are measurable up to the random time T.

Proposition 1.21 *If T is a stopping time, then $\mathcal{F}_T$ is a σ-algebra.*

Proof. Clearly, $\emptyset, \Omega \in \mathcal{F}_T$. If $A \in \mathcal{F}_T$ then, since $A^c \cap \{T \le n\} = \{T \le n\} \cap (A \cap \{T \le n\})^c$, we have $A^c \in \mathcal{F}_T$. Finally, if $A_i \in \mathcal{F}_T$, then

$$(\cup_{i=1}^{\infty} A_i) \cap \{T \le n\} = \cup_{i=1}^{\infty}(A_i \cap \{T \le n\}) \in \mathcal{F}_n$$

and therefore $\cup_{i=1}^{\infty} A_i \in \mathcal{F}_T$. □

Proposition 1.22 *Let S and T be stopping times with $S \le T$. Then $\mathcal{F}_S \subseteq \mathcal{F}_T$.*

Proof. Note that $\{T \le n\} \subseteq \{S \le n\}$ and therefore, if $A \in \mathcal{F}_S$, then

$$A \cap \{T \le n\} = A \cap \{S \le n\} \cap \{T \le n\}.$$

Since both $A \cap \{S \le n\}$ and $\{T \le n\}$ are in $\mathcal{F}_n$, we get $A \cap \{T \le n\} \in \mathcal{F}_n$. □

Proposition 1.23 *X_T is $\mathcal{F}_T$-measurable.*

Proof. Let B be a Borel set. We have to show that $\{X_T \in B\} \cap \{T \le n\} \in \mathcal{F}_n$ for all n. We have

$$\begin{aligned}\{X_T \in B\} \cap \{T \le n\} &= \cup_{k=0}^{n}\{X_T \in B\} \cap \{T = k\} \\ &= \cup_{k=0}^{n}\{X_k \in B\} \cap \{T = k\}.\end{aligned}$$

$\{X_k \in B\} \cap \{T = k\} \in \mathcal{F}_k \subseteq \mathcal{F}_n$ implies $\{X_T \in B\} \cap \{T \le n\} \in \mathcal{F}_n$. □

Theorem 1.24 *(Doob's Optional Sampling Theorem). Let $(X_n)_{n=0}^{\infty}$ be a martingale and let S, T be bounded stopping times with $S \le T$ a.s. Then, a.s.,*

$$\mathbb{E}(X_T \mid \mathcal{F}_S) = X_S. \tag{1.4}$$

Proof. Let T be bounded by N. X_T is integrable, since $|X_T| \le \sum_{k=0}^{N} |X_k|$. We have to prove that $\mathbb{E}(I_A X_T) = \mathbb{E}(I_A X_S)$ for every $A \in \mathcal{F}_S$. Let $R(\omega) = S(\omega)$ if $\omega \in A$ and $R(\omega) = T(\omega)$ if $\omega \in A^c$. R is a stopping time, since for all n,

$$\{R \le n\} = (\{R \le n\} \cap A) \cup (\{R \le n\} \cap A^c) = (\{S \le n\} \cap A) \cup (\{T \le n\} \cap A^c) \in \mathcal{F}_n.$$

Proposition 1.19 implies that $\mathbb{E}(X_R) = \mathbb{E}(X_0) = \mathbb{E}(X_T)$. Finally

$$\mathbb{E}(X_T) = \mathbb{E}(X_T I_A) + \mathbb{E}(X_T I_{A^c})$$

and

$$\mathbb{E}(X_0) = \mathbb{E}(X_R) = \mathbb{E}(X_S I_A) + \mathbb{E}(X_T I_{A^c}) = \mathbb{E}(X_S I_A) + \mathbb{E}(X_0) - \mathbb{E}(X_T I_A)$$

implies $\mathbb{E}(X_S I_A) = \mathbb{E}(X_T I_A)$. □

Theorem 1.25 *Let $(X_n)_{n=0}^\infty$ be adapted and integrable such that for all n, $\mathbb{E}(X_n) = \mathbb{E}(X_0)$. If for all bounded stopping times T, $\mathbb{E}(X_T) = \mathbb{E}(X_0)$, then $(X_n)_{n=0}^\infty$ is a martingale.*

Proof. We may assume that $X_0 = 0$, otherwise we replace X_n by $X_n - X_0$. Let $0 \le m < n$, $A \in \mathcal{F}_m$ and define $T(\omega) = m$ if $\omega \in A^c$ and $T(\omega) = n$ if $\omega \in A$. If $\mathbb{E}(X_T) = 0$, then

$$\begin{aligned} 0 = \mathbb{E}(X_T) &= \mathbb{E}(I_A X_n) + \mathbb{E}(I_{A^c} X_m) \\ &= \mathbb{E}(I_A X_n) + \mathbb{E}((1 - I_A) X_m) \\ &= \mathbb{E}(I_A X_n) - \mathbb{E}(I_A X_m) \end{aligned}$$

and therefore for all $\mathcal{F}_m$-measurable A, $\mathbb{E}(I_A X_n) = \mathbb{E}(I_A X_m)$, i.e. $X_m = \mathbb{E}(X_n \mid \mathcal{F}_m)$. □

1.4 Exercises

Exercise 1.1 *Let $(X_n)_{n=0}^\infty$ be a stochastic process and define $S_n = X_0 + \cdots + X_n$. Show that $(X_n)_{n=0}^\infty$ and $(S_n)_{n=0}^\infty$ have the same history.*

Exercise 1.2 *Let (X_n) be a martingale. Prove*

1. *If (X_n) is predictable, then it is constant in time, i.e. $X_n = X_0$ for all n.*

2. *If X_n is independent of $\mathcal{F}_{n-1}$, then (X_n) is constant, i.e. $X_n = \mathbb{E}(X_0)$ for all n.*

Exercise 1.3 *Let $(X_n)_{n=0}^\infty$ be a random walk with innovations (Z_n). Show that if e_n^Z is integrable, (Y_n) with $Y_n = e_n^X$ is a geometric random walk.*

Let (Y_n) a positive geometric random walk and let $X_n = \log Y_n$. Show that (X_n) is a random walk.

Exercise 1.4 *(Continued) Show that if (X_n) is a martingale, then (Y_n) is a submartingale, but except in the trivial case, not a martingale. On the other hand, if (Y_n) is a martingale, then (X_n) is a supermartingale, but not a martingale.*

Exercise 1.5 *Let (X_n) be a martingale. Prove that martingale differences $X_n - X_{n-1}$ are uncorrelated to $\mathcal{F}_{n-1}$, i.e. to all $\mathcal{F}_{n-1}$ measurable random variables. Conclude that*

$$\mathbb{V}(X_n) = \mathbb{V}(X_0) + \sum_{k=1}^{n} \mathbb{V}(X_k - X_{k-1}).$$

Exercise 1.6 *Let (X_n) be a geometric random walk that is a martingale. Let $H_n = X_{n-1}$. Find a martingale (U_n) s.t. (X_n) is the martingale transform of (U_n) and (H_n).*

Exercise 1.7 *Let* (X_n) *be a geometric random walk that is a martingale. Let* $D_n = (1+r)^n$*, where* $r > 0$ *and* $Y_n = D_n X_n$*. Compute the compensator of* (Y_n)*.*

Exercise 1.8 *Let* (X_n) *be a random walk with quadratic integrable innovations* $Z_n = X_n - X_{n-1}$ *that is a martingale. Derive the quadratic variation of* (X_n)*.*

Exercise 1.9 *Let* (X_n) *be a geometric random walk that is a martingale. Derive the quadratic variation of* (X_n)*.*

Exercise 1.10 *Let* $(X_n)_{n=0}^{\infty}$ *be a submartingale and let* $f : \mathbb{R} \to \mathbb{R}$ *be measurable, increasing and convex such that for all* n *the random variable* $f(X_n)$ *is integrable. Let* $Y_n = f(X_n)$*. Show that* (Y_n) *is a submartingale.*

Exercise 1.11 *Prove that a supermartingale* $(X_n)_{n=0}^{\infty}$ *may be written as* $X_n = M_m + A_n$*, where* (M_n) *is a martingale,* (A_n) *is predictable and nonincreasing.*

Exercise 1.12 *Let* Y *be integrable and* $X_n = \mathbb{E}(Y \mid \mathcal{F}_n)$*. Show that* $(X_n)_{n=0}^{\infty}$ *is a martingale.*

Exercise 1.13 *Let* (X_n) *be adapted. Prove that* (X_n) *is a martingale, if for all predictable processes* (H_n) *with* $H_0 = 0$*,* $\mathbb{E}((H \bullet X)_n) = 0$ *holds.*

Exercise 1.14 *Let* $(\mathcal{F}_n)_{n=0}^{\infty}$ *be a filtration and* $T : \Omega \to \mathbb{N}_0 \cup \{\infty\}$ *a random variable. Show that the following assertions are equivalent:*

1. $\{T \leq n\} \in \mathcal{F}_n$ *for all* n*.*
2. $\{T = n\} \in \mathcal{F}_n$ *for all* n*.*
3. $\{T < n\} \in \mathcal{F}_{n-1}$ *for all* n*.*
4. $\{T \geq n\} \in \mathcal{F}_{n-1}$ *for all* n*.*
5. $\{T > n\} \in \mathcal{F}_n$ *for all* n*.*

Exercise 1.15 *Let* S *and* T *be stopping times. Show that* $S \vee T = \max\{S, T\}$ *and* $S \wedge T = \min\{S, T\}$ *are stopping times.*

Exercise 1.16 *Let* S *and* T *be stopping times. Show that* $S + T$ *is a stopping time.*

Exercise 1.17 *Let* T *be a stopping time. Show that* kT *is a stopping time if* $k \geq 1$ *is an integer.*

Exercise 1.18 *Let* S *and* T *be stopping times. Show that* $\mathcal{F}_{S\wedge T} \subseteq \mathcal{F}_T \subseteq \mathcal{F}_{S\vee T}$*.*

Exercise 1.19 *Let* $(X_n)_{n=0}^N$ *be a simple random walk with* $X_0 = 0$.

1. *Write down carefully a suitable probability space* $(\Omega, \mathcal{F}, P)$.
2. *Let* $N = 3$ *and* $T = \min\{k > 0 \mid X_k = 0\} \wedge 3$. *Show that* T *is a stopping time.*
3. *Derive* $\mathcal{F}_T$ *and* X_T *and the probability distribution of* X_T.

Exercise 1.20 *(Cont.) Let* $(X_n)_{n=0}^N$ *be a simple random walk with* $X_0 = 0$ *and* $N = 3$.

1. *Let* $N = 3$ *and* $S_i = \min\{k \mid X_k = i\} \wedge N$ *for* $i = 1, 2$.
2. *Derive* $\mathcal{F}_{S_i}$ *and* X_{S_i} *and the probability distribution of* X_{S_i}.
3. *Compute* $\mathbb{E}(X_{S_2} \mid \mathcal{F}_{S_1})$.

Chapter 2

Random Walks and Boundary Crossing Probabilities

2.1 Gambling

Example 2.1 *(Gamblers Ruin I). Let* $(X_k)_{k=0}^{\infty}$ *be the independent payoffs of a series of gambling trials,* $X_0 = 0$*,* $S_n = X_0 + X_1 + \cdots + X_n$ *the accumulated payoff after n trials. In the simplest case,* $X_k \in \{-1, 1\}$ *with* $p = P(X_k = 1)$*,* $P(X_k = -1) = 1 - p =: q$*. Let player A with fortune a play against player B with fortune b.*

The game is repeated until either player A or player B has lost his fortune, i.e. until $S_n = -a$ *or* $S_n = b$*. Define*

$$\tau^{a,b} = \min\{n \mid S_n \notin [-a+1, b-1]\}, \tag{2.1}$$

with $\tau^{a,b} = \infty$ *if* $S_n \in [-a+1, b-1]$ *for all n.*

Remark 2.2 *If* X_k *has only the values* 1 *and* -1*, then, on* $\{\tau^{a,b} < \infty\}$*,* $S_{\tau^{a,b}} = -a$ *or* $S_{\tau^{a,b}} = b$*.*

Example 2.3 *(Gamblers Ruin II). Assume that gambler A has unlimited fortune* $(a = \infty)$*. Let*

$$\tau^{b} = \min\{n \mid S_n \geq b\}, \tag{2.2}$$

with $\tau^b = \infty$ *if* $S_n < b$ *for all n.*

Typically, the following problems are of interest:

1. Is $\tau^{a,b}$ finite, i.e. is $P(\tau^{a,b} < \infty) = 1$?

2. Is $\mathbb{E}(\tau^{a,b}) < \infty$?

3. If $\tau^{a,b}$ is finite, what is the probability of the ruin of A (of B)?

4. Is τ^b finite?

5. Is $\mathbb{E}(\tau^b) < \infty$?

6. Derive the distribution of $\tau^{a,b}$ (of τ^b).

The answers to 1. and 2. are "yes", to 5. "no", to 4. "yes" if $p \geq 1/2$. The distribution of τ^b and the probability in 3. can be computed in closed form, for the distribution of $\tau^{a,b}$ series expansions exist.

Closely related is the problem of the existence of strategies or stopping times with positive expected payoff. Assume that player A chooses at time $k-1$ stakes H_k depending on the information up to time $k-1$, $H_k = f_k(X_1, \dots, X_{k-1})$ (H_1 is constant). The payoff is then $S_n^H = H_1X_1 + \cdots + H_kX_k$.

7. Let $a < \infty$. Is there a stopping time T such that $\mathbb{E}(S_T) > 0$?

8. Let $a = \infty$. Is there a stopping time T such that $\mathbb{E}(S_T) > 0$?

9. Do predictable stakes (H_k) exist, such that $\mathbb{E}(S_n^H) > 0$?

Example 2.4 *(Ruin Problem). Denote by X_n the assets of an insurance company at the end of year n. Each year, the company receives premiums $b > 0$ and claims C_n are paid. Assume that (C_n) are i.i.d. with mean $\mu < b$. We have $X_{n+1} = X_n + b - C_{n+1}$. Ruin occurs, if $X_n \leq 0$ for at least one n.*

Compute $P(ruin)$ or give tight estimates of $P(ruin)$. How does $P(ruin)$ depend on X_0, μ, b and other parameters of C_n?

2.2 Calculations for the Simple Random walk

The random walk (S_n) with innovations $X_n = S_n - S_{n-1}$ is simple, if $X_0 = 0$ and $X_n \in \{-1, 1\}$. Let $p = P(X_n = 1)$, $q = 1 - p = P(X_n = -1)$. The simple random walk (s.r.w.) is called symmetric, if $p = q = 1/2$.

To derive the *distribution* of S_n, let

$$B_n = \#\{i \leq n \mid X_i = 1\}.$$

Then $S_n = B_n - (n - B_n) = 2B_n - n$ and thus $\{S_n = k\} = \{B_n = (n+k)/2\}$. Note that n is even iff S_n is even. Thus if n and k are both even or both odd,

$$P(S_n = k) = \binom{n}{(n+k)/2} p^{(n+k)/2} q^{(n-k)/2}. \tag{2.3}$$

Example 2.5 *Let* (S_n) *be a symmetric s.r.w. Then*

$$P(S_{2n}=0)=\binom{2n}{n}\left(\frac{1}{2}\right)^{2n}.$$

Since the innovations (X_n) are independent, (S_n) is a *Markov process* :

$$(S_{n+k})_{k=1}^{\infty} \mid (S_1,\ldots,S_n) \sim (S_{n+k})_{k=1}^{\infty} \mid S_n.$$

Since the innovations (X_n) are identically distributed, (S_n) is *homogeneous* , i.e. its distribution is invariant w.r.t. time-shifts : For $n,m\in\mathbb{N}$,

$$(S_{n+k})_{k=1}^{\infty} \mid S_n \sim (S_{m+k})_{k=1}^{\infty} \mid S_m.$$

Proposition 2.6 *Let the s.r.w.* (S_n) *be symmetric. The first exit time* τ^b*, defined by* (2.2)*, is finite a.s., i.e.* $P(\tau^b<\infty)=1$.

Proof. Let for $b\in\mathbb{N}_0$, $\pi_b=P(\tau^b=\infty)$ ($=P(S_n\neq b$ for all $n)$. We have $\pi_0=0$ and

$$\begin{aligned}\pi_1 &= P(S_1=-1 \text{ and } S_n\neq 1 \text{ for all } n\geq 2)\\ &= P(S_1=-1)P(S_n\neq 2 \text{ for all } n\geq 1)\\ &= \frac{1}{2}\pi_2.\end{aligned}$$

Similarly, for $k\geq 2$ we have

$$\begin{aligned}\pi_k &= P(S_1=1 \text{ and } S_n\neq k \text{ for all } n\geq 2)+P(S_1=-1 \text{ and } S_n\neq k \text{ for all } n\geq 2)\\ &= \frac{1}{2}\pi_{k+1}+\frac{1}{2}\pi_{k-1}.\end{aligned}$$

Therefore, $2\pi_k=\pi_{k+1}+\pi_{k-1}$,

$$\pi_{k+1}-\pi_k=\pi_k-\pi_{k-1}=\cdots=\pi_2-\pi_1=\pi_1$$

and

$$\pi_k=(\pi_k-\pi_{k-1})+\cdots+(\pi_2-\pi_1)+\pi_1=k\pi_1.$$

Since $\pi_k\leq 1$, we have $\pi_1=0$ and thus $\pi_k=0$ for all k. □

Proposition 2.7 *Let the s.r.w.* (S_n) *be asymmetric, let* $p<q$ *and* $\theta=p/q$*. Let the first exit time* τ^b *be defined by* (2.2)*. Then, for* $b\in\mathbb{N}_0$,

$$P(\tau^b<\infty)=\theta^b. \tag{2.4}$$

Proof. Again, let $\pi_k = 1 - P(\tau^k < \infty)$. We have $\pi_0 = 0$, $\pi_1 = q\pi_2$, $\pi_k = p\pi_{k-1} + q\pi_{k+1}$ and therefore

$$\pi_{k+1} - \pi_k = \theta(\pi_k - \pi_{k-1}) = \cdots = \theta^{k-1}(\pi_2 - \pi_1) = \theta^k \pi_1.$$

Thus $\pi_k - \pi_{k-1} = \theta^{k-1}\pi_1$ and

$$\begin{aligned} \pi_k &= \sum_{i=2}^{k} (\pi_i - \pi_{i-1}) + \pi_1 \\ &= \sum_{i=1}^{k} \theta^{i-1}\pi_1 = \frac{1-\theta^k}{1-\theta}\pi_1. \end{aligned}$$

From Lemma 2.9 ($\lim_{k\to\infty} \pi_k = 1$) it follows that $\pi_1 = (1-\theta)$ and $\pi_k = 1 - \theta^k$. □

Corollary 2.8 *Let (S_n) be a s.r.w. and $M = \max\{S_n \mid n \geq 0\}$ with $M = \infty$ if (S_n) is not bounded. If $p = q = 1/2$, then $M = \infty$ a.s. If $p < q$, then $M + 1$ is geometrically distributed with parameter $\theta = p/q$.*

Proof. Let $p < q$. We have for $k \geq 1$

$$P(M + 1 \geq k) = P(M \geq k - 1) = P(\tau^{k-1} < \infty) = \theta^{k-1}.$$

Therefore,

$$P(M + 1 = k) = P(M + 1 \geq k) - P(M + 1 \geq k + 1) = \theta^{k-1} - \theta^k = \theta^{k-1}(1 - \theta).$$

□

Lemma 2.9 *Let (S_n) and τ^b be the asymmetric s.r.w. and the first exit time of Proposition 2.7. Then $\lim_{k\to\infty} P(\tau^k < \infty) = 0$.*

Proof. Note that $pq < 1/4$ and

$$P(\tau^k < \infty) \leq \sum_{n=k}^{\infty} P(S_n = k) = \sum_{n=k, n+k \text{ even}}^{\infty} \binom{n}{(k+n)/2} p^{(n+k)/2} q^{(n-k)/2}.$$

Stirling's approximation for $n!$ ($n! \approx \sqrt{2\pi} n^{n+1/2} e^{-n}$) implies that there exists a constant c, independent of k, such that

$$\binom{n}{(k+n)/2} \leq c \frac{2^n}{\sqrt{n}}.$$

Therefore,

$$\begin{aligned} P(\tau^k < \infty) &\leq \sum_{n=k}^{\infty} c \frac{2^n}{\sqrt{n}} p^{(n+k)/2} q^{(n-k)/2} \\ &\leq \frac{c}{\sqrt{k}} \sum_{n=0}^{\infty} (4pq)^{n/2} = \frac{1}{\sqrt{k}} \frac{c}{(1 - \sqrt{4pq})}. \end{aligned}$$

□

2.3 Reflection Principle

Let the s.r.w. (S_n) be a symmetric. The computation of the probability of an event boils down to the counting of the number of paths defining the event. Probabilities can often be derived from the *reflection principle*. In its simplest form, the reflection principle is the following statement. Let (S_n) be a s.r.w. and let n, m, k be nonnegative integers. Then the number of paths with $S_n = m$ that hit the boundary $-k$ at some time up to n is the same as the number of paths with $S_n = -m - 2k$ (and therefore the number of paths with $S_n = m + 2k$). The idea is: take a path $(S_i(\omega))$ that hits $-k$ for the first time at $\tau \leq n$ and reflect it at $-k$. The reflected path ends in $S_\tau(\omega) - (S_n(\omega) - S_\tau(\omega)) = -k - (S_n(\omega) + k)$.

Let $T^k = \min\{n \geq 1 \mid S_n = k\}$, denote the first time after 0 that the symmetric s.r.w. (S_n) reaches k.

Proposition 2.10 *(Distribution of the first return to* 0*).*

$$P(T^0 > 2n) = P(S_{2n} = 0). \tag{2.5}$$

Proof. We have

$$\begin{aligned}
P(T^0 > 2n) &= P(S_2 = 2, S_k > 0 \text{ for } k = 3, \ldots, 2n) + P(S_2 = -2, S_k < 0 \text{ for } k = 3, \ldots, 2n) \\
&= 2P(S_2 = 2, S_k > 0 \text{ for } k = 3, \ldots, 2n) \\
&= \frac{1}{2} P(S_k > -2 \text{ for } k = 0, \ldots, 2n-2) \\
&= \frac{1}{2} \sum_{m=0}^{n-1} P(S_{2n-2} = 2m, S_k > -2 \text{ for } k = 0, \ldots, 2n-3) \\
&= \frac{1}{2} \sum_{m=0}^{n-1} \left(P(S_{2n-2} = 2m) - P(S_{2n-2} = 2m, S_k \text{ hits } -2, \text{ on } k = 0, \ldots, 2n-3) \right).
\end{aligned}$$

By the reflection principle,

$$P(S_{2n-2} = 2m, S_k \text{ hits } -2, \text{ on } k = 0, \ldots, 2n-3) = P(S_{2n-2} = 2m + 4)$$

and therefore

$$\begin{aligned}
P(T^0 > 2n) &= \frac{1}{2} \sum_{m=0}^{n-1} \left(P(S_{2n-2} = 2m) - P(S_{2n-2} = 2m + 4) \right) \\
&= \frac{1}{2} \left(P(S_{2n-2} = 0) + P(S_{2n-2} = 2) \right) \\
&= \frac{1}{2} P(S_{2n-2} = 0) + \frac{1}{4} P(S_{2n-2} = 2) + \frac{1}{4} P(S_{2n-2} = -2) \\
&= P(S_{2n} = 0 \text{ and } S_{2n-2} = 0) + P(S_{2n} = 0 \text{ and } S_{2n-2} = 2) + P(S_{2n} = 0 \text{ and } S_{2n-2} = -2) \\
&= P(S_{2n} = 0).
\end{aligned}$$

$\square$

Proposition 2.11 *Let $k \neq 0$. Then*

$$P(T^k = n) = \frac{k}{n} P(S_n = k). \tag{2.6}$$

Proof. The set of paths that hit k at time n consist of

1. paths that hit k at n for the first time,
2. paths with $S_n = k$ and $S_{n-1} = k + 1$,
3. paths with $S_n = k$, $S_{n-1} = k - 1$ and $S_i = k$ for some $i \leq n - 2$.

By the reflection principle, the sets 2. and 3. have the same size. Therefore

$$\begin{aligned} P(T^k = n) &= P(S_n = k) - 2P(S_{n-1} = k+1, S_n = k) \\ &= \binom{n}{(n+k)/2} \left(\frac{1}{2}\right)^n - 2\binom{n-1}{(n+k)/2} \left(\frac{1}{2}\right)^{n-1} \frac{1}{2} \\ &= \binom{n}{(n+k)/2} \left(\frac{1}{2}\right)^n \left(1 - \frac{2(n-k)/2}{n}\right) \\ &= \binom{n}{(n+k)/2} \left(\frac{1}{2}\right)^n \frac{k}{n}. \end{aligned}$$

□

2.4 Optional Stopping

Martingale theory and especially optional sampling theorems are powerful tools for proving boundary crossing probabilities. For motivation, consider the Gambler's Ruin problem for the symmetric s.r.w. (S_n). Note that (S_n) is a martingale and $\tau^{a,b}$ a stopping time, i.e. for all $n \geq 0$, the event $\{\tau^{a,b} \leq n\}$ is in $\sigma(S_1, \ldots, S_n)$, the σ-algebra generated by $S_1, \ldots, S_n$.

For bounded stopping times T, we have

$$\mathbb{E}(S_T) = \mathbb{E}(S_0). \tag{2.7}$$

If we knew that (2.7) holds for $T = \tau^{a,b}$ we could compute $P(S_{\tau^{a,b}} = b)$. We have

$$\begin{aligned} 0 &= \mathbb{E}(S_{\tau^{a,b}}) = -aP(S_{\tau^{a,b}} = -a) + bP(S_{\tau^{a,b}} = b) \\ &= -a(1 - P(S_{\tau^{a,b}} = b)) + bP(S_{\tau^{a,b}} = b) \end{aligned}$$

and get

$$P(S_{\tau^{a,b}} = b) = \frac{a}{a+b} \tag{2.8}$$

$$P(S_{\tau^{a,b}} = -a) = \frac{b}{a+b}. \tag{2.9}$$

Example 2.12 $\mathbb{E}(S_T) = \mathbb{E}(S_0)$ *does not hold for all martingales* (S_n) *and all stopping times* T. *Consider the strategy of doubling in a fair game. The player has unlimited fortune, doubles stakes until the first win. If the player loses in the games* $1, 2, \ldots, n-1$, *his loss is* $1+2+2^2+\cdots+2^{n-1} = 2^n - 1$. *We have* $S_n = S_{n-1} + 2^n$ *with probability* $1/2$ *and* $S_n = S_{n-1} - 2^n$ *again with probability* $1/2$. (S_n) *is a martingale. The game is stopped after the first win. Let* $T = T^1 = \min\{n \mid X_n = 1\}$. T^1 *is finite a.s.,* $S_T = 1$ *and therefore*

$$1 = \mathbb{E}(S_T) \neq 0 = \mathbb{E}(S_0).$$

A variety of theorems on optional stopping have been proved. The following is especially suitable for applications concerning random walks.

Theorem 2.13 *Let* (S_n) *be a martingale for which there exists a constant* c *such that for all* n

$$\mathbb{E}(|S_{n+1} - S_n| \mid \sigma(S_1, \ldots, S_n)) \leq c. \text{ for } n < T, \text{ a.s.} \tag{2.10}$$

If T *is a stopping time with* $\mathbb{E}(T) < \infty$, *then*

$$\mathbb{E}(S_T) = \mathbb{E}(S_0). \tag{2.11}$$

Proof. See [1], Chapter 6, Corollary 3.1.

Corollary 2.14 *Let* (S_n) *be a random walk,* $S_n = X_0 + X_1 + \cdots + X_n$ *with integrable* X_n. *Then* (2.11) *holds for any integrable stopping time* T.

Corollary 2.15 *For the symmetric s.r.w.* (S_n), (2.8) *and* (2.9) *hold.*

Example 2.16 *Let the s.r.w.* (S_n) *be symmetric. To compute* $\mathbb{E}(\tau^{a,b})$, *define* $Y_n = S_n^2 - n$. *Note that* (Y_n) *is a martingale (Exercise 2.4). Since* $X_{n+1} \in \{-1, 1\}$, *we have*

$$|Y_{n+1} - Y_n| = |X_{n+1}^2 + 2S_n X_{n+1} - 1| = |2S_n X_{n+1}| \leq 2(\max\{a, b\} - 1)$$

on $n < \tau^{a,b}$. *Now, since* $0 = \mathbb{E}(S^2_{\tau^{a,b}} - \tau^{a,b})$ *we have*

$$\begin{aligned} \mathbb{E}(\tau^{a,b}) &= \mathbb{E}(S^2_{\tau^{a,b}}) \\ &= b^2 \frac{a}{a+b} + a^2 \frac{b}{a+b} = ab. \end{aligned}$$

2.5 Exercises

Exercise 2.1 *Let* $a, b > 0$. *Show that for the s.r.w.* $P(\tau^{a,b} < \infty) = 1$ *and* $\mathbb{E}(\tau^{a,b}) < \infty$.

Hint. Note that if $X_{n+1} = \cdots = X_{n+a+b-1} = -1$, then at least one of $S_n, S_{n+1}, \ldots, S_{n+a+b-1}$ is not in $[-a+1, b-1]$. The probability of a block of -1 of length $a+b-1$ is q^{a+b-1}. Let $m = a+b-1$. Then

$$\begin{aligned}
P(\tau^{a,b} > mk) &\leq P((S_1, \ldots, S_m) \text{ is not a block of } -1's, \\
&\quad (S_{m+1}, \ldots, S_{2m} \text{ is not a block of } -1's, \\
&\quad (S_{m(k-1)+1}, \ldots, S_{mk} \text{ is not a block of } -1's) \\
&\leq (1-q^m)^k.
\end{aligned}$$

Conclude $P(\tau^{a,b} = \infty) = \lim_{k\to\infty} P(\tau^{a,b} > mk) = 0$ and $\mathbb{E}(\tau^{a,b}) = \sum_{n=0}^{\infty} P(\tau^{a,b} > n) < \infty$.

Exercise 2.2 *Show that* $P(T^0 < \infty) = 1$ *and* $\mathbb{E}(T^0) = \infty$.

Hint. Use Stirling's approximation to get

$$P(T^0 > 2n) = P(S_{2n} = 0) = \binom{2n}{n}\left(\frac{1}{2}\right)^n \approx \frac{c}{\sqrt{2n}}.$$

Therefore $\lim_{n\to\infty} P(T^0 > 2n) = 0$ and for a suitable positive constant c,

$$\mathbb{E}(T^0) = \sum_{n=0}^{\infty} P(T^0 > n) \geq \sum_{n=0}^{\infty} cn^{-1/2} = \infty.$$

Exercise 2.3 *Show that for* $k \geq 1$, $P(T^k < \infty) = 1$ *and* $\mathbb{E}(T^k) = \infty$.

Hint. Use Proposition 2.11 and Stirling's approximation to get $P(T^k = n) \approx k\sqrt{2/\pi}n^{-3/2}$ for $n+k$ odd. For a suitable positive constants c_1 and c_2 we have $c_1 n^{-1/2} \leq P(T^k > n) \leq c_2 n^{-1/2}$. Therefore, $\lim_{n\to\infty} P(T^k > n) = 0$ and $\mathbb{E}(T^k) = \sum_{n=0}^{\infty} P(T^1 > n) \geq \sum_{n=0}^{\infty} c_1 n^{-1/2} = \infty$.

Exercise 2.4 *Let* (S_n) *be a random walk with increments* (X_n) *having zero expectation and finite variance* σ^2. *Prove that* $Y_n = S_n^2 - n\sigma^2$ *is a martingale.*

Exercise 2.5 *Let* (S_n) *denote an asymmetric s.r.w. with* $p < q$. *Compute* $P(b) = P(S_{\tau^{a,b}} = b)$, $P(a) = P(S_{\tau^{a,b}} = -a)$ *and* $\mathbb{E}(\tau^{a,b})$.

Hint. First show that $Y_n = (q/p)^{S_n}$ is a martingale which satisfies the assumptions of Theorem 2.13. Then

$$1 = \mathbb{E}(Y_0) = \mathbb{E}(Y_{\tau^{a,b}})) = P(b)(q/p)^b + P(a)(q/p)^{-a}$$

implies

$$\begin{aligned}
P(b) &= \frac{1-(p/q)^a}{(q/p)^b - (p/q)^a}, \\
P(a) &= \frac{(q/p)^b - 1}{(q/p)^b - (p/q)^a}.
\end{aligned}$$

Then apply the optional sampling theorem to the martingale $(S_n - n(p-q))$ to derive

$$\mathbb{E}(\tau^{a,b}) = \frac{bP(b) + aP(a)}{p-q}.$$

Exercise 2.6 *(Wald's identity I). Let (S_n) be a random walk with integrable increments (X_n). Let $\mu = \mathbb{E}(X_1)$ and let τ be a finite stopping time with $\mathbb{E}(\tau) < \infty$. Show that*

$$\mathbb{E}(S_\tau) = \mu \mathbb{E}(\tau).$$

Hint. Apply the optional sampling theorem to $(S_n - \mu n)$.

Exercise 2.7 *(Wald's identity II). Let (S_n) be a random walk with increments (X_n) a having finite moment generating function*

$$m(\theta) = \mathbb{E}(e^{\theta X_1}).$$

Let θ satisfy $m(\theta) \geq 1$. Prove that for $\tau = \tau^{a,b}$,

$$\mathbb{E}(m(\theta)^{-\tau} e^{\theta S_\tau}) = 1.$$

Hint. Show that $(Y_n) = (m(\theta)^{-n} e^{\theta S_n})$ is a martingale that satisfies the assumptions of Theorem 2.13.

Exercise 2.8 *A typical application of Wald's identity II is the computation of $P(b)$ and $P(a)$, where $P(b)$ is the probability that a random walk (S_n) leaves the interval $]-a, b[$ through b, i.e. $P(b) = P(S_T \geq b)$ and $P(a) = P(S_\tau \leq -a)$. In the general case $S_\tau \notin \{-a, b\}$. Choose $\theta \neq 0$ s.t. $m(\theta) = 1$. Then $\mathbb{E}(e^{\theta S_\tau}) = 1$ implies*

$$1 \approx P(b)e^{\theta b} + P(a)e^{-\theta a}$$

and therefore

$$P(b) \approx \frac{1 - e^{-\theta a}}{e^{\theta b} - e^{-\theta a}}, \tag{2.12}$$

$$P(b) \approx \frac{e^{\theta b} - 1}{e^{\theta b} - e^{-\theta a}}. \tag{2.13}$$

Show that if (S_n) is simple and a and b are integers, (2.12) *and* (2.13) *holds exactly. Cf. Exercise 2.5, derive the nonzero solution of $m(\theta) = 1$.*

Chapter 3

The Radon-Nikodym Theorem

3.1 Absolute Continuity

Let two probability measures P and Q be defined on a measurable space $(\Omega, \mathcal{F})$.

Definition 3.1 *Q is* absolutely continuous *w.r.t. P ($Q << P$) if for all $A \in \mathcal{F}$, if $P(A) = 0$, then $Q(A) = 0$. If $Q << P$ and $P << Q$, then P and Q are called* equivalent *($P \sim Q$).*

Example 3.2 *Let $f : (\Omega, \mathcal{F}) \to (\mathbb{R}_+, \mathcal{B})$ with $f \geq 0$ and $\mathbb{E}^P(f) = 1$. Define Q by $Q(A) = \mathbb{E}^P(I_A f)$. Q is a probability distribution and since $\mathbb{E}^P(I_A f) = 0$ for all A with $P(A) = 0$, Q is absolutely continuous w.r.t. P.*

If $f > 0$ a.s., then $P(A) = \mathbb{E}^Q(I_A \frac{1}{f})$ and $P \sim Q$.

Theorem 3.3 *(Radon-Nikodym Theorem). Let $Q << P$. Then there exists a (P-a.s.) unique nonnegative and integrable random variable $\frac{dQ}{dP}$ such that for all $A \in \mathcal{F}$,*

$$Q(A) = \mathbb{E}^P(\frac{dQ}{dP} I_A).$$

Proof. See [2]. □

3.2 Exercises

Exercise 3.1 *Let $P(B) > 0$ and define Q by $Q(A) = P(A \cap B)/P(B)$. Show that $Q << P$ and derive the Radon-Nikodym derivative $\frac{dQ}{dP}$.*

Exercise 3.2 *Let P and Q be defined on a measurable space $(\Omega, \mathcal{F})$. Show that there always exist a probability measure R s.t. $P << R$ and $Q << R$. Moreover, let $P_1, P_2, \dots$ be countably many probability measures. Show that there exist a probability measure R s.t. for all n, $P_n << R$.*

Hint: Let $\pi_n > 0$ with $\sum_n \pi_n = 1$ and let $R = \sum_{n=1}^{\infty} \pi_n P_n$.

Exercise 3.3 *Let* $X \sim B(n,p)$ *and* $Y \sim N(0,1)$. *Let* $Z = X$ *with probability* $1/2$ *and* $Z = Y$ *with probability* $1/2$. *Let the distributions of* X, Y, Z *be denoted by* P, Q, R. *Show that* $P << R$ *and* $Q << R$. *Find* $\frac{dQ}{dR}$ *and* $\frac{dP}{dR}$.

Exercise 3.4 *Let* $\mu \neq 0$. *Denote by* P *and* Q *the normal distributions with variance* 1 *and expectations* 0 *and* μ. *Compute* $\frac{dQ}{dP}$. *Show that* $P \sim Q$.

Exercise 3.5 *Denote by* P *the standard normal distribution and by and* Q *the exponential distribution with rate* 1. *Compute* $\frac{dQ}{dP}$. *Show that* $Q << P$ *and that* P *is not absolutely continuous with respect to* Q.

Exercise 3.6 *Let* $Q << P$ *and* $P << R$. *Prove that* $Q << R$ *and that* $\frac{dQ}{dR} = \frac{dQ}{dP}\frac{dP}{dR}$.

Exercise 3.7 *Let* Z *be, under* P, *Gaussian with mean* 0 *and variance* 1, *i.e.* $Z \sim^P N(0,1)$. *Let* $X = \mu + \sigma Z$. *Find* $Q \sim P$, *s.t.* X *is Gaussian under* Q *and* $\mathbb{E}^Q(X) = -\mathbb{V}ar^Q(X)/2$. *Is* Q *unique?*

Chapter 4

Introduction to Mathematical Finance

4.1 Basic Concepts

Let a probability space $(\Omega, \mathcal{F}, P)$ be given. We consider a market model consisting of $m+1$ assets: $(S_n)_{n=0}^N$, with $S_n = (S_n^0, S_n^1, \dots, S_n^m)$. Typically, (S_n^0) is called the *riskless* asset and we always assume that $S_n^0 > 0$ for all $n = 0, \dots, N$ and $S_0^0 = 1$. Furthermore, we always assume that S_n is square integrable. In the simplest case $m = 1$ and the market consists of a riskless and a risky asset. In this case, we use the notation $S_n^0 = B_n$ and $S_n^1 = S_n$. N is always a finite time-horizon.

The discounted prices of the assets are denoted by $(\bar{S}_n)$, i.e.

$$\bar{S}_n^k = S_n^k / S_n^0.$$

Definition 4.1 *(Trading Strategy) 1. A trading strategy (portfolio) is an $\mathbb{R}^{m+1}$-valued predictable and square-integrable process (ϕ_n), with $\phi_n = (\phi_n^0, \dots, \phi_n^m)$. ϕ_n^k denotes the number of shares of asset k held in $(n-1, n]$ in the portfolio.*

2. The price of the portfolio at time n is

$$V_n(\phi) = \sum_{k=0}^m \phi_n^k \, S_n^k.$$

3. A trading strategy ϕ is self-financing, if for all n

$$\sum_{k=0}^m \phi_n^k \, S_n^k = \sum_{k=0}^m \phi_{n+1}^k \, S_n^k. \qquad \text{(Balancing condition)} \tag{4.1}$$

Proposition 4.2 *Let (ϕ_n) denote a trading strategy. The following statements are equivalent:*

1. (ϕ_n) is self-financing.

2. $(V_n) = ((\phi \bullet S)_n)$, i.e. for all n,

$$V_n(\phi) = V_0(\phi) + \sum_{k=1}^n \langle \phi_k, S_k - S_{k-1} \rangle.$$

3. $(\bar{V}_n) = ((\phi \bullet \bar{S})_n)$, *i.e. for all* n,

$$\bar{V}_n(\phi) = V_0(\phi) + \sum_{k=1}^{n} \langle \phi_k, \bar{S}_k - \bar{S}_{k-1} \rangle.$$

Proof. To establish the equivalence of 1. and 2., we have to prove that

$$V_n - V_{n-1} = \langle \phi_n, S_n - S_{n-1} \rangle$$

is equivalent to (ϕ_n) being self-financing. Note that

$$V_n - V_{n-1} = \langle \phi_n, S_n \rangle - \langle \phi_{n-1}, S_{n-1} \rangle$$

and that the balancing condition is

$$\langle \phi_n, S_{n-1} \rangle = \langle \phi_{n-1}, S_{n-1} \rangle.$$

The equivalence of 1. and 3. can be shown similarly:

$$\bar{V}_n - \bar{V}_{n-1} = \langle \phi_n, S_n \rangle / S_n^0 - \langle \phi_{n-1}, S_{n-1} \rangle / S_{n-1}^0$$

and therefore

$$\bar{V}_n - \bar{V}_{n-1} = \langle \phi_n, \bar{S}_n - \bar{S}_{n-1} \rangle + \langle \phi_n, S_{n-1} \rangle / S_{n-1}^0 - \langle \phi_{n-1}, S_{n-1} \rangle / S_{n-1}^0.$$

Since 1. is the same as

$$\langle \phi_n, S_{n-1} \rangle / S_{n-1}^0 = \langle \phi_{n-1}, S_{n-1} \rangle / S_{n-1}^0,$$

1. is equivalent to 3. □

Corollary 4.3 *Let Q denote a probability distribution such that $(\bar{S}_n)$ is a square-integrable martingale under Q. Then $(\bar{V}_n(\phi))$ is a martingale, for all self-financing trading strategies (ϕ_n) that are square-integrable under Q.*

Definition 4.4 *1. A self-financing strategy is* admissible*, if there exists a constant $c \geq 0$, such that for all $n \leq N$,*

$$V_n(\phi) \geq -c.$$

2. An arbitrage strategy is an admissible strategy (ϕ_n), such that $V_0(\phi) = 0$, $V_N(\phi) \geq 0$ and $P(V_N(\phi) > 0) > 0$. The market model is viable *(arbitrage-free, NA), if there exists no arbitrage strategy.*

No arbitrage is a fundamental assumption in mathematical finance. We aim at deriving a characterization of arbitrage-free markets. This characterization allows to compute prices of *contingent claims, (derivatives, options).* The price of a contingent claim depends on the so-called underlying. Typically, the underlying is traded in a market.

Example 4.5 Bank Account. *Typically, the riskless asset or numeraire is a bank account* (B_n)*. One Euro in time* $n = 0$ *gives* B_n *in* $n > 0$*. We assume that* $B_0 = 1$ *and* $B_n \leq B_{n+1}$ *for all* n*. If the bank account is deterministic with constant (nominal) interest rate* r*, we have*

$$B_n = e^{rn}.$$

Example 4.6 European Call. *The European call gives the holder (the buyer) the right, but not the obligation, to buy the underlying* S *at the* exercise date (expiration date), N *for a price* K*, the* exercise price (strike price) .

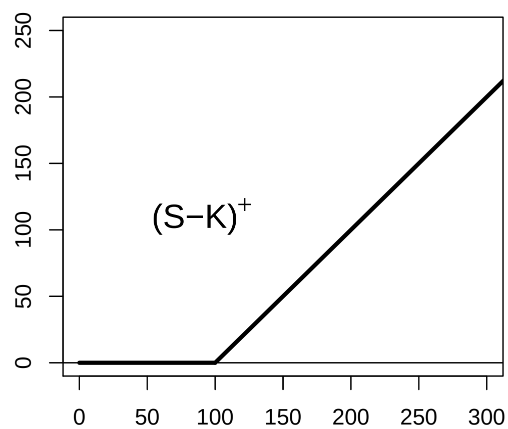

Figure 4.1: Call: Payoff $(S_N - K)^+$

Let C_n *denote the price (premium) of the option in time* n*. The contract specifies* C_n *at* $n = N$*: If* $S_T \leq K$*, then the option will not be exercised. If* $S_T > K$*, then* S_N *is bought for the price* K*. We have*

$$C_N = \begin{cases} 0 & \text{if } S_N \leq K \\ S_N - K & \text{if } S_N > K, \end{cases}$$

Thus

$$C_N = (S_T - K)^+ = \max\{S_N - K, 0\}.$$

The option is in the money, if $S_n > K$*, at the money, if* $S_n = K$ *or out of the money, if* $S_n < K$*.*

Example 4.7 American Call. *An American call is the right to buy the underlying for a price K at any time in the interval $[0, N]$. Again, if τ is the time when the option is exercised,*

$$C_\tau = (S_\tau - K)^+ = \max\{S_\tau - K, 0\}.$$

Example 4.8 European Put. *The European put gives the holder the right to sell the underlying at $n = N$ for the fixed price K.*

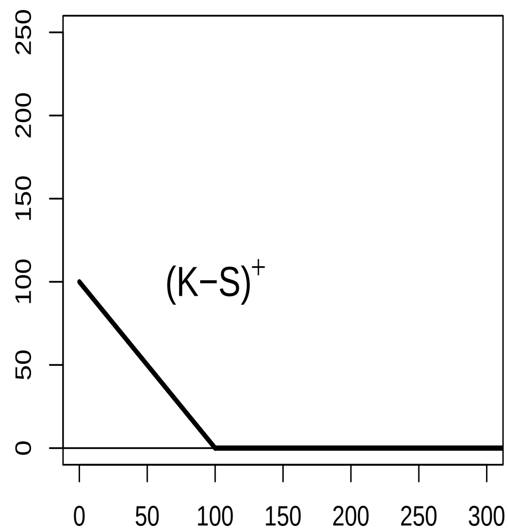

Figure 4.2: Put: Payoff $(K - S_N)^+$

Let P_n denote its price at time n. Again, P_N is specified by the contract. We have

$$P_N = (K - S_N)^+ = \max\{K - S_N, 0\}.$$

The American put is the right to sell the underlying for the price K at any time between 0 and N. Let $\tau \leq N$ be the exercise date, then

$$P_\tau = (K - S_\tau)^+ = \max\{K - S_\tau, 0\}.$$

The put is in the money, if $S_n < K$, at the money, if $S_n = K$ and out of the money, if $S_n > K$.

Example 4.9 Barrier Option. *Barriers allow, for instance, to exclude extreme situations and adapt options accordingly.*

$$(S_N - K)^+ I_{\{S_n \leq B \text{ for all } n \leq N\}}$$

is the payoff of a call with a knock-out barrier B. Let

$$M_n = \max\{S_n \mid 0 \le n \le N\}.$$

The payoff can be written as

$$(S_N - K)^+ I_{\{M_N \le B\}}.$$

A further example of a payoff with a barrier is

$$I_{\{M_N > B\}},$$

paying one unit, if the underlying is at least once (in the interval $[0, N]$) above B.

Example 4.10 Bonds, Interest Models. *The building block for interest rate models is the price of a zero-coupon bond. It pays one unit in $t = T$. Let $B_t(T)$ denote its price in t, $0 \le t \le T$. Various products may be written as linear combinations of zero-coupon bonds. For instance,*

$$C_1 B_t(T_1) + C_2 B_t(T_2) + \cdots + C_n B_t(T_n)$$

is a coupon-paying bond. It pays $C_1, C_2, \ldots, C_n$ at times $T_1, T_2, \ldots, T_n$.

Example 4.11 Exotic Options. *There exists a zoo of contingent claims and options that are typically traded "over the counter". For instance, the payoff of an* Asian option *is*

$$\left(\frac{1}{T}\int_0^T S_t\, dt - K\right)^+,$$

it is the payoff of a call on the average price of the asset (S_t). Other contracts have payoffs such as

$$(\max\{S_t \mid 0 \le t \le T\} - K)^+, \quad S_T - \min\{S_t \mid 0 \le t \le T\}, \quad \left(\max\{S_t \mid 0 \le t \le T\} - \frac{1}{T}\int_0^T S_t\, dt\right)^+.$$

It is often possible to derive prices or bounds on prices of certain contingent claims given prices of different assets by arbitrage considerations only, without having to specify a model. A simple but helpful tool for these parity considerations is the law of one price:

Proposition 4.12 *Let an arbitrage free model be given. Then the* law of one price (LOOP) *holds: If (X_n) and (Y_n) are two assets, such that $X_N = Y_N$ a.s., then for all $n = 0, \ldots, N$, $X_n = Y_n$ a.s.*

Proof. Suppose, $X_N = Y_N$, but $X_{n^*} > Y_{n^*}$ with positive probability. An arbitrage strategy is the following. In $n = n^*$, if $X_n > Y_n$, sell X and buy Y and put the positive gain (call it ϵ) into the bank account. At $n = N$ you have

$$Y_N - X_N + \epsilon \frac{S_N^0}{S_{n^*}^0} = \epsilon \frac{S_N^0}{S_{n^*}^0} > 0.$$

$\square$

Example 4.13 Put-Call Parity. *Assume that the model, containing a risky asset* (S_n) *and a deterministic riskless* (B_n) *is arbitrage free. Denote by* $P_n(N,K)$ *and* $C_n(N,K)$ *the prices of a put and a call on* (S_n) *with the same exercise date* N *and strike price* K*. Then for all* n,

$$C_n(N,K) - P_n(N,K) = S_n - KB_n/B_N. \qquad \textit{(put-call parity)}$$

To see that this parity holds, note that at $n = N$,

$$C_N(N,K) - P_N(N,K) = (S_N - K)^+ - (K - S_N)^+ = S_N - K.$$

The r.h.s. is the value of a portfolio, consisting of one unit of the risky asset and $K\,B_n/B_N$ *units of the bank account short (sold).*

Example 4.14 *Let the price of the asset be* $S_0 = 100$*, let* $B_t = e^{rt}$ *with* $r = 0.03$*. Furthermore* $C_0(1,100) = 26.8$*,* $P_0(1,100) = 24.5$*. We have to check whether an arbitrage opportunity exists. According to the put-call parity, the price of the put should be*

$$C_0(1,100) - S_0 + Ke^{-r} = 26.8 - 100 + 97.045 = 23.845 < 24.5.$$

Therefore we buy 1000 *calls, sell* 1000 *puts and* 1000 *units of the asset and put the difference,* 97700*, into the bank.*

In $n = 1$ *we have* 1000 *calls, and are* 1000 *puts and* 1000 *assets short and have*

$$97700 \times e^{0.03} = 100675.40$$

on the account. We can or have to buy the assets for a price of $1000 \times 100 = 100000$*. The gain is* 645.40*: If* $S_1 > K = 100$*, the puts expire, we exercise the calls and get* 1000 *assets. If* $S_1 \leq K = 100$*, the calls expire, the holder of the puts exercises, we have to buy the asset.*

Example 4.15 *Assume that the model, containing a risky asset* (S_n) *and a deterministic riskless* (B_n) *is arbitrage free. Denote by* $P_n(N,K)$ *and* $C_n(N,K)$ *the prices of a put and a call on* (S_n) *with exercise date* N *and strike price* K*.*

$$C_N(N,K) = (S_N - K)^+ \geq S_N - K$$

implies

$$C_n(N,K) \geq S_n - KB_n/B_N.$$

Since $C_n(N,K)$ *is always nonnegative, we get a lower bound for the call,*

$$C_n(N,K) \geq (S_n - KB_n/B_N)^+.$$

Analogously, we can show that

$$P_n(N,K) \geq (KB_n/B_N - S_n)^+.$$

Let $S_0 = 100$, $B_t = e^{rt}$ with $r = 0.03$, $K = 120$. The put $P_0(1, 120)$ costs 12.5.

It should cost at least $Ke^{-r} - S_0 = 16.45$. Since the price of the put is too low, we buy puts and assets S. For 100 assets and 100 puts we pay $100 \times 100 + 100 \times 12.5 = 11250$, which we finance by a credit from the bank account. In $n = 1$ we have 100 assets and puts and $-11250 \times e^{0.03} = -11592.61$ in the bank. We can sell the assets for a price of at least 120×100 gaining at least 407.39.

Example 4.16 *Assume that the model, containing a risky asset (S_n) and a deterministic riskless (B_n) is arbitrage free. Denote by $C_n(N, K)$ the price of a call on (S_n) with exercise date N and strike price K.*

We want to show that if $N_1 \leq N_2$, then

$$C_n(N_1, K) \leq C_n(N_2, K)$$

for $n \leq N_1$.

Assume, there exists a $n_0 \leq N_1 \leq N_2$ with $C_{n_0}(N_1, K) > C_{n_0}(N_2, K)$. In n_0, we buy the call with exercise date N_2 and sell the call with exercise date N_1. The difference is put into the bank.

We have to distinguish two cases: If $S_{N_1} \leq K$, the sold call expires, we have a call and a positive bank account. If $S_{N_1} > K$ the sold call is exercised, we get K for the asset, K is put into the bank. In $n = N_2$ we have

$$-S_{N_2} + (S_{N_2} - K)^+ + KB_{N_2}/B_{N_1} + (C_{n_0}(N_1, K) - C_{t_0}(N_2, K))B_{N_2}/B_{n_0}.$$

Since

$$-S_{N_2} + (S_{N_2} - K)^+ \geq -K,$$

we have at least

$$K(B_{N_2}/B_{N_1} - 1) + (C_{n_0}(N_1, K) - C_{n_0}(N_2, K))B_{N_2}/B_{n_0} > 0.$$

4.2 No Arbitrage

The following "Fundamental Theorem of Asset Pricing" gives a complete characterization of arbitrage free models. Recall that two probability distributions P and P^* are equivalent ($P \sim P^*$) if for all A, $P(A) = 0$ if and only if $P^*(A) = 0$.

Theorem 4.17 *(Fundamental Theorem). The market model satisfies NA if and only if there exists a probability distribution P^* on $(\Omega, \mathcal{F})$ such that $P \sim P^*$ and $(\bar{S}_n^i)$ is a martingale under P^* for all $i = 1, \ldots, m$.*

Remark 4.18 *Let $\mathcal{P} = \{P^* \mid P \sim P^*$ and $(\bar{S}_n^i)$ is a P^*-martingale$, i = 1, \ldots, m\}$. The elements of $\mathcal{P}$ are called* equivalent martingale measures *or* risk neutral distributions.

1. *The market model is arbitrage free if and only if $\mathcal{P} \neq \emptyset$.*

2. *Let the market model be arbitrage free, i.e. $\mathcal{P} \neq \emptyset$. Assume a (European) derivative with exercise date N and payoff h has to be priced. For any $P^* \in \mathcal{P}$,*

$$\bar{c}_n = \mathbb{E}^*(\bar{h} \mid \mathcal{F}_n) \tag{4.2}$$

gives a (discounted) arbitrage free price.

3. *Let $\mathcal{P} \neq \emptyset$. If $\mathcal{P}$ consists of exactly one element P^*, the (arbitrage free) prices of all derivatives are unique. If $\mathcal{P}$ has more than one element, it has infinitely many elements and the (arbitrage free) prices of derivatives are typically not unique. The set of all (arbitrage free) prices is then an interval.*

Proof (of Theorem 4.17). First assume that an equivalent martingale measure P^* exists. We show that arbitrage is not possible. Assume, that (V_n) is the value of a portfolio with $V_N \geq 0$ and $P(V_N > 0) > 0$. Then, since $S_N^0 > 0$ and $P^* \sim P$, we have $P^*(\bar{V}_N > 0) > 0$. Therefore, $\mathbb{E}^*(\bar{V}_N) > 0$. Since $\mathbb{E}^*(\bar{V}_N) = \mathbb{E}^*(\mathbb{E}^*(\bar{V}_N \mid \mathcal{F}_0)) = \mathbb{E}^*(V_0)$, it is impossible to have $V_0 = 0$ P^*-a.s.

The proof of the existence of a martingale measure in the case the market model is arbitrage free is much more complicated. We sketch the proof for the case of a finite probability space only. Let $\Omega = \{\omega_1, \ldots, \omega_K\}$, with $\mathcal{F}$ the power set of Ω and $P(\{\omega_i\}) = p_i > 0$. In this simple situation, expectations are scalar products, $\mathbb{E}^P(X) = \langle P, X \rangle = \sum_{i=1}^K p_i X(\omega_i)$, where $P = (p_1, \ldots, p_K)$ and $X = (X(\omega_1), \ldots, X(\omega_K))$.

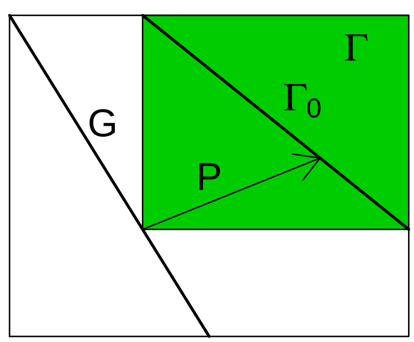

Figure 4.3: No Arbitrage

Denote by $\mathcal{G} \subseteq \mathbb{R}^K$ the set of all gains from self-financing strategies, i.e.

$$\mathcal{G} = \{\sum_{n=1}^{N} \langle \phi_n, \bar{S}_n - \bar{S}_{n-1} \rangle \mid (\phi_n) \text{ predictable}\}$$

and let

$$\Gamma = \{x \in \mathbb{R}^K \mid x_i \geq 0, x_i > 0 \text{ for at least one } i\}.$$

Furthermore, let $\Gamma_1 = \{x \in \Gamma \mid x_1 + \cdots + x_K = 1\}$. NA is equivalent to $\Gamma \cap \mathcal{G} = \emptyset$.

Let us have a closer look at the structure of Γ, Γ_1 and $\mathcal{G}$. $\mathcal{G}$ is a subspace of $\mathbb{R}^K$, i.e. if $G_1, G_2 \in \mathcal{G}$ and $a_1, a_2 \in \mathbb{R}$, then $a_1 G_1 + a_1 G_2 \in \mathcal{G}$. Γ is convex, Γ_1 is convex and compact (i.e. convex, bounded and closed).

Therefore, the sets $\mathcal{G}$ and Γ_1 can be separated by a hyperplane. There exists a $\tilde{P} = (\tilde{p}_1, \dots, \tilde{p}_K)$ in $\mathbb{R}^K$, s.t. for all $x \in \Gamma_1$, $\langle \tilde{P}, x \rangle > 0$ and $\langle \tilde{P}, G \rangle = 0$ for all $G \in \mathcal{G}$. Let e_i denote the unit vector, that has components 0 except the i-th, which is 1. $e_i \in \Gamma_1$ implies $0 < \langle \tilde{P}, e_i \rangle = \tilde{p}_i$. Let $P^* = \tilde{P} / \sum_{i=1}^{K} \tilde{P}_i$. P^* is a probability distribution and

$$\mathbb{E}^*(G) = 0$$

for all $G \in \mathcal{G}$. If the probability distribution P^* satisfies $\mathbb{E}^*(G) = 0$ for all $G \in \mathcal{G}$, then $(\bar{S}_n)$ is a martingale under P^*. □

4.3 Models

4.3.1 Model of Cox, Ross and Rubinstein

We have a deterministic bank account

$$B_n = e^{rn}$$

with $r \geq 0$ and one risky asset (S_n), a geometric random walk, $S_0 > 0$ is deterministic and for $k \geq 1$,

$$S_n = S_0 \prod_{k=1}^{n} Z_k$$

with (Z_k) i.i.d. and $P(Z_k = U) = p, P(Z_k = D) = 1 - p$, $0 < p < 1$ and and $0 < D < U$.

Let us first describe the distribution of $(S_n)_{n=0}^N$. It is discrete. There are exactly 2^N different paths. These paths $(s_n)_{n=0}^N$ can be written as $s_n = S_0 \prod_{k=1}^n z_k$ with $z_k \in \{U, D\}$. Let π_n denote the number of U's in the path up to time n. Then

$$S_n = S_0 U^{\pi_n} D^{n - \pi_n}.$$

Each path has probability

$$p^{\pi_N} (1-p)^{N - \pi_N} > 0.$$

Therefore, a distribution Q is equivalent to P if and only if each path has a strictly positive probability under Q.

Proposition 4.19 *The CRR-model is arbitrage free if and only if $0 < D < e^r < U$. The martingale measure P^* is then unique. Under P^* the process (S_n) is a geometric random walk with $P^*(Z_k = U) = p^*, P^*(Z_k = D) = 1 - p^*$, where*

$$p^* = \frac{e^r - D}{U - D}.$$

Proof. $(\bar{S}_n)$ is a geometric random walk with $\bar{S}_n = S_0 \prod_{k=1}^n \bar{Z}_k$, $(\bar{Z}_k)$ i.i.d. with $P(\bar{Z}_k = \bar{U}) = p, P(\bar{Z}_k = \bar{D}) = 1 - p$, where $\bar{U} = Ue^{-r}$ and $\bar{D} = De^{-r}$. Let P^* denote a probability distribution. $(\bar{S}_n)$ is a martingale under P^* if and only if

$$\mathbb{E}^*(\bar{S}_n \mid \mathcal{F}_{n-1}) = \bar{S}_{n-1}.$$

Since $\bar{S}_n = \bar{S}_{n-1}\bar{Z}_n$, this is equivalent to

$$\mathbb{E}^*(\bar{Z}_n \mid \mathcal{F}_{n-1}) = 1.$$

Let $p_n^* = p_n^*(S_1, \ldots, S_{n-1})$ denote the the conditional probability that $\bar{Z}_n = \bar{U}$, given $\mathcal{F}_{n-1}$, i.e. given $S_1, \ldots, S_{n-1}$.

$$\bar{U}p_n^* + \bar{D}(1 - p_n^*) = 1$$

holds if and only if

$$p_n^* = \frac{1 - \bar{D}}{\bar{U} - \bar{D}} \left(= \frac{e^r - D}{U - D} \right).$$

Note that the solution is unique, does not depend on n and not on $S_1, \ldots, S_{n-1}$. Therefore, even under P^*, (S_n) and $(\bar{S}_n)$ are geometric random walks, i.e. both processes (Z_n) and $(\bar{Z}_n)$ are i.i.d. $\square$

Remark 4.20 *The Radon-Nikodym derivative is*

$$\frac{dP^*}{dP} = \left(\frac{p^*}{p}\right)^{\pi_N} \left(\frac{1 - p^*}{1 - p}\right)^{N - \pi_N}.$$

Example 4.21 *The structure of the CRR-process, the binomial tree, allows a simple computation of conditional expectations and therefore of prices of derivatives.*

Let $U = 1.21$, $D = 0.88$, $P(Z_n = U) = 0.5$, $S_0 = 100$ and $e^r = 1.1$. We have to compute the price of a call with strike $K = 100$ and exercise time $N = 3$.

First, we have to compute the equivalent martingale measure. We have

$$p^* = \frac{e^r - D}{U - D} = \frac{1.1 - 0.88}{1.21 - 0.88} = \frac{2}{3}.$$

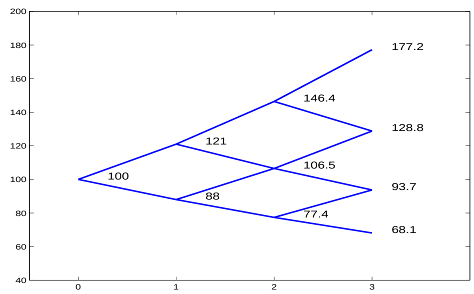

Figure 4.4: Binomial tree, $(S_n)_{n=0}^3$

To compute the price of the call, we may proceed as follows. Let $\bar{C}_n$ denote the discounted price at time n. Figure 4.4 shows the structure of the binomial tree. In $n = 0$ we have $S_n = 100$. In $n = 1$ the process is either in the knot $100 \times U = 121$ or in $100 \times D = 88$. Each knot has two successors. The upper knot corresponds to a jump up (a U), the lower to a jump down (a D). The probability (w.r.t. P^*) of an up is $2/3$, of a down $1/3$.

In the next step we compute the price at the exercise time $N = 3$, where the price is specified by the contract, i.e. by the payoff given. Note that the option is in the money only in the two upper knots. $\bar{C}_3$ is in the uppermost knot equal to

$$e^{-3r}(100U^3 - 100) = 58$$

and in the second knot

$$e^{-3r}(100U^2 D - 100) = 21.7.$$

Results are rounded. Then the tree is filled from right to left. Into each knot the weighted mean of the successor knots is written. For instance,

$$\begin{aligned} \bar{C}_2(S_2 = 100U^2) &= p^* \bar{C}_3(S_3 = 100U^3) + (1 - p^*)\bar{C}_3(S_3 = 100U^2 D) \\ &= \frac{2}{3} \times 58 + \frac{1}{3} \times 21.7 \\ &= 45.9 \end{aligned}$$

Finally one gets $C_0 = \bar{C}_0 = 26.8$.

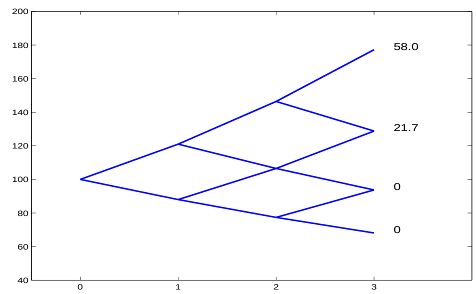

Figure 4.5: $\bar{C}_N$

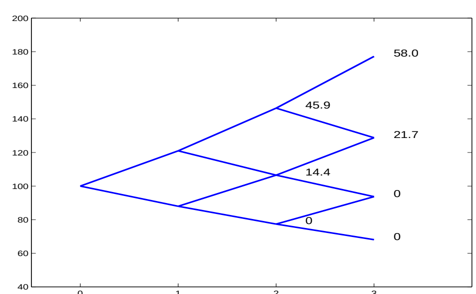

Figure 4.6: $\bar{C}_{N-1}$

Example 4.22 *The martingale measure and therefore the prices of derivatives depend on the size of the jumps and on the interest rate only, not on the physical probability, especially not on the probability that the option ends in the money.*

Let $U = 1.21$*,* $D = 0.99$*,* $e^r = 1.1$*, let* $S_0 = 100$ *and* $K = 220$ *the strike price of a call that*

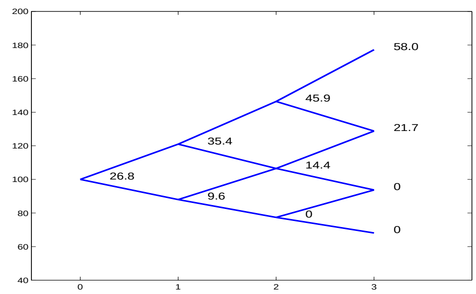

Figure 4.7: $\bar{C}_0$

expires in $N = 5$. *The option is exercised only if there are no downs, because*

$$100 \times 1.21^4 \times 0.99 = 212.22 < K.$$

We have $\bar{U} = 1.21/1.1 = 1.1$, $\bar{D} = 0.99/1.1 = 0.9$ *and*

$$p^* = \frac{1-0.9}{1.1-0.9} = \frac{0.1}{0.2} = \frac{1}{2}.$$

The price of the option is

$$C_0 = p^{*5} e^{-Nr}(S_0 U^5 - K) = 0.764.$$

Let $P(S_n/S_{n-1} = U) = p$. *If* $p = 0.01$, *then the probability that the option is exercised, is* $p^5 = 10^{-10}$. *If* $p = 0.99$ *it is* $p^5 = 0.951$.

4.3.2 Lognormal Returns

The lognormal model is a single period version of the famous model of Black and Scholes, a continuous time model. To derive the results in a "general form", we consider the two fixed time points $t < T$.

The bank account is deterministic, $B_u = e^{ru}, u \in \{t, T\}$. We have one risky asset, (S_u). $S_t > 0$ is the spot price (deterministic, i.e. known at t) and S_T the terminal value, it is random. We assume that $\log(S_T/S_t) \sim N((\mu - \sigma^2/2)(T-t), \sigma^2(T-t))$. Then

$$\begin{aligned} S_T &= S_t \exp\left((\mu - \sigma^2/2)(T-t) + \sigma\sqrt{T-t}Z\right), \\ \bar{S}_T &= \bar{S}_t \exp\left((\mu - r - \sigma^2/2)(T-t) + \sigma\sqrt{T-t}Z\right), \end{aligned}$$

with $Z \sim N(0,1)$.

Proposition 4.23 *1.*

$$\mathbb{E}(\bar{S}_T \mid \bar{S}_t) = \bar{S}_t e^{(\mu-r)(T-t)}.$$

$(\bar{S}_u)$ is a martingale (under P) if and only if $r = \mu$.
2. Let $\lambda = -(\mu - r)\sqrt{T-t}/\sigma$ and define $P^* \sim P$ by its Radon-Nikodym derivative

$$\frac{dP^*}{dP} = e^{\lambda Z - \lambda^2/2}.$$

$(\bar{S}_u)$ is a martingale under P^*, i.e. $\mathbb{E}^*(\bar{S}_T \mid \bar{S}_t) = \bar{S}_t$.

Proof. Remember that for $Z \sim N(0,1)$, $\mathbb{E}(e^{sZ}) = e^{s^2/2}$. Thus

$$\begin{aligned}
\mathbb{E}(\bar{S}_T \mid \bar{S}_t) &= \mathbb{E}(\bar{S}_t e^{(\mu-r-\sigma^2/2)(T-t)+\sigma\sqrt{T-t}Z}) \\
&= \bar{S}_t e^{(\mu-r-\sigma^2/2)(T-t)} \mathbb{E}(e^{\sigma\sqrt{T-t}Z}) \\
&= \bar{S}_t e^{(\mu-r-\sigma^2/2)(T-t)} e^{\sigma^2(T-t)/2} = \bar{S}_t e^{(\mu-r)(T-t)}.
\end{aligned}$$

Note that, again since $\mathbb{E}(e^{\lambda Z}) = e^{\lambda^2/2}$, $\mathbb{E}(dP^*/dP) = 1$ and therefore defines an equivalent distribution P^*. We show that under P^*, $Z \sim N(\lambda, 1)$. The characteristic function is

$$\begin{aligned}
\varphi_Z^*(s) &= \mathbb{E}^*(e^{isZ}) = \mathbb{E}(e^{isZ} e^{\lambda Z - \lambda^2/2}) \\
&= \int e^{isz} e^{\lambda z - \lambda^2/2} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz \\
&= \int e^{isz} \frac{1}{\sqrt{2\pi}} e^{-(z-\lambda)^2/2}\, dz \\
&= e^{is\lambda - s^2/2}.
\end{aligned}$$

Let $X = -\lambda + Z$. Then $X \sim^{P^*} N(0,1)$ and since

$$\bar{S}_T = \bar{S}_t e^{(\mu-r-\sigma^2/2)(T-t)+\sigma\sqrt{T-t}Z} = \bar{S}_t e^{-\sigma^2/2(T-t)+\sigma\sqrt{T-t}X},$$

$$\mathbb{E}^*(\bar{S}_T \mid \bar{S}_t) = \bar{S}_t.$$

□

Remark 4.24 *There exist infinitely many equivalent martingale measures for the lognormal model (see chapter 3). However, in the continuous time model the martingale measure is unique, it is the measure P^* of Proposition 4.23.*

Proposition 4.25 *(Pricing of contingent claims). Let a contingent claim (C_t, C_T) be defined by its payoff $C_T = h(S_T)$. An arbitrage free price is*

$$\bar{C}_t = \mathbb{E}^*(\bar{h}(S_T)) = \bar{F}(S_t),$$

where

$$F(x) = \int_{-\infty}^{\infty} e^{-r(T-t)} h(xe^{(r-\sigma^2/2)(T-t)+\sigma\sqrt{T-t}z})\phi(z)\,dz. \tag{4.3}$$

$\phi(z)$ is the density of the standard normal distribution.

Example 4.26 *(Put). To derive "a formula" for the price of the put we have to compute in closed form*

$$F(x) = \int e^{-r(T-t)}(K - xe^{(r-\sigma^2/2)(T-t)+\sigma\sqrt{T-t}z})^+\phi(z)dz.$$

First, we have to identify the interval an which the integrand is strictly positive. We abbreviate the time to expiration $T-t$ by θ. We have

$$K - xe^{(r-\sigma^2/2)\theta+\sigma\sqrt{\theta}z} > 0$$

iff

$$z < -\frac{\log(x/K) + (r-\sigma^2/2)\theta}{\sigma\sqrt{\theta}} =: -d_2(x).$$

Then

$$\begin{aligned} F(x) &= \int_{-\infty}^{-d_2(x)} e^{-r\theta}(K - xe^{(r-\sigma^2/2)\theta+\sigma\sqrt{\theta}z})\phi(z)dz \\ &= e^{-r\theta}K\Phi(-d_2(x)) - x\int_{-\infty}^{-d_2(x)} e^{-\theta\sigma^2/2+\sigma\sqrt{\theta}z}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz \\ &= e^{-r\theta}K\Phi(-d_2(x)) - x\int_{-\infty}^{-d_2(x)} \frac{1}{\sqrt{2\pi}}e^{-(z-\sqrt{\theta}\sigma)^2/2}dz \\ &= e^{-r\theta}K\Phi(-d_2(x)) - x\int_{-\infty}^{-d_2(x)-\sqrt{\theta}\sigma} \frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz \\ &= e^{-r\theta}K\Phi(-d_2(x)) - x\Phi(-d_1(x)), \end{aligned}$$

with

$$d_1(x) = \frac{\log(x/K) + (r+\sigma^2/2)\theta}{\sigma\sqrt{\theta}}, \tag{4.4}$$

$$d_2(x) = \frac{\log(x/K) + (r-\sigma^2/2)\theta}{\sigma\sqrt{\theta}}. \tag{4.5}$$

Figure 4.9 shows the price of the put for $t \in [0,1]$.

Example 4.27 *(Call) The Put-Call-parity simplifies the computation. $C_t(T,K,S_t) + Ke^{-r(T-t)} = P_t(T,K,S_t) + S_t$ (and $\theta = T-t$) implies $C_t(T,K,S_t) = F(S_t)$ with*

$$\begin{aligned} F(x) &= -Ke^{-r\theta} + e^{-r\theta}K\Phi(-d_2(x)) - x\Phi(-d_1(x)) + x \\ &= x(1-\Phi(-d_1(x))) - Ke^{-r\theta}(1-\Phi(-d_2(x))) \\ &= x\Phi(d_1(x)) - Ke^{-r\theta}\Phi(d_2(x)). \end{aligned} \tag{4.6}$$

Figure 4.10 shows the price of the call in the interval $t \in [0,1]$.

Table 4.1: "Greek Variables"

		Call	Put
Price	$F(x)$	$x\Phi(d_1(x)) - e^{-r\theta}K\Phi(d_2(x))$	$e^{-r\theta}K\Phi(-d_2(x)) - x\Phi(-d_1(x))$
Delta	$\frac{\partial F(x)}{\partial x}$	$\Phi(d_1(x))$	$-\Phi(-d_1(x))$
Gamma	$\frac{\partial^2 F(x)}{\partial x^2}$	$\frac{\phi(d_1(x))}{x\sigma\sqrt{\theta}}$	$\frac{\phi(d_1(x))}{x\sigma\sqrt{\theta}}$
Vega	$\frac{\partial F(x)}{\partial \sigma}$	$e^{-r\theta}K\sqrt{\theta}\phi(d_2(x))$	$e^{-r\theta}K\sqrt{\theta}\phi(d_2(x))$
Rho	$\frac{\partial F(x)}{\partial r}$	$e^{-r\theta}\theta K\Phi(d_2(x))$	$-e^{-r\theta}\theta K\Phi(-d_2(x))$

The sensitivity of the prices of options w.r.t. the underlying or the parameters σ and r are determined by the so-called Greeks, i.e. the partial derivatives.

Example 4.28 *The* Delta *is the sensitivity of the price w.r.t. the underlying,*

$$\Delta = \frac{\partial}{\partial x} F(x).$$

In the case of a call,

$$\Delta(x) = \Phi(d_1(x)).$$

The fluctuation of the Delta is the Gamma*:*

$$Gamma = \frac{\partial}{\partial x}\Delta(x).$$

The partial derivatives with respect to σ and r are called the Vega *and* Rho.

Example 4.29 *Binomial models are often used to approximate the lognormal model. This example discussed briefly the calibration, i.e. the choice of the parameters U, D and p.*

Let $\hat{S}_0 = S_t$ and let $\hat{S}_N = \hat{S}_0 \prod_{n=1}^{N} Z_n$ be an approximation to S_T. We have

$$\mathbb{E}(S_T \mid S_t) = S_t e^{\mu(T-t)} \qquad \text{and} \qquad \mathbb{V}(\log S_T) = \sigma^2(T-t).$$

We choose U, D and p to match these moments: Let $U = e^a$ and $D = 1/U = e^{-a}$. We have

$$\mathbb{E}(\hat{S}_N \mid \hat{S}_0) = \hat{S}_0 \left(pe^a + (1-p)e^{-a}\right)^n \qquad \text{and} \qquad \mathbb{V}(\log \hat{S}_N) = 4a^2 p(1-p)n.$$

The equations

$$\begin{aligned} \left(pe^a + (1-p)e^{-a}\right)^n &= e^{\mu(T-t)} \\ 4a^2 p(1-p)n &= \sigma^2(T-t) \end{aligned}$$

may be simplified to

$$\begin{aligned} p &= \frac{e^{\mu(T-t)/n} - e^{-a}}{e^a - e^{-a}}, \\ a &= \frac{\sigma\sqrt{T-t}}{2p(1-p)\sqrt{n}}. \end{aligned}$$

Note that for $n \to \infty$, $p \to 1/2$ *and* $a \approx \sigma\sqrt{T-t}/\sqrt{n}$.

4.4 Hedging

We assume that the market model is arbitrage free.

Definition 4.30 *A claim with payoff C_N is* attainable, *if there is a self-financing portfolio ϕ such that*
$V_N(\phi) = C_N$.

Proposition 4.31 *Let the claim C_N be attainable. Then $\mathbb{E}^*(\bar{C}_N)$ does not depend on the choice of $P^* \in \mathcal{P}$.*

Proof. *Let (V_n) denote the value of the portfolio that replicates C_N. The LOOP implies that (V_n) is unique. Then for all $P^* \in \mathcal{P}$,*

$$\mathbb{E}^*(\bar{C}_N) = \mathbb{E}^*(\bar{V}_N) = V_0.$$

□

Remark 4.32 *An attainable claim has a unique arbitrage free price.*

Definition 4.33 *A market model is* complete, *if every (square-integrable) claim is attainable.*

Theorem 4.34 *A viable market is complete if and only if the equivalent martingale measure is unique.*

Example 4.35 *Since the CRR-model has a unique martingale measure, it is complete. To find the replicating (hedging, duplicating) strategy $(\phi_n) = (\phi_n^0, \phi_n^1)$ of a claim C_N, we first identify (ϕ_n^1). Let (V_n) denote the value of the replicating portfolio. ϕ_n^1 is determined at time $n-1$. Note that*

$$\bar{V}_n = \phi_n^0 + \phi_n^1 \bar{S}_n.$$

Thus

$$\bar{V}_n - \phi_n^1 \bar{S}_n = \phi_n^0.$$

Since the r.h.s. is predictable, the l.h.s. is also predictable. Therefore $V_n - \phi_n^1 S_n$ is predictable and depends on S_{n-1}, not on S_n. Thus,

$$V_n(S_{n-1}U) - \phi_n^1 S_{n-1}U = V_n(S_{n-1}D) - \phi_n^1 S_{n-1}D$$

and

$$\phi_n^1 = \frac{V_n(S_{n-1}U) - V_n(S_{n-1}D)}{S_{n-1}U - S_{n-1}D}.$$

Finally, $\bar{V}_n = \phi_n^0 + \phi_n^1 \bar{S}_n$ gives ϕ_n^0. Note that ϕ_n^0 can alsobe drived from

$$\bar{V}_{n-1} = \phi_n^0 + \phi_n^1 \bar{S}_{n-1}.$$

Example 4.36 *The trinomial process is a geometric random walk,* $S_n = S_0 \prod_{k=1}^{n} Z_k$*, with* $Z_k \in \{U, M, D\}$*, where* $D < M < D$ *are constants. The riskless asset is deterministic,* $B_n = e^{rn}$*. Let* $p_U = P(Z_k = U)$*,* $p_M = P(Z_k = M)$*,* $p_D = P(Z_k = D)$ *and assume that these probabilities are strictly positive. Let* $\bar{U} = Ue^{-r}$*,* $\bar{M} = Me^{-r}$*,* $\bar{D} = De^{-r}$*.*

To find an equivalent martingale measure P^**, let* $p_n^*(U) = p_n^*(U; S_1, \ldots, S_{n-1})$ *denote the probability that* $S_n = S_{n-1}U$ *conditional on* $S_1, \ldots, S_{n-1}$*. Similarly define* $p_n^*(M)$ *and* $p_n^*(D)$*. These probabilities are the strictly positive solutions of*

$$\begin{aligned} p_n^*(U) + p_n^*(M) + p_n^*(D) &= 1 \\ p_n^*(U)\bar{U} + p_n^*(M)\bar{M} + p_n^*(D)\bar{D} &= 1 \text{ (martingale property).} \end{aligned}$$

If $U > e^r > D$ *there are infinitely many solutions. These solutions may also depend on* n *and on* $S_1, \ldots, S_{n-1}$*.*

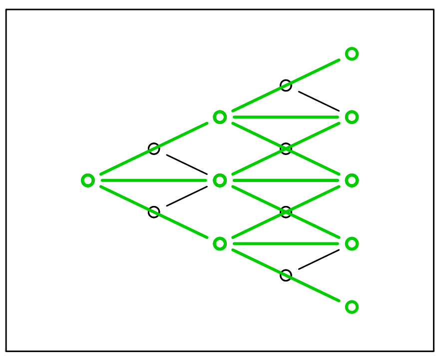

Figure 4.8: Trinomial tree

4.5 Exercises

Exercise 4.1 *The price of a call is* 57.1*. A put with the same exercise price* $K = 120$ *and expiration date* $T = 2$ *costs* 25.5*, the underlying costs* 130*. The interest rate is* 0.05*. Find an arbitrage opportunity!*

Exercise 4.2 *In an arbitrage-free market a call and a put cost* 57.1 *and* 35.5 *resp., the underlying*

120, the exercise price is $K = 120$ and $T = 1$ is the expiration date for both the call and the put. Compute the constant interest rate.

Exercise 4.3 *Let an arbitrage-free market consisting of an asset (S_n) and a deterministic riskless asset (B_n) be given. Let $P_n(N, K)$ and $C_n(N, K)$ denote the price of the put and the call at time n. K and N are the exercise price and the expiration date. Prove that for $K \leq K'$,*

$$C_n(N, K') \leq C_n(N, K) \leq C_n(N, K') + (K' - K)B_n/B_N,$$

$$P_n(N, K) \leq P_n(N, K') \leq P_n(N, K) + (K' - K)B_n/B_N.$$

Exercise 4.4 *Let an arbitrage-free market consisting of an asset (S_n) and a deterministic riskless asset (B_n) be given. Let $P_n(N, K)$ and $C_n(N, K)$ denote the price of the put and the call at time n. K and N are the exercise price and the expiration date. Let $\alpha \in [0, 1]$ and $K = \alpha K_1 + (1 - \alpha)K_2$. Prove that*

$$C_n(N, K) \leq \alpha C_n(N, K_1) + (1 - \alpha)C_n(N, K_2),$$

$$P_n(N, K) \leq \alpha P_n(N, K_1) + (1 - \alpha)P_n(N, K_2).$$

Exercise 4.5 *(Binary Option). A binary option has the payoff $h = I_A(S_N)$, where $A \subseteq \mathbb{R}$ is measurable (for instance an interval). Derive a parity relation for the European options with payoff $I_A(S_N)$ and $I_{A^c}(S_N)$.*

Exercise 4.6 *Derive a parity relation for calls, calls with knock-out barrier B and calls with knock-in barrier B.*

Exercise 4.7 *Show that if $\mathcal{P}$, the set of equivalent martingale measures, contains more than one element, it contains infinitely many.*
Hint. *Let $P_0, P_1 \in \mathcal{P}$. Show that for all $0 < \alpha < 1$, $P_\alpha = (1 - \alpha)P_0 + \alpha P_1 \in \mathcal{P}$.*

Exercise 4.8 *Prove that for any predictable process (ϕ_n^1) and any $\mathcal{F}_0$-measurable V_0 there exists a unique predictable process (ϕ_n^0) such that the strategy $(\phi_n) = ((\phi_n^0, \phi_n^1))$ is self-financing and its initial value is V_0.*

Exercise 4.9 *Consider the following modification of the CRR-model: Let $S_0 > 0$ be deterministic and $S_n = S_{n-1}Z_n$ with (Z_n) a sequence of independent random variables. Let $B_n = e^{r_1 + \cdots + r_n}$, $B_0 = 1$, r_n deterministic, and $Z_n \in \{U_n, D_n\}$, (U_n), (D_n) are nonrandom sequences. Derive the martingale measure.*

Exercise 4.10 *Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$, $P(\{\omega_i\}) = 1/3$, $B_0 = 1$, $B_1 = 1.1$, $B_2 = 1.2$, $S_0 = 2$, $S_1(\omega_1) = 2.2$, $S_1(\omega_2) = S_1(\omega_3) = 1.5$, $S_2(\omega_1) = 2.4$, $S_2(\omega_2) = 2.0$, $S_2(\omega_3) = 2.2$. Show that*

1. *The model allows arbitrage.*

2. *There exists a probability distribution Q, such that $(\bar{S}_n)_{n=0}^2$ is a Q-martingale.*

3. *There exists no equivalent martingale measure.*

Exercise 4.11 *Let X_0, $Z_1, \ldots, Z_N$, $U_1, \ldots, U_N$ be independent, (U_n) positive and integrable, $P(Z_n = 1) = P(Z_n) = -1) = 1/2$. Let $X_n = X_0 + U_1 Z_1 + \cdots + U_n Z_n$.*

Show that (X_n) is a martingale w.r.t. its history, but not w.r.t. the history of $(Y_n) = ((X_n, Z_{n+1}))$.

Discuss arbitrage opportunities (assume the interest rate is 0 and X_n is the price of an asset) of an investor A who observes (X_n) and of an investor B who knows additionally whether the asset moves up or down. What are the opportunities for investor C who observes (X_n, U_{n+1})?

Exercise 4.12 *Let (S_n) be a CRR-process with $P(S_{n+1}/S_n = U) = p = 0.7$, $U = 1.1$, $D = 0.9$, $S_0 = 100$, and nominal interest rate $r = 0$. Compute for the call with strike price $K = 100$ and expiration time $N = 4$ the prices and the probability that the options end in the money. Compute the replicating portfolio $\phi_3 = (\phi_3^0, \phi_3^1)$ for $S_2 = 99$.*

Exercise 4.13 *Let (S_n) be a CRR-process with $P(S_{n+1}/S_n = U) = p = 0.7$, $U = 1.1$, $D = 0.9$, $S_0 = 100$, and nominal interest rate $r = 0$. Compute for the binary option that pays one Euro if $S_4 \in [100, 130]$, the price and the probability that the option ends in the money. Compute the replicating portfolio $\phi_3 = (\phi_3^0, \phi_3^1)$ for $S_2 = 99$.*

Exercise 4.14 *Let (S_0, S_1) be a trinomial process (one-period model), with $U = 1.1, M = 1, D = 0.9$ and interest rate $r = 0$. Let $S_0 = 100$ and $P(S_1 = 110) = P(S_1 = 100) = P(S_1 = 90) = 1/3$.*

1. *Find all equivalent martingale measures.*

2. *Compute all arbitrage free prices of a call with strike price $K = 100$.*

3. *Let the price of a binary option, that pays 1 if $S_1 > 0.95$ cost 0.75. Compute the price of the call with strike price $K = 100$.*

Exercise 4.15 *Let (S_n) and (X_n) denote two assets. The option with payoff $(S_N - X_N)^+$ allows to exchange asset X for asset S in $n = N$. Assume that the two processes are independent CRR-processes with $S_0 = X_0 = 100$ and the same jump heights $U = 1.1$, $D = 0.8$. Furthermore assume that for both processes $p = 0.7$ and that $e^r = 1$. Let $N = 1$.*

Is the model arbitrage free? Is the martingale measure unique? Compute (all) prices.

Exercise 4.16 *Compute for the lognormal model the price of a binary option, which pays 1 Euro if $S_T \leq K$.*

Exercise 4.17 *Derive the Greeks for the Call in the lognormal model.*

Exercise 4.18 *Let a European option with expiration date T have the payoff $h(S_T)$, with*

$$h(x) = \begin{cases} 0 & \text{if } x \le K_1 \\ x - K_1 & \text{if } K_1 < x \le K_2 \\ K_2 - K_1 & \text{if } x > K_2. \end{cases} .$$

$K_1 < K_2$ are constants. Derive the price at t for the lognormal model.

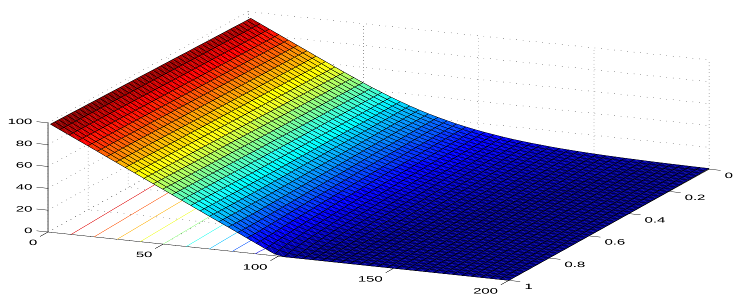

Figure 4.9: Put price surface, $K = 100, \sigma = 0.4, r = 0.05$

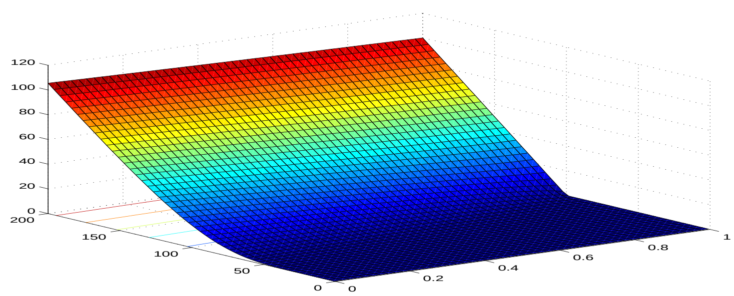

Figure 4.10: Call price surface, $K = 100, \sigma = 0.4, r = 0.05$

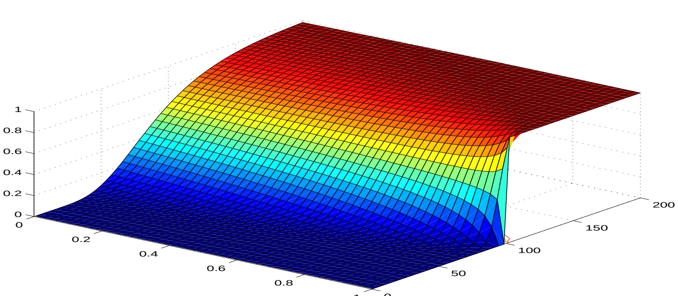

Figure 4.11: Delta of a Call, $K = 100, \sigma = 0.4, r = 0.05$

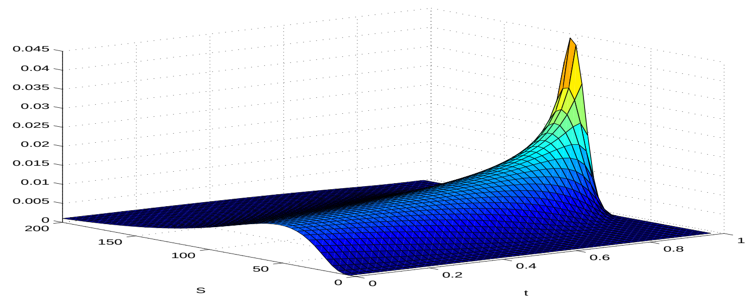

Figure 4.12: Gamma of a Call, $K = 100, \sigma = 0.4, r = 0.05$

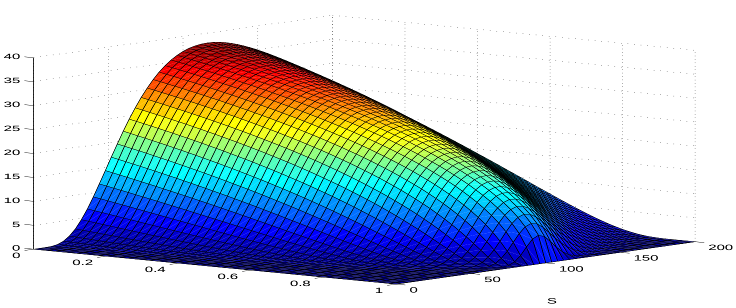

Figure 4.13: Vega, $K = 100, \sigma = 0.4, r = 0.05$

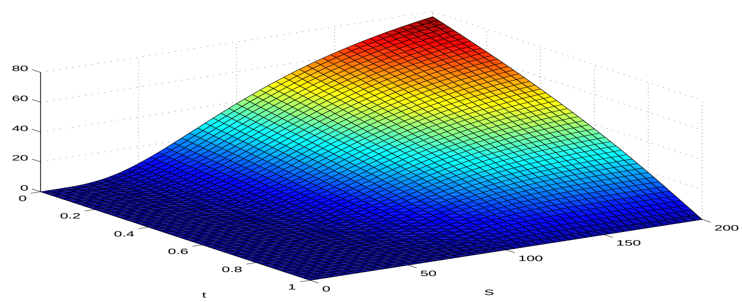

Figure 4.14: Rho of a Call, $K = 100, \sigma = 0.4, r = 0.05$

Chapter 5

American Options

5.1 The Problem

Let a multi-period financial model be given:

- $(\Omega, \mathcal{F}, P)$ a probability space.
- $(\mathcal{F}_n)_{n=0}^N$ a filtration $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_N \subseteq \mathcal{F}$, N a finite horizon.
- An adapted process $(S_n)_{n=0}^N$, $S_n = (S_n^0, S_n^1, \ldots, S_n^m)$. (S_n^0) is the numéraire (bank account), (S_n^k), $k > 0$ the underlyings, risky assets.
- $P^* \sim P$ the martingale measure, i.e. $(\bar{S}_n^k)$ are P^*-martingales, where $\bar{S}_n^k = S_n^k / S_k^0$ is the discounted k-th underlying. We assume that (discounted) prices of European claims with payoff Z in $n = N$ are given by $\bar{V}_n = \mathbb{E}^*(\bar{Z} \mid \mathcal{F}_n)$.

The American option is defined by its payoffs $(Z_n)_{n=0}^N$, it can be exercises at any time between 0 and N. If it is exercised at time n, the payoff is Z_n.

Let $m = 1$, $S_n = S_n^1$, $B_n = S_n^0$.

- $Z_n = (S_n - K)^+$ American call.
- $Z_n = (K - S_n)^+$ American put.
- $Z_n = b_n I_{C_n}(S_n)$ American binary option.

1. What is the price of an American option? For instance from the sellers perspective?
2. Choice of an optimal exercise time? Let $\mathcal{T}_{0,N}$ denote the class of stopping times with values in $\{0, 1, \ldots, N\}$. Choose $T \in \mathcal{T}_{0,N}$ that maximizes $\mathbb{E}^*(\bar{Z}_T)$ (among stopping times in $\mathcal{T}_{0,N}$).
3. If the market is complete, how can American options be hedged?

5.2 Prices

Let $(X_n)_{n=0}^N$ be adapted and let T denote a stopping time. $(X_n^T)_{n=0}^N$, denotes the process stopped at T, it is defined to be $(X_{T(\omega)\wedge n}(\omega))_{n=0}^N$, i.e. on $\{T = k\}$,

$$X_n^T = \begin{cases} X_k & \text{if } n \geq k, \\ X_n & \text{if } n < k. \end{cases}$$

Proposition 5.1 *If $(X_n)_{n=0}^N$ is a martingale (supermartingale, submartingale), the process $(X_n^T)_{n=0}^N$ is a martingale (supermartingale, submartingale).*

Proof. $(X_n^T)_{n=0}^N$ is a martingale transform,

$$X_n^T = X_0 + \sum_{k=1}^n I_{\{T \geq k\}}(X_k - X_{k-1}).$$

Therefore, it is a martingale, if $(X_n)_{n=0}^N$ is (see chapter 1). Similarly, if $(X_n)_{n=0}^N$ is a supermartingale, then

$$\begin{aligned} \mathbb{E}(X_n^T \mid \mathcal{F}_{n-1}) &= X_{n-1}^T + \mathbb{E}(I_{\{T \geq n\}}(X_n - X_{n-1}) \mid \mathcal{F}_{n-1}) \\ &= X_{n-1}^T + I_{\{T \geq n\}}\mathbb{E}(X_n - X_{n-1} \mid \mathcal{F}_{n-1}) \\ &= X_{n-1}^T + I_{\{T \geq n\}}(\mathbb{E}(X_n \mid \mathcal{F}_{n-1}) - X_{n-1}) \\ &\leq X_{n-1}^T, \end{aligned}$$

since $\mathbb{E}(X_n \mid \mathcal{F}_{n-1}) \leq X_{n-1}$ and $I_{\{T \geq n\}} \geq 0$. □

Let $(\bar{Z}_n)$ denote the (discounted) payoff and $(\bar{U}_n)$ the (discounted) price of the American claim. At time n, the seller of the options needs to have at least $\bar{Z}_n$, if the buyer exercises, or, if the buyer does not exercise, the amount necessary to generate $\bar{U}_{n+1}$ at time $n+1$, i.e. $\mathbb{E}^*(\bar{U}_{n+1} \mid \mathcal{F}_n)$. Therefore

$$\bar{U}_n \geq \max\{\bar{Z}_n, \mathbb{E}^*(\bar{U}_{n+1} \mid \mathcal{F}_n)\}.$$

Definition 5.2 *Let $(\bar{Z}_n)$ be adapted. Define $(\bar{U}_n)$ recursively by*

$$\bar{U}_n = \begin{cases} \bar{Z}_N & \text{if } n = N, \\ \max\{\bar{Z}_n, \mathbb{E}^*(\bar{U}_{n+1} \mid \mathcal{F}_n)\} & \text{if } 0 \leq n < N. \end{cases}$$

$(\bar{U}_n)$ is called the Snell envelope *of $(\bar{Z}_n)$.*

Proposition 5.3 *The Snell envelope $(\bar{U}_n)$ is the smallest supermartingale dominating $(\bar{Z}_n)$.*

Proof. Obviously, $(\bar{U}_n)$ is a supermartingale and $\bar{U}_n \geq \bar{Z}_n$ for all n. To show that it is the smallest, let $(\bar{V}_n)$ denote another supermartingale dominating $(\bar{Z}_n)$. We have to show that $\bar{U}_n \leq \bar{V}_n$.

Since $\bar{U}_N = \bar{Z}_N$ and $\bar{V}_N \geq \bar{Z}_N$, we have $\bar{U}_N \leq \bar{V}_N$. Now suppose, that the statement holds for $n+1, n+2, \ldots, N$. Then

$$\bar{V}_n \geq \mathbb{E}^*(\bar{V}_{n+1} \mid \mathcal{F}_n) \geq \mathbb{E}^*(\bar{U}_{n+1} \mid \mathcal{F}_n).$$

and $\bar{V}_n \geq \bar{Z}_n$ and therefore

$$\bar{V}_n \geq \max\{\bar{Z}_n, \mathbb{E}^*(\bar{U}_{n+1} \mid \mathcal{F}_n)\} = \bar{U}_n.$$

□

We define the discounted price of the American claim to be the Snell envelope of the discounts payoffs.

5.3 Optimal Exercise Times

Let $T \in \mathcal{T}_{0,N}$ be a stopping time. Note that

$$\bar{U}_T = \bar{U}_N^T.$$

Since $(\bar{U}_n)$ is a supermartingale, we have

$$\mathbb{E}^*(\bar{U}_0) = \mathbb{E}^*(\bar{U}_0^T) \geq \mathbb{E}^*(\bar{U}_N^T) = \mathbb{E}^*(\bar{U}_T) \geq \mathbb{E}^*(\bar{Z}_T).$$

Thus,

$$\max_{T \in \mathcal{T}_{0,N}} \mathbb{E}^*(\bar{Z}_T) \leq \mathbb{E}^*(\bar{U}_0).$$

We show that there is a stopping time T for which $\mathbb{E}^*(\bar{Z}_T) = \mathbb{E}^*(\bar{U}_0)$. Therefore, a stopping time T is optimal if and only if $\mathbb{E}^*(\bar{Z}_T) = \mathbb{E}^*(\bar{U}_0)$. Furthermore, note that for an optimal stopping time,

1. if we stop at $T = n$, then $\bar{Z}_n = \bar{U}_n$, and thus $\bar{Z}_n \geq \mathbb{E}^*(\bar{U}_{n+1} \mid \mathcal{F}_n)$, i.e. stopping when $\bar{Z}_n < \mathbb{E}^*(\bar{U}_{n+1} \mid \mathcal{F}_n)$ is not optimal,
2. if $\mathbb{E}^*(\bar{U}_0) = \mathbb{E}^*(\bar{U}_T)$, then the stopped supermartingale $(\bar{U}_n^T)$ is in fact a martingale.

Let $\bar{U}_n = M_n - A_n$ denote the Doob decomposition of the supermartingale $(\bar{U}_n)$ into a martingale (M_n) and an increasing predictable process (A_n). We define the minimal and the maximal optimal stopping time:

$$T_{\min} = \min\{n \mid \bar{Z}_n = \bar{U}_n\}, \qquad (5.1)$$

$$T_{\max} = \max\{n \mid A_n = 0\}. \qquad (5.2)$$

Note that $T_{\max} = \min\{n \mid A_{n+1} > 0\}$ is a stopping time, since (A_n) is predictable.

Theorem 5.4 *(a) T_{min} and T_{max} are optimal stopping times.*
(b) A stopping time T is optimal, iff $\bar{Z}_T = \bar{U}_T$ and $(\bar{U}_n^T)$ is a martingale.
(c) For all optimal stopping times T, $T_{min} \leq T \leq T_{max}$.

Proof. (a) Let $T = T_{\min}$.

$$\bar{U}_T = \bar{U}_0 + \sum_{k=1}^{N} I_{\{T \geq k\}}(\bar{U}_k - \bar{U}_{k-1})$$

and $\bar{U}_{k-1} = \mathbb{E}^*(\bar{U}_k \mid \mathcal{F}_{k-1})$ on $\{k \leq T\}$ implies

$$\mathbb{E}^*(I_{\{T \geq k\}}(\bar{U}_k - \bar{U}_{k-1})) = \mathbb{E}^*(I_{\{T \geq k\}}(\bar{U}_k - \mathbb{E}^*(\bar{U}_k \mid \mathcal{F}_{k-1})) = 0$$

and therefore $\mathbb{E}^*(\bar{U}_T) = \mathbb{E}^*(\bar{U}_0)$, i.e. $T_{\min}$ is optimal.

The definition of $T_{\max}$ implies that $\bar{U}_{T_{\max}} = M_{T_{\max}}$, which implies

$$\mathbb{E}^*(\bar{U}_{T_{\max}}) = \mathbb{E}^*(M_{T_{\max}}) = \mathbb{E}^*(M_0) = \mathbb{E}^*(\bar{U}_0).$$

(b) From

$$\mathbb{E}^*(\bar{U}_0) \geq \mathbb{E}^*(\bar{U}_T) \geq \mathbb{E}^*(\bar{Z}_T) = \mathbb{E}^*(\bar{U}_0)$$

we conclude that $\bar{Z}_T = \bar{U}_T$ a.s. and $\mathbb{E}^*(\bar{U}_{n+1}^T \mid \mathcal{F}_n) = \bar{U}_n^T$.

(c) (b) implies that optimal stopping times T satisfy $\bar{Z}_T = \bar{U}_T$ and therefore $T \geq T_{\min}$. Assume, $P(T > T_{\max}) > 0$. Then $\bar{U}_T < M_T$ with positive probability and thus

$$\mathbb{E}^*(\bar{U}_T) < \mathbb{E}^*(M_T) = \mathbb{E}^*(M_0) = \mathbb{E}^*(\bar{U}_0),$$

i.e. T is not optimal. □

Remark 5.5 *Although for all optimal stopping times T, $T_{min} \leq T \leq T_{max}$, in the general case not all stopping times between T_{min} and T_{max} are optimal!*

5.4 Call Options and Examples

Let $(\bar{U}_n)$ denote the price of an American claim with payoff $(\bar{Z}_n)$ and $(\bar{u}_n)$ the price of the European claim with payoff $\bar{Z}_N$. Note that $(\bar{u}_n)$ is a martingale. If $\bar{u}_n \geq \bar{Z}_n$ for all n, then $\bar{u}_n \geq \bar{U}_n$ for all n and thus $\bar{u}_n = \bar{U}_n$.

Example 5.6 *The prices of the European and the American call options are the same (if no dividends are paid). Denote the prices by $\bar{c}_n$ and $\bar{C}_n$. We assume that the bank account (B_n) is nondecreasing. We have*

$$\begin{aligned} \bar{c}_n &= \mathbb{E}^*(\frac{1}{B_N}(S_N - K)^+ \mid \mathcal{F}_n) \\ &= \mathbb{E}^*((\bar{S}_N - K/B_N)^+ \mid \mathcal{F}_n). \end{aligned}$$

Since

$$(\bar{S}_N - K/B_N)^+ \geq \bar{S}_N - K/B_N \geq \bar{S}_N - K/B_n$$

we get

$$\begin{aligned}\bar{c}_n &\geq \mathbb{E}^*(\bar{S}_N - K/B_n \mid \mathcal{F}_n)\\ &= \bar{S}_n - K/B_n.\end{aligned}$$

Furthermore, $\bar{c}_n \geq 0$, and therefore

$$\bar{c}_n \geq (\bar{S}_n - K/B_n)^+ = \bar{Z}_n.$$

Example 5.7 *Figure 5.1 shows the binomial tree of a CRR-process $(S_n)_{n=0}^{15}$ with $S_0 = 20$. We have $U = 1.05$, $D = 0.9$ and $r = 0$. The price of an American claim with payoff*

$$Z_n = \begin{cases} 0 & \text{if } S_n \leq 16 \\ S_n - 16 & \text{if } 16 < S_n < 25 \\ 9 & \text{if } S_n \geq 25 \end{cases}$$

has to be computed. This claim is equivalent to the difference of two calls with strike prices 16 and 25. A path starts at the root and runs through the tree from left to right. There are three kind of knots: Blue knots are continuation knots, here $U_n > Z_n$. If the path reaches such a knot, the option is not exercised. The remaining knots are green, they are in the exercise region. Here we have $U_n = Z_n$. If the path reaches such a knot, the option may be exercised. If the green knot is additionally labeled with a red circle, the maximal optimal stopping time is reaches, the option has to be exercised.

The price of the option is $U_0 = 4.6020$.

5.5 Exercises

Exercise 5.1 *Prove that if $(X_n)_{n=0}^{\infty}$ is a supermartingale, then for all $n \geq 0$, $\mathbb{E}(X_0) \geq \mathbb{E}(X_n)$ holds.*

Exercise 5.2 *Let the bank account be constant, $B_n = 1$. Show that the prices of the American and the European put are the same. Hint: Use Jensen's inequality.*

Exercise 5.3 *Consider the CRR-model with "up" $= 1.21$, "down" $= 0.99$, $B_n = 1.1^n$, $S_0 = 1$ and $N = 4$. Find the price of the American claim with discounted payoff $\bar{Z}_n = 1$ if $S_n \leq 0.985$ or $1.3 \leq S_n \leq 2$ and $\bar{Z}_n = 0$ in all other cases. Compute $(\bar{Z}_n)$, $(\bar{U}_n)$, $(\mathbb{E}^*(U_{n+1} \mid \mathcal{F}_n))$, (A_n), T_{max} and T_{min}.*

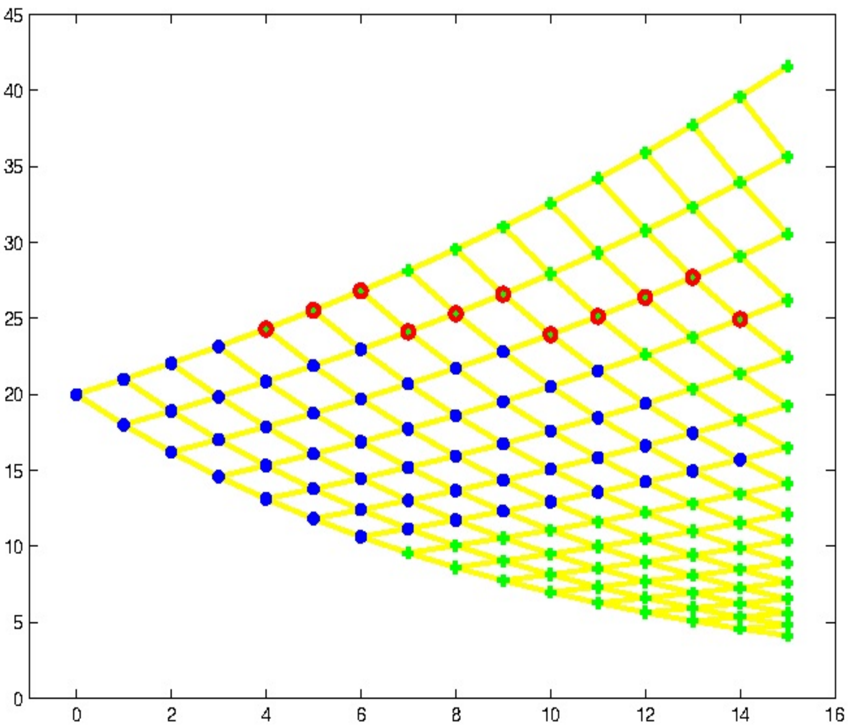

Figure 5.1: Claim $C_{16} - C_{25}$

Chapter 6

Brownian Motion and Boundary Crossing Probabilities

6.1 Definition, Properties and Related Processes

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A stochastic process in continuous time $(X_t)_{0\leq t<\infty}$ is a collection of random variables $X_t : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$.

Definition 6.1 *A stochastic process $(W_t)_{0\leq t<\infty}$ is called a* Brownian motion *if*

1. $W_0 = 0$,
2. *paths $t \mapsto W_t(\omega)$ are continuous a.s.,*
3. *for all $0 < t_1 < t_2 < \cdots < t_n$ the increments W_{t_1}, $W_{t_2} - W_{t_1}$, $\ldots W_{t_n} - W_{t_{n-1}}$ are independent and normally distributed with $W_{t_i} - W_{t_{i-1}} \sim N(0, t_i - t_{i-1})$.*

The Brownian motion is also called *Wiener process*.

Proposition 6.2 *Let (W_t) be a Brownian motion. Then*

1. *(W_t) is Gaussian process with $\mathbb{E}(W_t) = 0$ and $\mathbb{C}ov(W_s, W_t) = s \wedge t$.*
2. *(W_t) is a martingale.*
3. *$(W_t^2 - t)$ is a martingale.*

Remark 6.3 *A process (X_t) is Gaussian, if for all $t_1, \ldots, t_n$, the distribution of $(X_{t_1}, \ldots, X_{t_n})$ is Gaussian. To check the distribution of a Gaussian process, it is sufficient to check the distribution of (X_s, X_t) for all s, t, i.e. to check $\mathbb{E}(W_t)$ and $\mathbb{C}ov(W_s, W_t)$.*

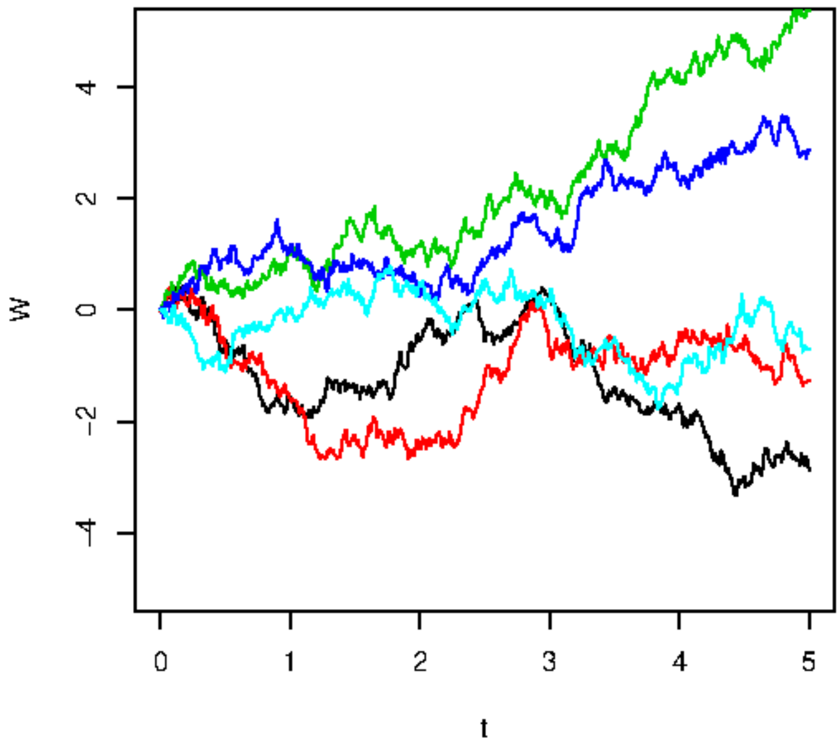

Figure 6.1: Paths of Brownian motion

Proof. Let $s \leq t$. $W_t - W_s$ is independent of $\mathcal{F}_s = \sigma(W_u, u \leq s)$.

1. (W_t) has independent normally distributed increment. Therefore it is a Gaussian process. $\mathbb{E}(W_t) = 0$ by definition. It follows that $\mathbb{C}ov(W_s, W_t) = \mathbb{E}(W_s W_t)$. Furthermore $\mathbb{E}(W_s(W_t - W_s)) = 0$. Then

$$\begin{aligned}
\mathbb{E}(W_s W_t) &= \mathbb{E}(W_s(W_t - W_s + W_s)) \\
&= \mathbb{E}(W_s(W_t - W_s)) + \mathbb{E}(W_s^2) \\
&= 0 + s = s.
\end{aligned}$$

2. Since for $W_t - W_s$ is independent of $\mathcal{F}_s = \sigma(W_u, u \leq s)$, we have

$$\mathbb{E}(W_t \mid \mathcal{F}_s) = \mathbb{E}(W_t - W_s + W_s \mid \mathcal{F}_s) = \mathbb{E}(W_t - W_s \mid \mathcal{F}_s) + W_s = W_s.$$

3. We have

$$(W_t^2 - t) - (W_s^2 - s) = (W_t - W_s)^2 + 2W_s(W_t - W_s) - (t - s)$$

and

$$\begin{aligned}
\mathbb{E}((W_t - W_s)^2 + 2W_s(W_t - W_s) - (t-s) \mid \mathcal{F}_s) &= \mathbb{E}((W_t - W_s)^2 \mid \mathcal{F}_s) \\
&\quad +2\mathbb{E}(W_s(W_t - W_s) \mid \mathcal{F}_s) - (t-s) \\
&= \mathbb{E}((W_t - W_s)^2) + 2W_s\mathbb{E}(W_t - W_s \mid \mathcal{F}_s) - (t-s) \\
&= t - s + 0 - (t-s) = 0.
\end{aligned}$$

□

The paths of the Brownian motion are continuous. However, with probability 1, they are nowhere differentiable and the length of the paths on each finite interval is not finite.

Proposition 6.4 *Let (W_t) be a Brownian motion. Then, for every $t > 0$ and every sequence of $0 = t_0^n < t_1^n < \cdots < t_n^n = t$ of partitions of $[0, t]$ with $\lim_{n\to\infty} \max_{i\le n} |t_{i-1}^n - t_i^n| = 0$,*

$$\sum_{i=1}^n (W_{t_i^n} - W_{t_{i-1}^n})^2 \to^P t.$$

Cf. Proposition 6.2. The quadratic variation of the Brownian motion is deterministic and equals the identity, $Q(t) = t$.

Proof. Let

$$Q_n(t) = \sum_{i=1}^n (W_{t_i^n} - W_{t_{i-1}^n})^2.$$

$Q_n(t)$ is a random variable. We have to show that for all $\epsilon > 0$, $\lim_{n\to\infty} P(|Q_n(t) - t| > \epsilon) = 0$. We apply Chebyshev's inequality:

$$P(|Q_n(t) - \mathbb{E}(Q_n(t))| > \epsilon) \le \frac{\sigma^2(Q_n(t))}{\epsilon^2}.$$

We have

$$\mathbb{E}(Q_n(t)) = \sum_{i=1}^n \mathbb{E}((W_{t_i^n} - W_{t_{i-1}^n})^2) = \sum_{i=1}^n (t_i^n - t_{i-1}^n) = t$$

and

$$\sigma^2(Q_n(t)) = \sum_{i=1}^n \sigma^2((W_{t_i^n} - W_{t_{i-1}^n})^2) = \sum_{i=1}^n 2(t_i^n - t_{i-1}^n)^2$$

and therefore

$$\sigma^2(Q_n(t)) \le 2\max_{i\le n} |t_i^n - t_{i-1}^n| \sum_{i=1}^n (t_i^n - t_{i-1}^n) = 2\max_{i\le n} |t_i^n - t_{i-1}^n| t \to 0.$$

□

Definition 6.5 *Let (W_t) be a Brownian motion, $\mu, x_0 \in \mathbb{R}$, $\sigma, s_0 > 0$.*

1. *The* generalized Brownian motion $(X_t)_{0\leq t<\infty}$ *is defined by* $X_t = x_0 + \mu t + \sigma W_t$.

2. *The geometric* Brownian motion $(S_t)_{0\leq t<\infty}$ *is the process* $S_t = s_0 e^{\mu t+\sigma W_t}$.

3. *A* Brownian bridge $(U_t)_{0\leq t\leq 1}$ *is a Gaussian process with continuous paths and* $\mathbb{E}(U_t) = 0$ *and* $\mathbb{C}ov(U_s, U_t) = s(1-t)$ *for* $0 \leq s \leq t \leq 1$.

6.2 First Exit Times

6.2.1 Linear One-Sided Boundaries

Let (W_t) be a Brownian motion with $(\mathcal{F}_t)$ its filtration. Let $T > 0$ a a finite or infinite time-horizon, $a, b : [0,T] \to \mathbb{R}$ continuous with $a(t) \leq b(t)$ for all $t \in [0,T]$ and $a(0) \leq 0 \leq b(0)$. a and b are the lower - and the upper boundary. We consider the boundary crossing time

$$\begin{aligned} \tau = & \quad \inf\{t \geq 0 \mid W_t \geq b(t)\} \quad && \text{(one-sided case)}, \\ \text{or} & \\ \tau = & \quad \inf\{t \geq 0 \mid W_t \geq b(t) \text{ or } W_t \leq a(t)\} && \text{(two-sided case)}. \end{aligned}$$

τ is a stopping time, also called the first exit time or the first passage time. The corresponding boundary crossing probabilities are

$$P(b\,;T) = P\left(W_t \geq b(t) \text{ for a } t \in [0,T]\right), \qquad \text{(one-sided case)} \tag{6.1}$$

$$P(a,b\,;T) = P\left(W_t \geq b(t) \text{ or } W_t \leq a(t) \text{ for a } t \in [0,T]\right). \qquad \text{(two-sided case)} \tag{6.2}$$

Note that $P(b\,;T) = P(\tau \leq T)$. In the one-sided case the distribution of τ is related to the distribution of $\sup_{0\leq s\leq t}(W_s - b(s))$:

$$\begin{aligned} P(\tau \leq t) &= P(W_s \geq b(s) \text{ for an } s \in [0,t]) \\ &= P(W_s - b(s) \geq 0 \text{ for an } s \in [0,t]) \\ &= P\left(\sup_{0\leq s\leq t}(W_s - b(s)) \geq 0\right). \end{aligned}$$

Let us consider the one-sided case with either a constant boundary $b(t) = \alpha$ or a linear boundary $b(t) = \alpha + \beta t$, where $\alpha > 0$. We denote the boundary crossing time τ by τ^α or $\tau^{\alpha,\beta}$ to indicate the constant or the linear problem. Furthermore, define

$$W_t^* = \sup_{0\leq s\leq t} W_t.$$

Theorem 6.6 *Let* $\alpha \geq 0$. *Define* $\bar{\Phi}(x) = 1 - \Phi(x)$. *Then*

$$P(W_t^* \geq \alpha) = 2\bar{\Phi}(\alpha/\sqrt{t}). \tag{6.3}$$

Proof. To derive the law of W_t^*, we apply the following reflection principle (for a proof see Proposition 6.8): Define for a path $(W_t(\omega))$ that crosses α its reflection at α $\tilde{W}_t(\omega)$ as $\tilde{W}_t(\omega) = W_t(\omega)$ for $t \le \tau$ and $\tilde{W}_t(\omega) = \alpha - (W_t(\omega) - \alpha) = 2\alpha - W_t(\omega)$ for $t > \tau$. τ denotes the first crossing time of the boundary α. Then the events $\{W_t \le \alpha \text{ and } W_t^* \ge \alpha\}$ and $\{\tilde{W}_t \ge \alpha\}$ have the same probability. Note that $(\tilde{W}_t)$ is again a Brownian motion. Therefore $P(\tilde{W}_t \ge \alpha) = P(W_t \ge \alpha)$.

Assume that this reflection principle holds. Then, since $\{W_t \ge \alpha\} \subseteq \{W_t^* \ge \alpha\}$,

$$\begin{aligned}
P(W_t^* \ge \alpha) &= P(W_t^* \ge \alpha, W_t > \alpha) + P(W_t^* \ge \alpha, W_t \le \alpha) \\
&= P(W_t^* \ge \alpha, W_t \ge \alpha) + P(W_t^* \ge \alpha, W_t \le \alpha) \\
&= P(W_t \ge \alpha) + P(W_t \ge \alpha) \\
&= 2P(W_t \ge \alpha) = 2\bar{\Phi}(\alpha/\sqrt{t}).
\end{aligned}$$

□

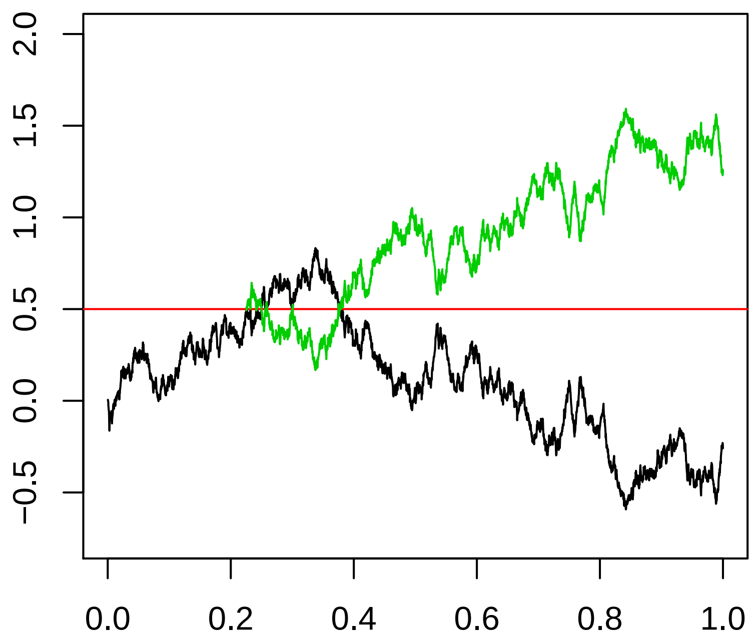

Figure 6.2: Paths of Brownian motion

For the proof of the reflection principle we need the following lemma.

Lemma 6.7 *Let f be bounded and measurable and define $g(s) = \mathbb{E}(f(W_s + \alpha))$. Then*

$$\mathbb{E}(f(W_t)I_{\{\tau^\alpha \le t\}}) = \mathbb{E}(g(t - \tau^\alpha)I_{\{\tau^\alpha \le t\}}).$$

Proof. We have

$$\begin{aligned}
\mathbb{E}(f(W_t)I_{\{\tau^\alpha \le t\}}) &= \mathbb{E}(f(W_t - W_{\tau^\alpha} + W_{\tau^\alpha})I_{\{\tau^\alpha \le t\}}) \\
&= \mathbb{E}(f(W_t - W_{\tau^\alpha} + \alpha)I_{\{\tau^\alpha \le t\}})
\end{aligned}$$

Approximate τ^α by a sequence of stopping times τ_n that have only countably many values $(t_{n,i})_{i=1}^\infty$ s.t. $\tau_n \downarrow \tau^\alpha$. For such stopping times we have

$$\mathbb{E}(f(W_t - W_{\tau_n} + \alpha)I_{\{\tau_n \le t\}}) = \sum_{i:\, t_{n,i} \le t} \mathbb{E}(f(W_t - W_{t_{n,i}} + \alpha)I_{\{\tau_n = t_{n,i}\}}).$$

Note that $W_t - W_{t_{n,i}}$ has the same distribution as $W_{t-t_{n,i}}$ and is independent of τ_n (strong Markov property). Therefore we get

$$\begin{aligned}
\sum_{i:\, t_{n,i} \le t} \mathbb{E}(f(W_t - W_{t_{n,i}} + \alpha)I_{\{\tau_n = t_{n,i}\}}) &= \sum_{i:\, t_{n,i} \le t} \mathbb{E}(g(t - t_{n,i})I_{\{\tau_n = t_{n,i}\}}) \\
&= \mathbb{E}(g(t - \tau^n)I_{\{\tau_n \le t\}}) \\
&\to \mathbb{E}(g(t - \tau^\alpha)I_{\{\tau^\alpha \le t\}}).
\end{aligned}$$

□

Proposition 6.8 *(Reflection Principle). Let f be bounded and measurable. Then*

$$\mathbb{E}(f(W_t)I_{\{\tau^\alpha \le t\}}) = \mathbb{E}(f(2\alpha - W_t)I_{\{\tau^\alpha \le t\}}).$$

Proof. Note that $\mathbb{E}(f(W_s + \alpha)) = \mathbb{E}(f(-W_s + \alpha))$ and apply Lemma 6.7. Let $(\tilde{W}_t)$ be a Brownian motion independent of (W_t).

$$\begin{aligned}
\mathbb{E}(f(W_t)I_{\{\tau^\alpha \le t\}}) &= \mathbb{E}(\mathbb{E}(f(\tilde{W}_s + \alpha) \mid s = t - \tau^\alpha)I_{\{\tau^\alpha \le t\}}) \\
&= \mathbb{E}(\mathbb{E}(f(-\tilde{W}_s + \alpha) \mid s = t - \tau^\alpha)I_{\{\tau^\alpha \le t\}}) \\
&= \mathbb{E}(f(-(W_t - W_{\tau^\alpha}) + \alpha)I_{\{\tau^\alpha \le t\}}) \\
&= \mathbb{E}(f(2\alpha - W_t)I_{\{\tau^\alpha \le t\}}).
\end{aligned}$$

□

Theorem 6.9 *(Joint distribution of W_t and W_t^*). (a) Let $\alpha \ge 0$, $\mu \le \alpha$.*

$$P(W_t \le \mu, W_t^* \le \alpha) = P(W_t \le \mu) - P(W_t \ge 2\alpha - \mu) \tag{6.4}$$

$$P(W_t^* \le \alpha) = 2P(W_t \le \alpha) - 1 \tag{6.5}$$

(b) (W_t, W_t^*) *has a density* $f(x, y)$, *with* $f(x, y) = 0$ *if* $y < 0$ *or* $x > y$ *and*

$$f(x, y) = \frac{2(2y - x)}{\sqrt{2\pi t^3}} e^{-(2y-x)^2/2t}, \qquad \textit{else.}$$

(c) Let $x \leq \alpha$. *Then*

$$P(W_t^* \geq \alpha \mid W_t = x) = e^{-2\alpha(\alpha - x)/t}. \tag{6.6}$$

Proof. (a) Apply the reflection principle to $f(W_t) = I_{\{W_t \leq \mu\}}$. Then

$$\begin{aligned} P(W_t \leq \mu, W_t^* \geq \alpha) &= P(2\alpha - W_t \leq \mu, W_t^* \geq \alpha) \\ &= P(W_t \geq 2\alpha - \mu, W_t^* \geq \alpha) \\ &= P(W_t \geq 2\alpha - \mu), \end{aligned}$$

since $2\alpha - \mu \geq \alpha$ and therefore $\{W_t \geq 2\alpha - \mu\} \subseteq \{W_t^* \geq \alpha\}$. Furthermore,

$$\begin{aligned} P(W_t \leq \mu, W_t^* \leq \alpha) &= P(W_t \leq \mu) - P(W_t \leq \mu, W_t^* \geq \alpha) \\ &= P(W_t \leq \mu) - P(W_t \geq 2\alpha - \mu). \end{aligned}$$

$\mu = \alpha$ gives (6.5).

(b) $f(x, y)$ is

$$\frac{\partial^2}{\partial\mu\partial\alpha} P(W_t \leq \mu, W_t^* \leq \alpha) \mid_{x=\mu,\, y=\alpha} .$$

(c) Note that the density of W_t is

$$f_{W_t}(x) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}.$$

Thus

$$f_{W_t^*|W_t=x}(y \mid x) = \frac{f(x, y)}{f_{W_t}(x)} = \frac{2(2y - x)}{t} e^{-2(y^2 - xy)/t}.$$

Therefore

$$P(W_t^* \geq \alpha \mid W_t = x) = \int_\alpha^\infty \frac{2(2y - x)}{t} e^{-2(y^2 - xy)/t}\, dy = e^{-2\alpha(\alpha - x)/t}.$$

□

To derive the boundary crossing probabilities for the Brownian motion and a linear boundary $b(t) = \alpha + \beta t$, note that

$$\{W_s \geq \alpha + \beta s \text{ for an } s \in [0, t]\} = \{W_s - \beta s \geq \alpha \text{ for an } s \in [0, t]\}.$$

Let $X_t = W_t - \beta t$ and $X_t^* = \sup_{0 \leq s \leq t} X_s$.

Remark 6.10 (X_t) *is a Brownian motion with drift. The conditional distribution of* X_s *given* $X_t = x$ *is, for* $0 \leq s \leq t$ *independent of* β, *see Exercise 6.6. A process* $(U_s)_{0 \leq s \leq t}$ *with the same distribution as* $(X_s)_{0 \leq s \leq t} \mid (X_t = x)$ *is called a Brownian bridge on* $[0, t]$.

Theorem 6.11 *Let $\alpha > 0$. Then*

$$P(X_t^* \geq \alpha, X_t \leq \lambda) = \begin{cases} e^{-2\alpha\beta}\Phi\left(\frac{\lambda-2\alpha+\beta t}{\sqrt{t}}\right), & \text{if } \lambda \leq \alpha, \\ \Phi\left(\frac{\lambda+\beta t}{\sqrt{t}}\right) - \Phi\left(\frac{\alpha+\beta t}{\sqrt{t}}\right) + e^{-2\alpha\beta}\Phi\left(\frac{\beta t-\alpha}{\sqrt{t}}\right), & \text{if } \lambda > \alpha, \end{cases} \tag{6.7}$$

and

$$P(X_t^* \geq \alpha) = 1 - \Phi\left(\frac{\alpha + \beta t}{\sqrt{t}}\right) + e^{-2\alpha\beta}\Phi\left(\frac{\beta t - \alpha}{\sqrt{t}}\right). \tag{6.8}$$

Proof. (6.7) implies (6.8). To prove (6.7), let $\lambda \leq \alpha$. Then

$$P(X_t^* \geq \alpha, X_t \leq \lambda) = \int_{-\infty}^{\lambda} e^{-2\alpha(\alpha-x)/t}\frac{1}{\sqrt{2\pi t}}e^{-(x+\beta t)^2/2t}dx.$$

Since

$$4\alpha(\alpha - x) + (x + \beta t)^2 = (x + \beta t - 2\alpha)^2 + 4\alpha\beta t,$$

we have

$$\begin{aligned} P(X_t^* \geq \alpha, X_t \leq \lambda) &= \int_{-\infty}^{\lambda} e^{-2\alpha\beta}\frac{1}{\sqrt{2\pi t}}e^{-(x+\beta t-2\alpha)^2/2t}dx \\ &= e^{-2\alpha\beta}\Phi\left(\frac{\lambda - 2\alpha + \beta t}{\sqrt{t}}\right). \end{aligned}$$

If $\lambda > \alpha$, we have

$$\begin{aligned} P(X_t^* \geq \alpha, X_t \leq \lambda) &= P(X_t^* \geq \alpha, X_t \leq \alpha) + P(X_t^* \geq \alpha, \alpha < X_t \leq \lambda) \\ &= e^{-2\alpha\beta}\Phi\left(\frac{\beta t - \alpha}{\sqrt{t}}\right) + P(\alpha < X_t \leq \lambda) \\ &= e^{-2\alpha\beta}\Phi\left(\frac{\beta t - \alpha}{\sqrt{t}}\right) + \Phi\left(\frac{\lambda + \beta t}{\sqrt{t}}\right) - \Phi\left(\frac{\alpha + \beta t}{\sqrt{t}}\right). \end{aligned}$$

□

Remark 6.12 *Let $\beta > 0$. Then*

$$P(W_t \geq \alpha + \beta t \textit{ for a } t \geq 0) = e^{-2\alpha\beta}. \tag{6.9}$$

Let $\beta \leq 0$. Then

$$P(W_t \geq \alpha + \beta t \textit{ for a } t \geq 0) = 1. \tag{6.10}$$

6.2.2 Constant Two-Sided Boundaries

We consider two constant boundaries only. Let $b(t) = \beta$, $a(t) = -\alpha$, $\alpha, \beta > 0$. Let

$$\tau^{\alpha,\beta} = \inf\{t \geq 0 \mid W_t \geq \beta \text{ or } W_t \leq -\alpha\}. \tag{6.11}$$

Proposition 6.13 *Let $s \in \mathbb{R}$. Then*

$$\mathbb{E}(e^{-s^2\tau^{\alpha,\beta}/2}) = \frac{e^{\alpha s} - e^{-\alpha s} - e^{\beta s} - e^{-\beta s}}{e^{(\alpha+\beta)s} - e^{-(\alpha+\beta)s}}. \tag{6.12}$$

Proof. Let

$$M_t = e^{sW_t - ts^2/2}.$$

(M_t) is a martingale. $\tau^{\alpha,\beta} \wedge n$ is a bounded stopping time. We may apply the optional sampling theorem to get

$$\begin{aligned} 1 &= \mathbb{E}(M_{\tau^{\alpha,\beta}\wedge n}) \\ &= e^{-s\alpha}\mathbb{E}\left(e^{-(\tau^{\alpha,\beta}\wedge n)s^2/2} I_{\{\tau^{\alpha,\beta}\wedge n \le n,\, W_{\tau^{\alpha,\beta}\wedge n} = -\alpha\}}\right) \\ &+ e^{s\beta}\mathbb{E}\left(e^{-(\tau^{\alpha,\beta}\wedge n)s^2/2} I_{\{\tau^{\alpha,\beta}\wedge n \le n,\, W_{\tau^{\alpha,\beta}\wedge n} = \beta\}}\right) \\ &+ \mathbb{E}\left(e^{sW_{\tau^{\alpha,\beta}\wedge n} - ns^2/2} I_{\{\tau^{\alpha,\beta}\wedge n > n\}}\right). \end{aligned}$$

Now let $n \to \infty$. The third term is bounded by

$$e^{|s|(\alpha\vee\beta) - ns^2/2}$$

and converges to 0. Define events $A = \{W_{\tau^{\alpha,\beta}\wedge n} = -\alpha\}$, $B = \{W_{\tau^{\alpha,\beta}\wedge n} = \beta\}$ and abbreviate

$$\mathbb{E}\left(e^{-\tau^{\alpha,\beta}s^2/2} I_A\right)$$

by x and

$$\mathbb{E}\left(e^{-\tau^{\alpha,\beta}s^2/2} I_B\right)$$

by y. Then we have

$$1 = e^{-s\alpha}x + e^{s\beta}y.$$

Replace s by $-s$ to get

$$1 = e^{s\alpha}x + e^{-s\beta}y$$

with solutions

$$\begin{aligned} x &= \frac{e^{\beta s} - e^{-\beta s}}{e^{(\alpha+\beta)s} - e^{-(\alpha+\beta)s}}, \\ y &= \frac{e^{\alpha s} - e^{-\alpha s}}{e^{(\alpha+\beta)s} - e^{-(\alpha+\beta)s}}. \end{aligned}$$

Since $\mathbb{E}(e^{-s^2\tau^{\alpha,\beta}/2}) = x + y$, (6.12) follows.

□

The Laplace-transform of $\tau^{\alpha,\beta}$, i.e. the mapping $v \mapsto \mathbb{E}(e^{-v\tau^{\alpha,\beta}})$ characterizes the distribution of $\tau^{\alpha,\beta}$ uniquely. It may be inverted to derive a series expansion of the c.d.f. of $\tau^{\alpha,\beta}$. We sketch the derivation of the c.d.f. for the symmetric case $\alpha = \beta$ only by applying again a reflection principle.

Let $\alpha = \beta$. We abbreviate $\tau^{\alpha,\beta}$ by τ^{β}. τ^{β} has Laplace-transform

$$\mathbb{E}(e^{-v\tau^{\alpha}}) = \frac{2}{e^{\beta\sqrt{2v}} + e^{-\beta\sqrt{2v}}} = \frac{1}{\cosh(\beta\sqrt{2v})}.$$

Let $|W|_t^* = \sup_{0\le s\le t} |W_s|$. We define events $A_+ = \{W_s = \beta \text{ for an } s \in [0,t]\}$, $A_- = \{W_s = -\beta \text{ for an } s \in [0,t]\}$. Then

$$P(|W|_t^* \ge \beta) = P(A_+ \cup A_-) = P(A_+) + P(A_-) - P(A_+ \cap A_-).$$

A path in $A_+ \cap A_-$ crosses β and $-\beta$. It is either in $A_{+-} = \{$ paths that cross β before $-\beta\}$ or in $A_{-+} = \{$ paths that cross $-\beta$ before $\beta\}$.

Paths in A_{+-} are reflected at β and correspond to (certain) paths crossing 3β. The event $\{W_t^* \ge 3\beta\}$ contains additionally to the reflected paths from A_{+-} those from A_{-+-} reflected at β. Similarly, paths in A_{-+} are reflected at $-\beta$ and correspond to paths crossing -3β. The procedure is iterated by considering A_{+-+} and A_{-+-} and reflecting its reflected paths a second time at 3β and -3β respectively. We get

$$\begin{aligned} P(A_+ \cup A_-) &= P(A_+) - P(A_{+-}) + P(A_{+-+}) - P(A_{+-+-}) + \cdots \\ &+ P(A_-) - P(A_{-+}) + P(A_{-+-}) - P(A_{-+-+}) + \cdots \\ &= 2\left(P(A_+) - P(A_{+-}) + P(A_{+-+}) - P(A_{+-+-}) + \cdots\right) \\ &= 2\left(P(W_t^* \ge \beta) - P(W_t^* \ge 3\beta) + P(W_t^* \ge 5\beta) - \cdots\right) \\ &= 2\left(2\bar{\Phi}(\beta/\sqrt{t}) - 2\bar{\Phi}(3\beta/\sqrt{t}) + 2\bar{\Phi}(5\beta/\sqrt{t}) - \cdots\right). \end{aligned}$$

This proves

Proposition 6.14 *Let $\beta > 0$. Then*

$$P(|W|_t^* \ge \beta) = 4\sum_{k=1}^{\infty} (-1)^{k-1} \bar{\Phi}((2k-1)\beta/\sqrt{t}). \tag{6.13}$$

6.3 Exercises

Exercise 6.1 *Prove that if $Z \sim N(0,h)$, then $\mathbb{E}(Z^2) = h$ and $\sigma(Z^2) = 2h^2$.*

Exercise 6.2 *Let (W_t) be a Brownian motion. Show that the following processes are again Brownian motions:*

1. *Let $c \ne 0$ and $\tilde{W}_t = cW_{t/c^2}$ (transformation of time).*
2. *Let $h > 0$ and $(\tilde{W}_t) = (W_{t+h} - W_h)$.*
3. *Let $\tau > 0$ be fixed and $\tilde{W}_t = W_t$ if $t \le \tau$ and $\tilde{W}_t = 2W_\tau - W_t$ if $t > \tau$. $(\tilde{W}_t)$ is the Brownian motion reflected in $w = W_\tau$.*

4. $\tilde{W}_t = tW_{1/t}$ *for* $t > 0$ *and* $\tilde{W}_0 = 0$. *(To check the continuity of* $(\tilde{W}_t)$ *in* 0, *show only that for* $\epsilon > 0$, $P(|\tilde{W}_t| > \epsilon) \to 0$ *for* $t \to 0+$.*)*

Exercise 6.3 *A geometric Brownian motion* (S_t) *is a submartingale, if* $\mu \geq -\sigma^2/2$, *a supermartingale if* $\mu \leq -\sigma^2/2$. *It is a martingale if and only if* $\mu = -\sigma^2/2$.

Exercise 6.4 *Let a geometric Brownian motion* (S_t) *be given. Prove that* $(V_t) = (1/S_t)$ *is again a geometric Brownian motion. Furthermore, prove that if* (S_t) *is a supermartingale, then* (V_t) *is a submartingale. Is it possible that both* (S_t) *and* (V_t) *are submartingales or both supermartingales?*

Exercise 6.5 *Let* $X_t = x_0 + \mu t + \sigma W_t$ *be a (generalized) Brownian motion. Prove that* (X_t) *is a martingale if and only if* $\mu = 0$.

Exercise 6.6 *Let* (W_t) *be a Brownian motion and* $X_t = W_t + \mu t$. *Derive the distribution of* $(X_t)_{0 \leq t \leq u}$ *conditional on* $W_u = x$.

Exercise 6.7 *Let* (W_t) *be a Brownian motion and* $U_t = W_t - tW_1$ *(for* $0 \leq t \leq 1$*). Prove that* (U_t) *is a Brownian bridge.*

Exercise 6.8 *Let* $(U_t)_{0 \leq t \leq 1}$ *be a Brownian bridge and* $W_t = (1+t)U_{t/(1+t)}$. *Prove that* $(W_t)_{0 \leq t < \infty}$ *is a Brownian motion.*

Exercise 6.9 *Let* (W_t) *and* (V_t) *be independent Brownian motions. Let* $-1 \leq r \leq 1$. *Show that* $(rW_t + \sqrt{1-r^2}V_t)$ *is a Brownian motion.*

Exercise 6.10 *Let* (U_t) *and* (V_t) *be independent Brownian bridges. Let* $-1 \leq r \leq 1$. *Show that* $(\sqrt{1-r^2}U_t - rV_t)$ *is a Brownian bridge.*

Exercise 6.11 *Let* $(U_t)_{0 \leq t \leq 1}$ *be a Brownian bridge. Prove that* $(U_{t/2} - U_{1-t/2})$ *is again a Brownian bridge.*

Exercise 6.12 *Show that* τ^α, *the first hitting time to the constant boundary, has the density (w.r.t. Lebesgue measure)*

$$f(t) = \frac{\alpha}{\sqrt{2\pi t^3}} e^{-\alpha^2/2t}$$

for $t > 0$ *and* $f(t) = 0$ *for* $t \leq 0$. *Derive that* $1/\tau^\alpha$ *is gamma-distributed with shape parameter* $1/2$ *and rate parameter* $\alpha^2/2$ *and that* $\mathbb{E}(\tau^\alpha) = \infty$.

Exercise 6.13 *Let* (S_t) *be a geometric Brownian motion,*

$$S_t = S_0 e^{(\mu - \sigma^2/2)t + \sigma W_t},$$

with μ *constant and* $\sigma > 0$, $S_0 > 0$ *constant.*

1. *Let* $\nu > 0$. *Compute*

$$P(S_u \geq \nu \textit{ for a } u \leq t).$$

2. *Let* $\mu < \sigma^2/2$. *Compute*

$$P(S_u \textit{ ever crosses } \nu).$$

Chapter 7

Poisson Process

7.1 Definition and Properties

Let $(\tau_i)_{i=1}^{\infty}$ be a sequence of independent exponentially (i.e. $\Gamma(1,\lambda)$) distributed random variables. Let $T_n = \sum_{i=1}^{n} \tau_i$ and

$$N_t = \sum_{n=1}^{\infty} I_{\{T_n \le t\}}. \tag{7.1}$$

$(N_t)_{t\in[0,\infty)}$ is called the *Poisson process* with rate (intensity) λ.

(T_n) are the times, when an "event" occurs, N_t the number of "events" up to t. (T_n) are also called arrival times, $\tau_i = T_i - T_{i-1}$ the interarrival or waiting times.

Lemma 7.1 *(Absence of memory). If $T \sim \Gamma(1,\lambda)$, then*

$$P(T > t+s \mid T > t) = P(T > s).$$

Proof. We have $P(T > t) = e^{-\lambda t}$. Therefore

$$P(T > t+s \mid T > t) = \frac{P(T > t+s)}{P(T > t)} = e^{-\lambda s}.$$

□

Lemma 7.2 *Let $X_1, \ldots, X_n$ be independent and exponentially distributed with rate λ. Then $X_1 + \cdots + X_n \sim \Gamma(n,\lambda)$.*

Proof. Let $X \sim \Gamma(n,\lambda)$ and $Y \sim \Gamma(1,\lambda)$ be independent. We prove that $X+Y \sim \Gamma(n+1,\lambda)$.
Version I: The characteristic function $\varphi_Z(s)$ of a $\Gamma(n,\lambda)$-distributed random variable Z is

$$(1 - is/\lambda)^{-n}.$$

Therefore $\varphi_X(s)\varphi_Y(s) = (1-is/\lambda)^{-(n+1)}$ and $X+Y \sim \Gamma(n+1,\lambda)$.

Version II: Denote by f_X, f_Y and f_Z the densities of X, Y and $Z = X + Y$. f_Z is the convolution of f_X and f_Y. For $z > 0$ we have

$$\begin{aligned}
f_Z(z) &= \int f_X(z-y) f_Y(y) dy \\
&= \int_0^\infty I_{[0,\infty)}(z-y) \frac{\lambda^n}{\Gamma(n)} (z-y)^{n-1} e^{-\lambda(z-y)} \lambda e^{-\lambda y} dy \\
&= \int_0^z \frac{\lambda^{n+1}}{\Gamma(n)} (z-y)^{n-1} e^{-\lambda z} dy \\
&= \frac{\lambda^{n+1}}{\Gamma(n)} e^{-\lambda z} \int_0^z (z-y)^{n-1} dy \\
&= \frac{\lambda^{n+1}}{\Gamma(n)} e^{-\lambda z} \frac{z^n}{n} \\
&= \frac{\lambda^{n+1}}{\Gamma(n+1)} z^n e^{-\lambda z}.
\end{aligned}$$

Therefore $Z = X + Y \sim \Gamma(n+1, \lambda)$ □

Lemma 7.3 *Let (N_t) be a Poisson process with rate λ. Then $N_t \sim P(\lambda t)$.*

Proof. Let g_n denote the density of T_n. Then

$$\begin{aligned}
P(N_t = n) &= P(T_n \leq t, T_{n+1} > t) \\
&= \int_0^t P(T_{n+1} > t \mid T_n = s) g_n(s) ds \\
&= \int_0^t e^{-\lambda(t-s)} \frac{\lambda^n s^{n-1}}{\Gamma(n)} e^{-\lambda s} ds \\
&= e^{-\lambda t} \frac{\lambda^n}{(n-1)!} \int_0^t s^{n-1} ds \\
&= e^{-\lambda t} \frac{\lambda^n}{(n-1)!} \frac{t^n}{n} \\
&= e^{-\lambda t} \frac{(\lambda t)^n}{n!}.
\end{aligned}$$

□

Lemma 7.4 *Let $t > 0$ and $U_1, \ldots, U_n$ be independent and $\mathbb{U}[0,t]$-distributed. Let $Y_1, \ldots, Y_n$ denote the order statistics of $U_1, \ldots, U_n$, i.e. $Y_1 = U_{(1)} \leq Y_2 = U_{(2)} \leq \cdots \leq Y_n = U_{(n)}$. Then $(Y_1, \ldots, Y_n)$ has the density*

$$f(y_1, \ldots, y_n) = \begin{cases} n!/t^n & \text{if } (y_1, \ldots, y_n) \in \mathcal{S}, \\ 0 & \text{else,} \end{cases}$$

where $\mathcal{S} = \{(y_1, \ldots, y_n) \mid 0 < y_1 < \cdots < y_n < t\}$.

Proof. Note that for all $i = 1, \ldots, n-1$, $P(Y_i = Y_{i+1}) = 0$. $(U_1, \ldots, U_n)$ has a density, that is constant $(= 1/t^n)$ on $[0,t]^n$ and 0 on the complement of $[0,t]^n$. Furthermore, $(Y_1, \ldots, Y_n)$ is a function of $(U_1, \ldots, U_n)$. If $(u_1, \ldots, u_n) \in [0,t]^n$, with all u_i's different, there exists a unique permutation π of $\{1, \ldots, n\}$ s.t. $(u_1, \ldots, u_n) = (y_{\pi(1)}, \ldots, y_{(n)})$ with $(y_1, \ldots, y_n) \in \mathcal{S}$. Therefore, for all bounded and measurable functions g,

$$\begin{aligned}
\mathbb{E}(g(Y_1, \ldots, Y_n)) &= \int_0^t \cdots \int_0^t g(y_1, \ldots, y_n) \frac{1}{t^n} du_1 \cdots du_n \\
&= \sum_\pi \int_{\mathcal{S}} g(y_1, \ldots, y_n) \frac{1}{t^n} dy_1 \cdots dy_n \\
&= \int_{\mathcal{S}} g(y_1, \ldots, y_n) \frac{n!}{t^n} dy_1 \cdots dy_n.
\end{aligned}$$

□

Remark 7.5 *The probability distribution of the order statistics of i.i.d.* $\mathbb{U}[0,t]$*-distributed random variables is called* Dirichlet *distribution* $D_{n,t}$.

Proposition 7.6 *Let* (T_i) *be the arrival times of a Poisson process* (N_t) *with rate* λ*. Then, conditionally on* $T_{n+1} = t$*,*

$$(T_1, \ldots, T_n) \sim D_{n,t}.$$

Remark 7.7 *(a)* $(T_1/T_{n+1}, \ldots, T_n/T_{n+1})$ *is independent of* T_{n+1} *and* $D_{n,1}$*-distributed.*

(b) Let $U_1, \ldots, U_n, V$ *be independent with* $U_i \sim \mathbb{U}[0,1]$ *and* $V \sim \Gamma(n+1, \lambda)$*. Then*

$$(T_1, \ldots, T_n) \sim (VU_{(1)}, \ldots, VU_{(n)}).$$

Proof of Proposition 7.6. Let $\tau_1, \ldots, \tau_{n+1}$ be the exponentially distributed waiting times and $T_k = \tau_1 + \cdots + \tau_k$ the arriving times. Consider the mapping $T_k : (\tau_1, \ldots, \tau_n) \mapsto \tau_1 + \cdots + \tau_k$. Then

$$\frac{\partial T_k}{\partial \tau_i} = \begin{cases} 1 & \text{if } i \leq k \\ 0 & \text{if } i > k. \end{cases}$$

We have

$$\begin{vmatrix} 1 & 0 & & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & 0 & . \\ . & & & & . \\ 1 & . & . & . & 1 \end{vmatrix} = 1$$

and therefore the density of $(T_1, \ldots, T_n) \mid (T_{n+1} = t)$ is

$$\begin{aligned}
f(t_1, \ldots, t_n \mid t) &= \frac{\lambda e^{-\lambda t_1} \lambda e^{-\lambda(t_2 - t_1)} \cdots \lambda e^{-\lambda(t - t_n)}}{\lambda^{n+1} t^n e^{-\lambda t} / \Gamma(n+1)} \\
&= \frac{n!}{t^n} \frac{e^{-\lambda(t_1 + (t_2 - t_1) + \cdots + (t - t_n))}}{e^{-\lambda t}} \\
&= \frac{n!}{t^n}
\end{aligned}$$

on $0 < t_1 < \cdots < t_n < t$. $\square$

Corollary 7.8 *Conditional on $N_t = n$,*

$$(T_1, \ldots, T_n) \sim D_{n,t}.$$

Proof. The density of $(T_1, \ldots, T_n)$ is

$$f(t_1, \ldots, t_n) = e^{-\lambda t_n} \frac{\lambda^n t_n^{n-1}}{\Gamma(n)} \frac{(n-1)!}{t_n^{n-1}} = e^{-\lambda t_n} \lambda^n$$

on $0 < t_1 < \cdots < t_n$. Let $A \subseteq [0,t]^n$ be measurable. We have

$$\begin{aligned} P((T_1, \ldots, T_n) \in A \mid N_t = n) &= \frac{P((T_1, \ldots, T_n) \in A, T_{n+1} > t)}{P(N_t = n)} \\ &= \frac{\int_A P(\tau_{n+1} > t - t_n) f(t_1, \ldots, t_n) dt_1 \cdots dt_n}{P(N_t = n)} \\ &= \frac{\int_A e^{-\lambda(t-t_n)} e^{-\lambda t_n} \lambda^n I_{\{0<t_1<\cdots<t_n\}} dt_1 \cdots dt_n}{e^{-\lambda t} (\lambda t)^n / n!} \\ &= \int_A \frac{n!}{t^n} I_{\{0<t_1<\cdots<t_n\}} dt_1 \cdots dt_n. \end{aligned}$$

$\square$

Theorem 7.9 *Let (N_t) be a Poisson process with rate λ. Then, for $0 \le t_1 < \cdots < t_n$, $N_{t_1}, N_{t_2} - N_{t_1}, N_{t_3} - N_{t_2}, \ldots, N_{t_n} - N_{t_{n-1}}$ are independent and $N_{t_i} - N_{t_{i-1}} \sim P(\lambda(t_i - t_{i-1}))$.*

Proof. Note that $N_{t_1} = k_1, N_{t_2} - N_{t_1} = k_2, \ldots, N_{t_n} - N_{t_{n-1}} = k_n$ if and only if $T_{k_1} \le t_1, T_{k_1+1} > t_1, T_{k_1+k_2} \le t_2, T_{k_1+k_2+1} > t_2, \ldots, T_{k_1+\cdots+k_n} \le t_n, T_{k_1+\cdots+k_n+1} > t_n$.

Let $N = k_1 + \cdots + k_n$. We condition on $T_{N+1} = s$ with $s > t_n$. We draw N independent $\mathbb{U}[0,s]$-distributed random variables $U_1, \ldots U_N$ and compute the probability of the event that exactly k_1 if the U_i's are in $[0, t_1]$, k_2 in $(t_1, t_2], \ldots, k_n$ in $(t_{n-1}, t_n]$. This probability is

$$\frac{N!}{k_1! k_2! \cdots k_n!} \left(\frac{t_1}{s}\right)^{k_1} \left(\frac{t_2 - t_1}{s}\right)^{k_2} \cdots \left(\frac{t_n - t_{n-1}}{s}\right)^{k_n}.$$

Therefore,

$$\begin{aligned} & P(N_{t_1} = k_2, N_{t_2} - N_{t_1} = k_2, \ldots, N_{t_n} - N_{t_{n-1}} = k_n) \\ = & \int_{t_n}^{\infty} \frac{N!}{k_1! k_2! \cdots k_n!} \left(\frac{t_1}{s}\right)^{k_1} \left(\frac{t_2 - t_1}{s}\right)^{k_2} \cdots \left(\frac{t_n - t_{n-1}}{s}\right)^{k_n} \frac{\lambda^{N+1} s^N}{\Gamma(N+1)} e^{-\lambda s}\, ds \\ = & \frac{t_1^{k_1} (t_2 - t_1)^{k_2} \cdots (t_n - t_{n-1})^{t_n}}{k_1! k_2! \cdots k_n!} \lambda^N \int_{t_n}^{\infty} \lambda e^{-\lambda s}\, ds \\ = & \frac{t_1^{k_1} (t_2 - t_1)^{k_2} \cdots (t_n - t_{n-1})^{t_n}}{k_1! k_2! \cdots k_n!} \lambda^N e^{-\lambda t_n} \\ = & \frac{(\lambda t_1)^{k_1} e^{-\lambda t_1}}{k_1!} \frac{(\lambda(t_2 - t_2))^{k_1} e^{-\lambda(t_2 - t_1)}}{k_2!} \cdots \frac{(\lambda(t_n - t_{n-1}))^{k_n} e^{-\lambda(t_n - t_{n-1})}}{k_n!}. \square \end{aligned}$$

Summary: Let (N_t) be a Poisson process with rate λ.

1. The sample paths $t \mapsto N_t(\omega)$ are piecewise constant with jumps of size 1.

2. The sample paths $t \mapsto N_t(\omega)$ are continuous from the right with limits from the left (cadlag).

3. N_t is Poisson distributed with parameter λt. Increments $N_t - N_u$ $(u < t)$ are Poisson distributed with parameter $\lambda(t-u)$ and if $u_1 < t_1 \leq u_2 < t_2$, then $N_{t_1} - N_{u_1}$ and $N_{t_2} - N_{u_2}$ are independent.

4. The increments are homogeneous, i.e. the distribution of $N_t - N_u$ depends on $t-u$ only and is the same as the distribution of N_{t-u}.

5. (N_t) has the Markov property, for $u < t$ and all integrable f,

$$\mathbb{E}(f(N_t) \mid \sigma(N_v, v \leq u)) = \mathbb{E}(f(N_t) \mid \sigma(N_u)).$$

Summary A Poisson process (N_t) is a

- *Counting process*: Let (random) times T_n with $T_1 \leq T_2 \leq \cdots$ and $T_n \to \infty$ with probability 1. Let $X_t = n$ if $T_n \leq t < T_{n+1}$. (X_t) "counts" the number of "events" up to time t.

- *Renewal process*: Let (τ_i) be i.i.d., $\tau_i \geq 0$, $T_n = \tau_1 + \cdots + \tau_n$. A renewal process is a counting process with independent and identically distributed waiting times.

- *Point process*: (N_t) "produces" a value of 1 at T_n. More generally, a point process consists of $(T_n, X_n)_{n=1}^{\infty}$, where (T_n) are times and (X_n) are values. If (N_t) is the counting process of (T_n), $Y_t = \sum_{n=1}^{N_t} X_n$ is a point process. In case (X_n) are i.i.d. and independent of (N_t) and (N_t) is a Poisson process, (Y_t) is a compound Poisson process.

The Poisson process may be generalized in different ways. One is the *Compound Poisson process*. Let (N_t) denote independent Poisson processes with intensity λ, Q a probability distribution on $(\mathbb{R}, \mathcal{B})$, $(X_n)_{n=1}^{\infty}$ i.i.d. with distribution Q, also independent of (N_t). Let $Y_t = \sum_{n=1}^{N_t} X_n$, i.e.

$$Y_t = \begin{cases} \sum_{n=1}^{k} X_n & \text{on } N_t = k, k > 0, \\ 0 & \text{on } N_t = 0. \end{cases}$$

(Y_t) is a compound Poisson process with intensity λ and jump distribution Q.

Remark 7.10 *Let (Y_t) denote a compound Poisson process with intensity λ and jump distribution Q.*

1. (Y_t) has piecewise constant cadlag paths.

2. *Assume $Q(\{0\}) = 0$. The jump times (T_n) have the same distribution as the jump times of the Poisson distribution (N_t).*

3. *The jump sizes (X_n) are i.i.d.*

4. *(Y_t) has stationary and independent increments.*

7.2 Exercises

Exercise 7.1 *Let (N_t) and (N_t') denote independent Poisson processes with intensities λ and λ' resp. Show that $(N_t + N_t')$ is a Poisson processes with intensity $\lambda + \lambda'$.*

Exercise 7.2 *(Thinning). Let (N_t) denote a Poisson processes with intensity λ. Let (T_n) denote its arrival times. For each n, keep (delete) T_n with probability p $(1-p)$, independently for different n's. Denote by (T_n') the arrival times that have not been deleted. Define*

$$N_t' = \sum_{n=1}^{\infty} I_{\{T_n' \le t\}}.$$

Prove that (N_t') is a Poisson process with intensity λp.

Exercise 7.3 *(Compensator). Let (N_t) denote a Poisson processes with intensity λ. Show that $(\tilde{N}_t)$, defined by $\tilde{N}_t = N_t - \lambda t$, is a martingale.*

Exercise 7.4 *(Continued). Let (N_t) denote a Poisson processes with intensity λ. Compute the quadratic variation of $(\tilde{N}_t)$, where $\tilde{N}_t = N_t - \lambda t$.*

Exercise 7.5 *Let (N_t) denote a Poisson processes with intensity λ. Show that $P(N_t < \infty) = 1$.*

Hint: Apply Chebyshev's inequality to prove that $\lim_{n\to\infty} P(T_n \le t) = 0$.

Exercise 7.6 *Let (Y_t) denote a compound Poisson process with intensity λ and jump distribution Q. Let μ_X and σ_X^2 denote the expectation and the variance of $X \sim Q$. Compute the expectation and variance of (Y_t).*

Solution: $\mathbb{E}(Y_t) = \lambda t \mu_X$, $\sigma_{Y_t}^2 = \lambda t \sigma_X^2 + \lambda t \mu_X^2$.

Exercise 7.7 *Let (Y_t) denote a compound Poisson process with intensity λ and jump distribution Q. Let μ_X denote the expectation of $X \sim Q$. Show that (Y_t) is a martingale, if $\mu_X = 0$. Compute the compensator of (Y_t) in case $\mu_X \neq 0$.*

Exercise 7.8 *Let (Y_t) denote a compound Poisson process with intensity λ and jump distribution Q. Compute the characteristic function of Y_t.*

Solution: $\varphi_{Y_t}(s) = e^{-\lambda t(1-\varphi_X(s))}$.

Chapter 8

Appendix: Conditional Expectation

As a motivation for the concept of the conditional expectation, consider the following problem of predicting a random variable Y: Let a probability space $(\Omega, \mathcal{F}, P)$ and a square integrable random variable $Y : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$ be given.

To predict Y without further information, a real number $c = \hat{Y}$ has to chosen, s.t. $\mathbb{E}((Y-c)^2)$ is a small as possible. We know that the solution is the expectation $c = \mathbb{E}(Y)$. With this choice $\mathbb{E}((Y-c)^2)$ is then the variance of Y.

Now assume that the prediction of Y may be based on information provided by a random variable X, i.e. one has to choose a function g and predict Y by $\hat{Y} = g(X)$. If (X, Y) has a joint distribution and if a conditional distribution of Y given $X = x$ can be defined, then $g(x)$ is the expectation of this conditional distribution and $\hat{Y} = g(X)$. Note that $g(X)$ is a random variable, since X is random.

Theorem 8.1 *(Causality Theorem). Let X be an $\mathbb{R}^n$-valued random variable on the measurable space $(\Omega, \mathcal{F})$. Let Y be an $\mathbb{R}$ valued random variable. Y is $\sigma(X)$-measurable if and only if an measurable $g : \mathbb{R}^n \to \mathbb{R}$ exists such that $Y = g(X)$.*

Proof. Only one direction has to be shown. Let us recall that $\sigma(X) = \{X^{-1}(B) \mid B \in \mathcal{B}^n\}$. As usual, we first prove the theorem for the special case that Y is simple, i.e. $Y = \sum_{i=1}^k c_i I_{A_i}$ with $A_i \in \sigma(X)$ and $(A_1, \ldots, A_k)$ a partition. Therefore, $Y = \sum_{i=1}^k c_i I_{X^{-1}(B_i)}$ with $B_i \in \mathcal{B}^n$ and $(B_1, \ldots, B_k)$ a partition. Then, if we define g by $g(x) = c_i$ if $x \in B_i$, we have $Y = g(X)$.

In the general case, $Y = \lim_{m\to\infty} Y_m$ with Y_m simple and $\sigma(X)$-measurable. Therefore, $Y_m = g_m(X)$. If we define $g(x) = \limsup_{m\to\infty} g_m(x)$, then g is measurable and $Y = g(X)$. □

Given the *Causality Theorem* the problem of predicting Y may be generalized. Let a probability space $(\Omega, \mathcal{F}, P)$, a sub σ-algebra $\mathcal{G} \subseteq \mathcal{F}$ and a square integrable random variable Y be given. Find the $\mathcal{G}$-measurable random variable $\hat{Y}$ that minimizes $\mathbb{E}((Y - \hat{Y})^2)$. Note that $L^2(\mathcal{G})$, the set of $\mathcal{G}$-measurable random variables, is a subspace of $L^2(\mathcal{F})$ and $\hat{Y}$ is the projection of Y onto this subspace. Projections have (in this case) the property that $Y - \hat{Y}$ is orthogonal to $L^2(\mathcal{G})$, i.e. for all $Z \in L^2(\mathcal{G})$, $\mathbb{E}(Z(Y - \hat{Y})) = 0$, which may be written as

$$\mathbb{E}(Z\hat{Y}) = \mathbb{E}(ZY).$$

The random variable $\hat{Y}$ is called the conditional expectation of Y given $\mathcal{G}$ and denoted by $\mathbb{E}(Y \mid \mathcal{G})$. It is uniquely defined by two properteis: $\mathbb{E}(Y \mid \mathcal{G})$ is $\mathcal{G}$-measurable and $\mathbb{E}(Z\hat{Y}) = \mathbb{E}(ZY)$ holds for all $Z \in L^2(\mathcal{G})$.

Definition 8.2 *Let a probability space $(\Omega, \mathcal{F}, P)$, a sub σ-algebra $\mathcal{G} \subseteq \mathcal{F}$ and a random variable Y be given. A random variable $\mathbb{E}(Y \mid \mathcal{G})$ is called the conditional expectation of Y given $\mathcal{G}$ if it satisfies*

$$\begin{aligned} \mathbb{E}(Y \mid \mathcal{G}) & \quad \textit{is } \mathcal{G}\textit{-measurable,} && (8.1) \\ \mathbb{E}(Z\mathbb{E}(Y \mid \mathcal{G})) &= \mathbb{E}(ZY) \textit{ for all bounded and } \mathcal{G}\textit{-measurable } Z. && (8.2) \end{aligned}$$

Theorem 8.3 *Let a probability space $(\Omega, \mathcal{F}, P)$, a sub σ-algebra $\mathcal{G} \subseteq \mathcal{F}$ and a random variable Y be given. If Y is integrable, then $\mathbb{E}(Y \mid \mathcal{G})$ exists and is a.s. unique in the sense that if any other r.v. $\hat{Y}$ satisfies* (8.1) *and* (8.2)*, then $\hat{Y} = \mathbb{E}(Y \mid \mathcal{G})$ a.s.*

Theorem 8.4 *(Properties). Let a probability space $(\Omega, \mathcal{F}, P)$, a sub σ-algebra $\mathcal{G} \subseteq \mathcal{F}$ and an integrable random variables Y, Y_1, Y_2 be given.*

1. *$\mathbb{E}(Y \mid \mathcal{G}) = Y$ if and only if Y is $\mathcal{G}$-measurable.*
2. *$\mathbb{E}(Y \mid \mathcal{G}) = \mathbb{E}(Y)$ if Y is independent of $\mathcal{G}$.*
3. *$\mathbb{E}(\alpha Y_1 + \beta Y_2 \mid \mathcal{G}) = \alpha\mathbb{E}(Y_1 \mid \mathcal{G}) + \beta\mathbb{E}(Y_2 \mid \mathcal{G})$. $\alpha, \beta \in \mathbb{R}$.*
4. *If $Y \geq 0$ then $\mathbb{E}(Y \mid \mathcal{G}) \geq 0$.*
5. *$\mathbb{E}(\mathbb{E}(Y \mid \mathcal{G})) = \mathbb{E}(Y)$.*
6. *If $\mathcal{H} \subseteq \mathcal{G}$ is a sub σ-algebra, then $\mathbb{E}(\mathbb{E}(Y \mid \mathcal{G}) \mid \mathcal{H}) = \mathbb{E}(Y \mid \mathcal{H})$ (Tower property).*
7. *If Z is bounded and $\mathcal{G}$-measurable, then $\mathbb{E}(ZY \mid \mathcal{G}) = Z\mathbb{E}(Y \mid \mathcal{G})$.*
8. *If $\mathcal{G} = \{\emptyset, \Omega\}$, then $\mathbb{E}(Y \mid \mathcal{G}) = \mathbb{E}(Y)$.*
9. *If $Y = g(X, Z)$, Z independent of X, then $\mathbb{E}(Y \mid \sigma(X)) = h(X)$, with $h(x) = \mathbb{E}(g(x, Z))$.*
10. *If f is measurable and convex, then $\mathbb{E}(f(Y) \mid \mathcal{G}) \geq f(\mathbb{E}(Y \mid \mathcal{G}))$.*
11. *Let Y be square integrable. Then* (8.2) *holds for all square integrable and $\mathcal{G}$-measurable Z.*

12. Let $\hat{Y}$ be $\mathcal{G}$-measurable. If $\mathbb{E}(YI_A) = \mathbb{E}(\hat{Y}I_A)$ for all $A \in \mathcal{G}$, then $\hat{Y} = \mathbb{E}(Y \mid \mathcal{G})$.

Proof.

1. Trivial.

2. We have to show that the constant function $\mathbb{E}(Y)$ is the conditional expectation of Y given $\mathcal{G}$. Constant functions are measurable w.r.t. all σ-algebras. To check (8.2), let Z be bounded and $\mathcal{G}$-measurable. Then Z and Y are independent and therefore

$$\mathbb{E}(ZY) = \mathbb{E}(Z)\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y)Z).$$

3. We have to show that $\alpha\mathbb{E}(Y_1 \mid \mathcal{G}) + \beta\mathbb{E}(Y_2 \mid \mathcal{G})$ is the conditional expectation of $\alpha Y_1 + \beta Y_2$ given $\mathcal{G}$. It is obviously $\mathcal{G}$-measurable. To check (8.2), let Z be bounded and $\mathcal{G}$-measurable. Then

$$\mathbb{E}(Z(\alpha Y_1 + \beta Y_2)) = \alpha\mathbb{E}(ZY_1) + \beta\mathbb{E}(ZY_2)$$
$$= \alpha\mathbb{E}(Z\mathbb{E}(Y_1 \mid \mathcal{G})) + \beta\mathbb{E}(Z\mathbb{E}(Y_2 \mid \mathcal{G})) = \mathbb{E}(Z(\alpha\mathbb{E}(Y_1 \mid \mathcal{G}) + \beta\mathbb{E}(Y_2 \mid \mathcal{G}))).$$

4. Let A denote the event $\{\mathbb{E}(Y \mid \mathcal{G}) < 0\}$. I_A is bounded and $\mathcal{G}$-measurable. $P(A) > 0$ is not possible, since otherwise

$$0 > \mathbb{E}(I_A\mathbb{E}(Y \mid \mathcal{G})) = \mathbb{E}(I_AY) \geq 0.$$

5. Let, in (8.2), $Z = 1$.

6. We have to prove that $\mathbb{E}(\mathbb{E}(Y \mid \mathcal{G}) \mid \mathcal{H})$ is the conditional expectation of Y given $\mathcal{H}$. It is obviously $\mathcal{H}$-measurable. Let Z be bounded and $\mathcal{H}$-measurable. Since $\mathcal{H} \subseteq \mathcal{G}$, it is also $\mathcal{G}$-measurable. Therefore

$$\mathbb{E}(Z\mathbb{E}(\mathbb{E}(Y \mid \mathcal{G}) \mid \mathcal{H})) = \mathbb{E}(Z\mathbb{E}(Y \mid \mathcal{G})) = \mathbb{E}(ZY).$$

7. We have to prove that $Z\mathbb{E}(Y \mid \mathcal{G})$ is the conditional expectation of ZY given $\mathcal{G}$. It is obviously $\mathcal{G}$-measurable. Let U be bounded and $\mathcal{G}$-measurable. Then UZ is bounded and $\mathcal{G}$-measurable. Therefore

$$\mathbb{E}(U(Z\mathbb{E}(Y \mid \mathcal{G}))) = \mathbb{E}(UZ\mathbb{E}(Y \mid \mathcal{G})) = \mathbb{E}(UZY).$$

8. If $\mathcal{G} = \{\emptyset, \Omega)$, then only the constant functions are measurable. From 5, it follows that this constant is $\mathbb{E}(Y)$.

9. Let $h(x) = \mathbb{E}(g(x, Z))$. The Theorem of Tonelli-Fubini implies that h is measurable (w.r.t. $\sigma(X)$) and $h(X)$ is integrable. Let Z be bounded and measurable w.r.t. $\sigma(X)$. The Causality

Theorem implies that $Z = u(X)$ for a bounded and measurable function u. To show that $h(X)$ is the conditional expectation of Y given $\sigma(X)$, note that

$$\mathbb{E}(ZY) = \mathbb{E}(u(X)g(X,Z)) \quad \text{and} \quad \mathbb{E}(Zh(X)) = \mathbb{E}(u(X)h(X)),$$

and the two expectations are the same, again by the Theorem of Tonelli-Fubini.

10. No proof.

11. No proof.

12. No proof.

□

$\mathbb{E}(Y \mid \sigma(X))$ is abbreviated by $\mathbb{E}(Y \mid X)$.

Example 8.5 *Let the σ-algebra $\mathcal{G}$ be generated by the partition $(B_1, \ldots, B_n)$. Any $\mathcal{G}$-measurable function is a linear combination of the indicator functions $I_{B_1}, \ldots, I_{B_n}$. Therefore $\mathbb{E}(Y \mid \mathcal{G}) = \sum_{i=1}^n c_i I_{B_i}$. To identify the numbers c_k, let $Z = I_{B_k}$. Z is bounded and $\mathcal{G}$-measurable. From* (8.2) *we get*

$$\mathbb{E}(ZY) = \mathbb{E}(Z\mathbb{E}(Y \mid \mathcal{G})),$$

i.e.

$$\mathbb{E}(I_{B_k}Y) = \mathbb{E}(I_{B_k}\sum_{i=1}^n c_i I_{B_i}) = \sum_{i=1}^n c_i \mathbb{E}(I_{B_k}I_{B_i}) = c_k\mathbb{E}(I_{B_k}) = c_k P(B_k),$$

and therefore

$$c_k = \frac{\mathbb{E}(I_{B_k}Y)}{P(B_k)}.$$

Exercise 8.1 *Let (X,Y) be bivariate Gaussian. Compute $\mathbb{E}(Y \mid X)$ and $\mathbb{E}(Y^2 \mid X)$.*

Exercise 8.2 *Let Y be square integrable. Prove that $\mathbb{E}(Y \mid \mathcal{G})$ and $Y - \mathbb{E}(Y \mid \mathcal{G})$ are uncorrelated.*

Exercise 8.3 *Let Y be square integrable. Prove that*

$$\sigma^2 = \mathbb{E}((Y - \mathbb{E}(Y \mid \mathcal{G}))^2) + \mathbb{E}((\mathbb{E}(Y \mid \mathcal{G}) - \mathbb{E}(Y))^2).$$

Conclude that $\mathbb{E}(Y \mid \mathcal{G})$ is also square integrable.

Exercise 8.4 *Let $X_1, \ldots, X_n$ be i.i.d. and integrable. Let $S = X_1 + X_2 + \cdots + X_n$. Find $\mathbb{E}(X_1 \mid S)$.*

Exercise 8.5 *Let $X \sim \mathbb{U}([-1,1])$. Compute $\mathbb{E}(|X| \mid X)$ and $\mathbb{E}(X \mid |X|)$. Compute also $\mathbb{E}(X \mid |X|)$ for $X \sim \mathbb{U}([-1,2])$ and for $X \sim f$, with f a density.*

Exercise 8.6 *Let* S_t*,* $t = 0, 1, 2$ *denote the value of an asset at time* t*. Assume that* $S_1 = S_0 e^{\mu+\sigma X_1}$ *and* $S_2 = S_0 e^{2\mu+\sigma(X_1+X_2)}$*, with* $\sigma > 0$*,* S_0, X_1, X_2 *independent and both* X_1 *and* X_2 *Gaussian with expectation* 0 *and variance* 1*. Compute* $\mathbb{E}(S_2 \mid S_1)$*.*

Exercise 8.7 *Show that if* $|Y| \le c$*, then* $|\mathbb{E}(Y \mid \mathcal{G})| \le c$*.*

Exercise 8.8 *Let* Y *be square integrable,* $\mathbb{E}(Y \mid X) = X$ *and* $\mathbb{E}(Y^2 \mid X) = X^2$*. Show* $Y = X$ *a.s.*

Exercise 8.9 *Let* $X \sim P(\lambda)$ *(Poisson). Let, conditional on* $X = x$*,* $Y \sim B(x, p)$*. Compute* $\mathbb{E}(Y \mid X)$ *and* $\mathbb{E}(X \mid Y)$*.*

Bibliography

[1] Karlin, S. and Taylor H. M. (1975). *A First Course in Stochastic Processes*, 2nd ed., Academic Press.

[2] Jacod, J. and Protter P. (2004). Probability Essentials. 2nd edition. Springer. ISBN 3-540-43871-8.