ARBITRAGE AND MARKET DYNAMICS: A PROBABILITY-BASED APPROACH

LEILA OUKACHA

Abstract. This paper explores the application of probability theory to arbitrage. While no advanced probability knowledge is assumed, familiarity with basic probability is required. After introducing probability spaces, random variables, and distributions to define martingales, we apply these concepts to arbitrage-free markets. We cover stock arbitrage in a single-period market and conclude with the fundamental theorem of arbitrage pricing. Ultimately, we extend our analysis to multi-period markets by deriving the martingale representation theorem.

Contents

Date: Ausgut 26th, 2024.

1. Preliminaries

1.1. Probability Spaces.

Definition 1.1.1. A **probability space** is a triple (Ω, $\mathcal{F}$, P) where Ω is a set of "outcomes", $\mathcal{F}$ is a set of "events", and $\mathcal{P} : \mathcal{F} \to [0,1]$ is a function that assigns probability to events.

Definition 1.1.2. (Ω, $\mathcal{F}$) is a **measurable space**. A **measure** is a nonnegative countably additive set function such that $\mu : \mathcal{F} \to \mathbb{R}$ satisfies the following properties:

(i) $\mu(A) \geq \mu(\emptyset) = 0$ for all $A \in \mathcal{F}$.
(ii) If $A_i \in \mathcal{F}$ is a countable sequence of disjoint sets, then $\mu\left(\bigcup_i A_i\right) = \sum_i \mu(A_i)$.
(iii) If $\mu(\Omega) = 1$, then μ is defined as a **probability measure** usually denoted by P.

Example 1.1.3. Discrete probability spaces. Let Ω be a finite or infinite countable set and $\mathcal{F}$ be the set of all subsets of Ω. Also, assume that $A \subset \mathcal{F}$.

Then, $P(A) = \sum_{\omega \in A} p(\omega)$, where $p(\omega) \geq 0$ and $\sum_{\omega \in \Omega} p(\omega) = 1$.

Proposition 1.1.4. *A* ***sigma-algebra*** *$\mathcal{F}$ on a set Ω is a* ***collection of subsets*** *of Ω satisfying the following properties:*
(i) $\Omega \in \mathcal{F}$.
(ii) If $A \in \mathcal{F}$, then $\Omega \setminus A \in \mathcal{F}$.
(iii) If $A_1, A_2, A_3, \ldots \in \mathcal{F}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Definition 1.1.5. The **Borel sigma-algebra** $\mathcal{B}(\mathbb{R})$ on $\mathbb{R}$ is the smallest sigma-algebra containing all open subsets of $\mathbb{R}$.

1.2. Random Variables.

Definition 1.2.1. A function $X : \Omega \to S$ is said to be a **measurable map** from $(\Omega, \mathcal{F})$ to $(S, \mathcal{S})$ if $\mathrm{X}^{-1}(\mathcal{B}) \equiv \{\omega : X(\omega) \in \mathcal{B}\} \in \mathcal{F}, \quad \forall \mathcal{B} \in \mathcal{S}$.

Definition 1.2.2. If $(S, \mathcal{S}) = (\mathbb{R}, \mathcal{B}(\mathbb{R})), then$ X is called a **random variable**.

Example 1.2.3. The indicator function of a set $A \in \mathcal{F}$ is a **random variable**:

$$1_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$$

Theorem 1.2.4. *If $\{\omega : X(\omega) \in A\} \in \mathcal{F}$ for all $A \in \mathcal{A}$ and $\mathcal{A}$ generates $\mathcal{S}$ such that $\mathcal{S}$ is the smallest σ-field that contains $\mathcal{A}$, then X is* ***measurable****.*

Proof. Letting $\{X \in B\}$ be equivalent to $\{\omega : X(\omega) \in B\}$, we have

$\{X \in \bigcup_i B_i\} = \bigcup_i \{X \in B_i\}$ and $\{X \in B^c\} = \{X \in B\}^c$.

So the class of sets $\mathcal{B} = \{B : \{X \in B\} \in \mathcal{F}\}$ is a σ-field.

Since $\mathcal{B} \supset \mathcal{A}$ and $\mathcal{A}$ generates $\mathcal{S}$, $\mathcal{B} \supset \mathcal{S}$. □

Theorem 1.2.5. *If $X : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ and $f : (S, \mathcal{S}) \to (T, \mathcal{T})$ are measurable maps, then $f(X)$ is a* ***measurable map*** *from $(\Omega, \mathcal{F})$ to $(T, \mathcal{T})$.*

Proof. Let $\mathcal{B} \in \mathcal{T}$. Then $f^{-1}(\mathcal{B}) \in \mathcal{S}$.

Thus, $\{\omega : f(X(\omega)) \in \mathcal{B}\} = \{\omega : X(\omega) \in f^{-1}(\mathcal{B})\} \in \mathcal{F}$. □

Theorem 1.2.6. *If $X_1, \ldots, X_n$ are random variables and $f : (\mathbb{R}^n, \mathcal{R}^n) \to (\mathbb{R}, \mathcal{R})$ is measurable, then $f(X_1, \ldots, X_n)$ is a* ***random variable****.*

Proof. First, observe that if $A_1, \ldots, A_n$ are Borel sets, then

$$\{(X_1, \ldots, X_n) \in A_1 \times \cdots \times A_n\} = \bigcap_i \{X_i \in A_i\} \in \mathcal{F}.$$

Since sets of the form $A_1 \times \cdots \times A_n$ generate $\mathcal{R}^n$, $(X_1, \ldots, X_n)$ is a random vector.

It follows by Theorem 1.2.4 that $X_1, \ldots, X_n$ is measurable.

So, by Theorem 1.2.5, $f(X_1, \ldots, X_n)$ is a measurable map from $(\mathrm{R}^n, \mathcal{R}^n) \to (\mathbb{R}, \mathcal{R})$, which proves that $f(X_1, \ldots, X_n)$ is a random variable. □

Theorem 1.2.7. *If $X_1, \ldots, X_n$ are random variables, then so are $\inf_n X_n$, $\sup_n X_n$, $\limsup_n X_n$, and $\liminf_n X_n$.*

Proof. First, observe that $\{\inf X_n < a\} = \bigcup_n \{X_n < a\} \in \mathcal{F}$. Similarly, $\{\sup X_n > a\} = \bigcup_n \{X_n > a\} \in \mathcal{F}$.

Now, remark that

$$\liminf_{n \to \infty} X_n = \sup_n \inf_{m \geq n} X_m$$

Note that $Y_n = \inf_{m \geq n} X_m$ is a random variable for each n, so $\sup_n Y_n$ is as well.

Similarly,

$$\limsup_{n \to \infty} X_n = \inf_n \sup_{m \geq n} X_m$$

$Z_n = \sup_{m \geq n} X_m$ is a random variable for each n, so $\inf_n Z_n$ is as well, which completes the proof. □

1.3. **Distributions.**

Definition 1.3.1. A **distribution function** of a random variable is a function $F : \mathbb{R} \to [0, 1]$ such that $F(x) = P(X \leq x)$ where X is a random variable.

Proposition 1.3.2. *Any distribution function F has the following properties:*

(i) F *is nondecreasing:* $F(x_1) \leq F(x_2)$ *whenever* $x_1 \leq x_2$.

(ii) $\lim_{x \to \infty} F(x) = 1, \quad \lim_{x \to -\infty} F(x) = 0.$

(iii) F *is right continuous:* $\lim_{y \downarrow x} F(y) = F(x)$.

(iv) *If* $F(x-) = \lim_{y \uparrow x} F(y)$, *then* $F(x-) = P(X < x)$.

(v) $P(X = x) = F(x) - F(x-)$.

Proof.

(i) Given $x \leq y$, we have $\{X \leq x\} \subset \{X \leq y\}$. By the monotonicity of probability measures,
$$P(X \leq x) \leq P(X \leq y).$$
Hence, $F(x) \leq F(y)$. Therefore, $F(x)$ is nondecreasing.

(ii)
- $\lim_{x\to\infty} F(x) = 1$:
  As $x \to \infty$, $\{X \leq x\} \to \Omega$, so $F(x) \to 1$.
- $\lim_{x\to-\infty} F(x) = 0$:
  As $x \to -\infty$, $\{X \leq x\} \to \emptyset$, so $F(x) \to 0$.

(iii) For any sequence $y_n \downarrow x$, $\{X \leq y_n\} \downarrow \{X \leq x\}$. By the right-continuity of measures,
$$\lim_{n\to\infty} F(y_n) = F(x).$$
Hence, F is right continuous.

(iv) If $F(x-) = \lim_{y\uparrow x} F(y)$, then $F(x-) = P(X < x)$:
$$\text{If } F(x-) = \lim_{y\uparrow x} F(y), \text{ then}$$
$$\{X \leq y_n\} \uparrow \{X < x\} \text{ as } y_n \uparrow x.$$
$$\text{Thus, } F(x-) = P(X < x).$$

(v)
$$P(X = x) = P(X \leq x) - P(X < x).$$
Using properties (iii) and (iv), we get $P(X = x) = F(x) - F(x-)$. □

Theorem 1.3.3. *If a function $F : \mathbb{R} \to [0,1]$ satisfies (i), (ii), and (iii) of Proposition 1.3.2, then it is the **distribution function** of some random variable.*

Proof. Let $\Omega = (0,1)$, $\mathcal{F} =$ the Borel sets, and $P =$ Lebesgue measure.

Fix $\omega \in (0,1)$ such that $\mathrm{X}(\omega) = \sup\{y : F(y) < \omega\}$.

We must show that

$$\{\omega : X(\omega) \leq x\} = \{\omega : \omega \leq F(x)\} \tag{1}$$

Observe that

- If $\omega \leq F(x)$, then $X(\omega) \leq x$, since $x \notin \{y : F(y) < \omega\}$.
- On the other hand, if $\omega > F(x)$, then, since F is right continuous, there exists $\epsilon > 0$ such that $F(x+\epsilon) < \omega$ and $X(\omega) \geq x + \epsilon > x$.

Therefore, $P(\{\omega : \omega \leq F(x)\}) = F(x)$ which proves (1). □

Example 1.3.4. Exponential distribution with rate λ such that $f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$

Distribution function: $F(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 - e^{-\lambda x} & \text{if } x \geq 0. \end{cases}$

2. Martingales

2.1. Conditional Expectation.

Definition 2.1.1. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$, and $X \in \mathcal{L}^1$ be a random variable. Then the random variable ξ, denoted by $\mathbb{E}[X \mid \mathcal{G}]$, is the **conditional expectation** of X with respect to $\mathcal{G}$ if

(i) $\xi \in \mathcal{L}^1$,
(ii) ξ is $\mathcal{G}$-measurable,
(iii) $\mathbb{E}[\xi \cdot 1_A] = \mathbb{E}[X \cdot 1_A]$, for all $A \in \mathcal{G}$.

Proposition 2.1.2. ***Properties of Conditional Expectation***
*Let $X \in L^2(\Omega, \mathcal{F}, P)$ and let $\mathcal{G}$ be a σ-algebra contained in $\mathcal{F}$. Then **Conditional Expectation** satisfies the following properties:*

(1) ***Linearity***: $\mathbb{E}(aX_1 + bX_2|\mathcal{G}) = a\mathbb{E}(X_1|\mathcal{G}) + b\mathbb{E}(X_2|\mathcal{G})$.
(2) ***Orthogonality***: $X - \mathbb{E}(X|\mathcal{G}) \perp L^2(\Omega, \mathcal{G}, P)$.
(3) ***Best Prediction***: $\mathbb{E}(X|\mathcal{G})$ *minimizes* $\mathbb{E}[(X-Y)^2]$ *among all* $Y \in L^2(\Omega, \mathcal{G}, P)$.
(4) ***Tower Property***: *If $\mathcal{H}$ is a σ-algebra contained in $\mathcal{G}$, so that $\mathcal{H} \subseteq \mathcal{G} \subseteq \mathcal{F}$, then*
$$\mathbb{E}(X|\mathcal{H}) = \mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{H}).$$
(5) ***Covariance Matching***: $\mathbb{E}(X|\mathcal{G})$ *is the unique random variable* $Z \in L^2(\Omega, \mathcal{G}, P)$ *such that for every* $Y \in L^2(\Omega, \mathcal{G}, P)$,
$$\mathbb{E}(XY) = \mathbb{E}(ZY).$$
(6) ***Normalization***: $\mathbb{E}(1|\mathcal{G}) = 1$ *almost surely.*
(7) ***Positivity***: *For any nonnegative, bounded random variable X,*
$$\mathbb{E}(X|\mathcal{G}) \geq 0 \text{ almost surely.}$$
(8) ***Monotonicity***: *If X, Y are bounded random variables such that $X \leq Y$ almost surely, then*
$$\mathbb{E}(X|\mathcal{G}) \leq \mathbb{E}(Y|\mathcal{G}) \text{ almost surely.}$$
(9) ***Jensen's Inequalities***: *If $\phi : \mathbb{R} \to \mathbb{R}$ is convex and $\mathbb{E}|X| < \infty$, then*
$$\mathbb{E}(\phi(X)) \geq \phi(\mathbb{E}(X)) \quad \text{and} \quad \mathbb{E}(\phi(X)|Y) \geq \phi(\mathbb{E}(X|Y)).$$

Proof. Omitted. □

Theorem 2.1.3. ***Monotone Covergence Theorem*** *Let X_n be a sequence of random variables such that $X_n \geq 0$ for all n, and $X_n \to X$ as $n \to \infty$.*

Then, $\mathbb{E}(X_n) \to \mathbb{E}(X)$ as $n \to \infty$.

Proof. Omitted. □

Theorem 2.1.4. ***Conditional Monotone Convergence Theorem***

Let X_n be a nondecreasing sequence of nonnegative random variables on a probability space $(\Omega, \mathcal{F}, P)$, and let $X = \lim_{n\to\infty} X_n$. Then for any σ-algebra $\mathcal{G} \subset \mathcal{F}$,

$$\mathbb{E}(X_n \mid \mathcal{G}) \uparrow \mathbb{E}(X \mid \mathcal{G}).$$

Proof. By Linearity (1) and Positivity (7) in Proposition 2.1.2,

$$\mathbb{E}(X_n \mid \mathcal{G}) \le \mathbb{E}(X_{n+1} \mid \mathcal{G}) \le \mathbb{E}(X \mid \mathcal{G}) \quad \forall n \in \mathbb{N}.$$

Then, the limit $V := \lim_{n\to\infty} \mathbb{E}(X_n \mid \mathcal{G})$ exists almost surely, and $V \le \mathbb{E}(X \mid \mathcal{G})$. Moreover, since each conditional expectation is $\mathcal{G}$-measurable, so is V.

To prove $V = \mathbb{E}(X \mid \mathcal{G})$ almost surely, define the set

$$B = \{V < \mathbb{E}(X \mid \mathcal{G})\}.$$

We need to show that P(B) = 0.

Since $B \in \mathcal{G}$, we have by Definition 2.1.1 and the Tower Property (4) in Proposition 2.1.2,

$$\mathbb{E}(X \cdot 1_B) = \mathbb{E}(\mathbb{E}(X \mid \mathcal{G}) \cdot 1_B) \quad \text{and} \quad \mathbb{E}(X_n \cdot 1_B) = \mathbb{E}(\mathbb{E}(X_n \mid \mathcal{G}) \cdot 1_B).$$

But, by Theorem 2.1.3,

$$\mathbb{E}(X \cdot 1_B) = \lim_{n\to\infty} \mathbb{E}(X_n \cdot 1_B) \quad \text{and} \quad \mathbb{E}(V \cdot 1_B) = \lim_{n\to\infty} \mathbb{E}(\mathbb{E}(X_n \mid \mathcal{G}) \cdot 1_B),$$

so $\mathbb{E}(X \cdot 1_B) = \mathbb{E}(V \cdot 1_B)$ which is a contradiction since $V < X$ on B.

Thus, $P(B) = 0$. □

2.2. Martingale.

Definition 2.2.1. Let $\{\mathcal{F}_n\}$ denote the **information** in $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_\infty$ such that **information** refers to the collection of data or events available up to a given time t. Then a **filtration** of a set Ω (finite or infinite) is defined to be a collection $\mathcal{F}_t$, indexed by a time parameter t (discrete or continuous), such that

(1) Each $\mathcal{F}_t$ is a σ-algebra of subsets (events) of Ω;
(2) If $s < t$, then $\mathcal{F}_s \subseteq \mathcal{F}_t$.

Example 2.2.2. Suppose $X_1, X_2, \dots$ are independent, identically distributed random variables with mean μ.

Let S_n denote the partial sum $S_n = X_1 + \cdots + X_n$ and $\mathcal{F}_n$ denote the information in $X_1, \dots, X_n$.

Suppose $m < n$. Then by linearity in Proposition 2.1.2,

$$\mathbb{E}(S_n \mid \mathcal{F}_m) = \mathbb{E}(X_1 + \cdots + X_m \mid \mathcal{F}_m) + \mathbb{E}(X_{m+1} + \cdots + X_n \mid \mathcal{F}_m). \tag{1}$$

Since $X_1 + \cdots + X_m$ is measurable with respect to $\mathcal{F}_m$, by the constants property stating that for any scalar a, $\mathbb{E}(a \mid \mathcal{G}) = a$, we have:

$$\mathbb{E}(X_1 + \cdots + X_m \mid \mathcal{F}_m) = X_1 + \cdots + X_m = S_m. \tag{2}$$

Since $X_{m+1} + \cdots + X_n$ is independent of $\mathcal{F}_m$, we have:

$$\mathbb{E}(X_{m+1} + \cdots + X_n \mid \mathcal{F}_m) = \mathbb{E}(X_{m+1} + \cdots + X_n) = (n-m)\mu. \tag{3}$$

Therefore, by (2) and (3) in (1),

$$\mathbb{E}(S_n \mid \mathcal{F}_m) = S_m + (n-m)\mu.$$

Definition 2.2.3. A sequence of random variables $M_0, M_1, M_2, \ldots$ with $\mathbb{E}(|M_i|) < \infty$ is a **martingale** with respect to $\{\mathcal{F}_n\}$ if each M_n is measurable with respect to $\mathcal{F}_n$ such that for each $m < n$,

$$\mathbb{E}(M_n \mid \mathcal{F}_m) = M_m. \tag{1}$$

Equivalently,

$$\mathbb{E}(M_n - M_m \mid \mathcal{F}_m) = 0. \tag{2}$$

Example 2.2.4. "Martingale betting strategy"

Suppose $X_0, X_1, \ldots$ are independent random variables with

$\mathbb{P}\{X_i = 1\} = \mathbb{P}\{X_i = -1\} = \frac{1}{2}$.

Then think of the random variables X_i as the results of a game where one flips a coin where:

(i) If it comes up heads, then one wins \$1.
(ii) If it comes up tails, then one loses \$1.

To beat the game, we will keep doubling our bet until we eventually win. At this point, we stop.

Let W_n denote the winnings (or losses) up through n flips of the coin using this strategy and let $W_0 = 0$. Whenever we win, we stop playing, so our winnings stop changing such that

$$\mathbb{P}\{W_{n+1} = 1 \mid W_n = 1\} = 1.$$

Now, suppose that the first n flips of the coin have turned up tails. After each flip, we have doubled our bet, so we have lost $1 + 2 + 4 + \ldots + 2^{n-1} = 2^n - 1$ dollars and $W_n = -(2^n - 1)$. At this time, we double our bet again and wager 2^n on the next flip. This gives

$$\mathbb{P}\left\{W_{n+1} = -\left(2^{n+1} - 1\right) \mid W_n = -\left(2^n - 1\right)\right\} = \frac{1}{2}.$$

Therefore,

$$\mathbb{E}(W_{n+1} \mid \mathcal{F}_n) = W_n.$$

Hence, by Definition 2.2.3, W_n is a martingale with respect to $\mathcal{F}_n$.

Definition 2.2.5. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(\mathcal{F}_t)_{0\leq t\leq T}$ or $(\mathcal{F}_t)_{0\leq t<\infty}$ be a filtration by sub-σ-algebras of $\mathcal{F}$. An adapted sequence X_t of integrable random variables is defined to be a

- **Martingale** if $\mathbb{E}(X_{t+1} \mid F_t) = X_t, \forall t$.
- **Submartingale** if $\mathbb{E}(X_{t+1} \mid F_t) \geq X_t, \forall t$.
- **Supermartingale** if $\mathbb{E}(X_{t+1} \mid F_t) \leq X_t, \forall t$.

Example 2.2.6. Let $\{X_n\}_{n\geq 0}$ be a martingale relative to the filtration $\{\mathcal{F}_n\}_{n\geq 0}$, and $\phi : \mathbb{R} \to \mathbb{R}$ be a convex function such that $\mathbb{E}[\phi(X_n)] < \infty$ for each $n \geq 0$. Then the sequence $\{Z_n\}_{n\geq 0}$ defined by

$$Z_n = \phi(X_n)$$

is a **submartingale** relative to the filtration $\{\mathcal{F}_n\}_{n\geq 0}$ by Definition 2.2.5.

This is a consequence of Jensen's inequality and the martingale property of $\{X_n\}_{n\geq 0}$:

$$\begin{aligned}\mathbb{E}[Z_{n+1} \mid \mathcal{F}_n] &= \mathbb{E}[\phi(X_{n+1}) \mid \mathcal{F}_n] \\ &\geq \phi(\mathbb{E}[X_{n+1} \mid \mathcal{F}_n]) \\ &= \phi(X_n) = \ Z_n.\end{aligned}$$

Theorem 2.2.7. ***Martingale Convergence Theorem***

Suppose $M_0, M_1, \ldots$ is a martingale with respect to $\{\mathcal{F}_n\}$ such that there exists a $C < \infty$ with $\mathbb{E}(|M_n|) < C$ for all n. Then there exists a random variable M_∞ such that

$$M_n \to M_\infty$$

Proof. We will show is that for every $0 < a < b < \infty$, the probability that the martingale fluctuates infinitely often (*ie: the probability that the martingale diverges*) between a and b is 0.

Fix $a < b$. We will consider the following betting strategy:

1) Think of M_n as giving the cumulative results of some fair game and $M_{n+1} - M_n$ as being the result of the game at time $n + 1$.

2) Whenever $M_n < a$, bet 1 on the martingale. Continue this procedure until $M_n > b$.

3) Stop betting until $M_n < a$ again and return to betting 1. Continue this process, changing the bet to 0 when $M_n > b$ and changing back to 1 when $M_n < a$.

After n steps, the winnings in this strategy are given by

$$W_n = \sum_{j=1}^{n} B_j(M_j - M_{j-1}),$$

where B_j is the bet which equals 1 if $M_n < a$ or 0 if $M_n > b$.

Note that

$$W_n \geq (b-a)U_n - |M_n - a|,$$

where U_n denotes the number of times that the martingale goes between a and b (ie: *the number of upcrossings*) and $M_n - a$ gives an estimate for the amount lost in the last interval. Since W_n is a martingale, we have

$$\mathbb{E}(W_0) = \mathbb{E}(W_n) \geq (b-a)\mathbb{E}(U_n) - \mathbb{E}(|M_n - a|). \tag{1}$$

By the triangle inequality,

$$\mathbb{E}(|M_n - a|) \leq \mathbb{E}(|M_n|) + a \leq C + a. \tag{2}$$

Thus, by (1) and (2),

$$\mathbb{E}(U_n) \leq \frac{\mathbb{E}(W_0) + C + a}{b-a}.$$

Since this holds for every n, the expected number of upcrossings from a to b in $\mathbf{R}$ is bounded for any a,b in $\mathbf{R}$. Thus, the number of upcrossings is finite almost surely. □

Example 2.2.8. Let $X_1, X_2, \ldots$ be independent random variables with

$$\mathbb{P}\{X_i = \frac{3}{2}\} = \mathbb{P}\{X_i = \frac{1}{2}\} = \frac{1}{2}.$$

Let $M_0 = 1$ and for $n > 0$, let $M_n = X_1 X_2 \cdots X_n$ such that

$$\mathbb{E}[M_n] = \mathbb{E}[X_1]\mathbb{E}[X_2]\cdots\mathbb{E}[X_n] = 1.$$

If $\mathcal{F}_n$ denotes the information contained in $X_1, X_2, \ldots, X_n$, then

$$\begin{aligned}
\mathbb{E}[M_{n+1}|\mathcal{F}_n] &= \mathbb{E}[X_1 X_2 \cdots X_n X_{n+1}|\mathcal{F}_n] \\
&= X_1 X_2 \cdots X_n \mathbb{E}[X_{n+1}|\mathcal{F}_n] \\
&= X_1 X_2 \cdots X_n \mathbb{E}[X_{n+1}] \\
&= M_n.
\end{aligned}$$

Hence, by Definition 2.2.3, M_n is a martingale with respect to $X_1, X_2, \ldots X_n$.

Since $\mathbb{E}[|M_n|] = \mathbb{E}[M_n] = 1$, the conditions of the martingale convergence theorem hold. Thus,

$$M_n \to M_\infty \text{ for some random variable } M_\infty.$$

3. Application to arbitrage and no-arbitrage pricing in finance

3.1. Arbitrage.

Definition 3.1.1. An ***asset*** is a resource owned or controlled by an individual, organization, or entity that is expected to provide future economic benefits.

Definition 3.1.2. A ***position*** refers to the amount of a particular asset, security, or financial instrument that an individual, company, or entity holds. The term is commonly used in the context of investments, trading, and portfolio management.

Definition 3.1.3. ***Arbitrage*** is the act of taking simultaneous positions in different assets to guarantee a riskless profit higher than the risk-free rate, such as that from US Treasury bills.

Remark 3.1.4. By identifying and exploiting market inefficiencies, traders search for price discrepancies in large, liquid markets and execute rapid, high-volume trades to capitalize on these differences. They buy a security in one market and sell it in another at a higher price, using the sale proceeds to pay for the purchase and pocketing the profit.

3.2. **Arbitrage-free single-period market.**

Remark 3.2.1. In the absence of arbitrage, the market imposes a probability distribution, called a ***risk-neutral*** or ***equilibrium measure***, on the set of possible market scenarios. This probability measure determines market prices through ***discounted expectation***, which refers to the process of adjusting the expected future value of a financial asset or cash flow to account for the time value of money. This leads us to the ***Fundamental Theorem of Arbitrage Pricing***.

Definition 3.2.2. Single Period market

Consider a market in which

(i) The share price of asset A_j at time $t = 0$ is S_0^j. Assume that $S_0^1 = 1$.
(ii) $A_1, A_2, \ldots, A_K$ are K freely traded assets such that A_1 is riskless (ie: S_0^1 is independent of the market scenario).
(iii) The finite set $\Omega = \omega_1, \omega_2, \ldots, \omega_N$ corresponds to N possible market scenarios.
(iv) The share prices $S_1^2, S_1^3, \ldots, S_1^K$ of the $K - 1$ assets at time $t = 1$ are functions of the market scenario.

Example 3.2.3. In scenario ω_i, there is an $N \times K$ matrix with entries $S_1^j(\omega_i)$ corresponding to the price of a share of A_j at time $t = 1$.

Since A_1 is riskless by Definition 3.2.2, the share price S_1^1 of A_1 in any scenario ω_i where r is the riskless rate of return is defined as follows:

$$S_1^1(\omega_i) = e^r \quad \forall i = 1, 2, \ldots, N.$$

Definition 3.2.4. Portfolios. A portfolio is a vector

$$\theta = (\theta_1, \theta_2, \ldots, \theta_K) \in \mathbb{R}^K$$

of K real numbers such that θ_j is the number of shares of asset A_j that are owned. If $\theta_j < 0$, then the portfolio is said to be short $|\theta_j|$ shares of asset A_j.

Proposition 3.2.5.

(i) The value of the portfolio θ at time $t = 0$ is

$$V_0(\theta) = \sum_{j=1}^{K} \theta_j S_0^j$$

(ii) The value of the portfolio θ at time $t = 1$ in market scenario ω_i is

$$V_1(\theta; \omega_i) = \sum_{j=1}^{K} \theta_j S_1^j(\omega_i)$$

Definition 3.2.6. Arbitrage. An arbitrage is a portfolio θ that *makes money from nothing*. Formally, an arbitrage is a portfolio θ such that either

$$V_0(\theta) \leq 0 \quad \text{and} \quad V_1(\theta; \omega_j) > 0, \quad \forall j = 1, 2, \ldots, N$$

or

$$V_0(\theta) < 0 \quad \text{and} \quad V_1(\theta; \omega_j) \geq 0, \quad \forall j = 1, 2, \ldots, N.$$

Definition 3.2.7. Equilibrium Measure. If, for every asset A, the share price of A at time $t = 0$ is the discounted expectation, under π, of the share price at time $t = 1$ such that

$$S_0^j = e^{-r} \sum_{i=1}^{N} \pi(\omega_i) S_1^j(\omega_i) \quad \forall j = 1, 2, \ldots, K, \tag{1}$$

then a probability distribution $\pi_i = \pi(\omega_i)$ on the set Ω is an *equilibrium measure (or risk-neutral measure).*

Lemma 3.2.8. *Let F be a closed, bounded, convex subset of $\mathbb{R}^k$ and let $x \in \mathbb{R}^k - F$. Then there is a unique nonzero vector $v \in \mathbb{R}^k$ such that*

$$v \cdot x \leq v \cdot y \quad \forall y \in F,$$

Proof. Assume that $x = 0$ (the origin in $\mathbb{R}^k$). Let $v \in F$ be the element of F closest to the origin 0. Then such a point exists since F is closed and bounded; it is unique since F is convex; the vector v cannot be the zero vector since $0 \notin F$.

Since $x = 0$, we will show that $v \cdot y > 0$ for all elements $y \in F$.

The dot product is unchanged by rotations of $\mathbb{R}^k$ about the origin, so assume that the vector v lies on the first coordinate axis such that

$$v = (a, 0, 0, \ldots, 0) \quad \text{for some} \quad a > 0.$$

Thus, to prove that $v \cdot y > 0$ for all elements $y \in F$, it suffices to show that there is no $y \in F$ whose first coordinate is nonpositive.

Suppose, for the sake of contradiction, that there exists a point $y \in F$ with a nonpositive first coordinate. Then the line segment L has endpoints v and y such that L $\subset F$. Because this line segment must cross the hyperplane consisting of points with first coordinate 0, we may suppose that y has the form

$$y = (0, y_2, y_3, \ldots, y_k),$$

Let L consist of all points of the form $y(\epsilon) := ((1-\epsilon)a, \epsilon y_2, \epsilon y_3, \ldots, \epsilon y_k)$, where $0 \leq \epsilon \leq 1$.

Now, the closest point to the origin on L must be v. But, for all sufficiently small $\epsilon > 0$, the point $y(\epsilon)$ is actually closer to the origin than v, which is a contradiction. □

Theorem 3.2.9. ***Fundamental Theorem of Arbitrage Pricing.*** *There exists an equilibrium measure if and only if arbitrages do not exist.*

Proof.

(i) There exists an equilibrium measure $\implies$ arbitrages do not exist.

Suppose that there is an equilibrium measure π. Then for any portfolio θ, the portfolio values at time $t = 0$ and $t = 1$ are related by the discounted expectation:

$$V_0(\theta) = \sum_{i=1}^{N} \pi(\omega_i) e^{-r} V_1(\theta; \omega_i). \tag{1}$$

In the case for an arbitrage portfolio, if $V_1(\theta; \omega_i) > 0$ for every market scenario ω_i, then (1) implies that $V_0(\theta) > 0$, and so θ cannot be an arbitrage by Definition 3.2.6. Thus, arbitrages do not exist.

(ii) Arbitrages do not exist $\implies$ there exists an equilibrium measure.

We must show that if the market does not admit arbitrages, then it has an equilibrium measure π, that is, a probability distribution $\pi(\omega_i)$ on the set Ω of market scenarios ω_i such that (1) in Definition 3.2.7 holds.

First, for $j = 1$, asset 1 is the riskless asset, so its share price at time $t = 0$ is $S_0^1 = 1$ and its share price at time $t = 1$, under any scenario ω_i, is e^r. So for any probability distribution π on the set of market scenarios,

$$1 = S_0^1 = e^{-r} \sum_{i=1}^{N} \pi(\omega_i) e^r = e^{-r} \sum_{i=1}^{N} \pi(\omega_i) S_1^1(\omega_i).$$

Thus, if $j = 1$, then equation (1) in Definition 3.2.7 holds.

Second, for $2 \leq j \leq K$, consider the set E of all vectors $y = (y_2, y_3, \ldots, y_K)$ that can be obtained from the discounted share prices by averaging against some probability distribution π on Ω. Then,

$$y_j = e^{-r} \sum_{i=1}^{N} \pi(\omega_i) S_1^j(\omega_i) \quad \forall j = 2, 3, \ldots, K. \tag{2}$$

The set E is a bounded, closed, convex polytope in $\mathbb{R}^{K-1}$. Now, we must show that, in the absence of arbitrages, we have $S = (S_0^2, S_0^3, \ldots, S_0^K)$ where $S \subset E$.

Equivalently, we will show that if $S \notin E$, then there would be an arbitrage. Using Lemma 3.2.8, we will show that if the time-zero price vector S is not an element of E, then there is an arbitrage. Suppose, then, that $S \notin E$.

Since E is a closed, bounded, convex set, the Separating Hyperplane Theorem implies that there is a nonzero vector

$$\theta^* = (\theta_2, \theta_3, \dots, \theta_K).$$

Then by Lemma 3.2.8,

$$S \cdot \theta^* < y \cdot \theta^*, \quad \forall y \in E.$$

Because E includes points of the form (2) where the probability distribution π puts all its mass on a single scenario ω_i, it follows that, for each scenario ω_i,

$$e^{-r} \sum_{j=2}^{K} \theta_j S_1^j(\omega_i) > \sum_{j=2}^{K} \theta_j S_0^j.$$

We can choose a real number $-\theta_1$ that lies between these two values. Then adding θ_1 to both sides of the previous inequality shows that for every market scenario ω_i,

$$e^{-r} \sum_{j=1}^{K} \theta_j S_1^j(\omega_i) > 0 > \sum_{j=1}^{K} \theta_j S_0^j.$$

This implies that the portfolio $\theta = (\theta_1, \theta_2, \dots, \theta_K)$ is an arbitrage by Definition 3.2.6. Thus, if θ is not an arbitrage, then there exists an equilibrium measure.

□

3.3. The Martingale Representation Theorem and Hedging in a multi-period market.

Remark 3.3.1. In a ***multi-period market***, information about the market scenario is revealed in ***stages***. Some events may be completely determined by the end of the first trading period, others by the end of the second, and others not until the termination of all trading, which suggests the following classification of events.

Definition 3.3.2. Filtration in a multi-period market

For each $t \leq T$,

$$\mathcal{F}_t = \{\text{all events determined in the first } t \text{ trading periods}\}.$$

The finite sequence $(\mathcal{F}_t)_{0 \leq t \leq T}$ is a *filtration* (ie: Definition 2.2.1) of the space Ω of market scenarios.

Remark 3.3.3. The **share prices of assets** in a multiperiod market depend on market scenarios, but evolve in such a way that their values at any time t, being observable at time t, **do not depend on** the unobservable post-t futures of the scenarios.

Definition 3.3.4. Adapted Processes. The price process S_t of a traded asset is *adapted* to the natural filtration $(\mathcal{F}_t)_{0\leq t\leq T}$ by Definition 3.3.2. Formally, a sequence X_t of random variables is *adapted* to a filtration $(\mathcal{F}_t)_{0\leq t\leq T}$ if, for each t, the random variable X_t is $\mathcal{F}_t$*-measurable.*

Example 3.3.5. The Two-Period Binary Market.

There are four market scenarios:

$$\Omega = \{++, +-, -+, --\}.$$

(i) For each scenario, the first (respectively, second) entry indicates whether the share price of the asset stock increased or decreased in the first (respectively, second) trading period.

(ii) There are $4^2 = 16$ events (subsets of Ω) in all.

(iii) The only events that are determined before the first trading period are the trivial events $\emptyset$ and Ω.

(iv) There are two other events determined by time $t = 1$:

$$\mathcal{F}^+ = \{++, +-\} \quad \text{and} \quad \mathcal{F}^- = \{-+, --\}.$$

Consequently, the natural filtration (as specified in Definition 3.3.2) is

$$\mathcal{F}_0 = \{\emptyset, \Omega\},$$

$$\mathcal{F}_1 = \{\emptyset, \Omega, \mathcal{F}_+, \mathcal{F}_-\},$$

$$\mathcal{F}_2 = \{\text{all subsets of } \Omega\}.$$

Proposition 3.3.6. *The market $\mathcal{M}$ has scenario space $\Omega = \{+, -\}^T$, the set of all sequences of pluses and minuses of length T. Moreover, there is a **riskless asset bond** with rate of return r, and a **risky asset stock** whose price process evolves according to the rule*

$$S_{t+1}(\omega_1\omega_2\dots\omega_t+) = S_t(\omega_1\omega_2\dots\omega_t)u;$$

$$S_{t+1}(\omega_1\omega_2\dots\omega_t-) = S_t(\omega_1\omega_2\dots\omega_t)d.$$

Proposition 3.3.7. *If the riskless rate of return is $r = 0$, then the risk-neutral probability measure P is the probability measure on Ω under which the coordinate random variables ξ_t, defined by*

$$\xi_t(\omega_1\omega_2\dots\omega_T) = \omega_t \cdot 1 \quad \textit{for} \quad t = 1, 2, \dots, T,$$

are independent and identically distributed, with distribution

$$P\{\xi_t = +1\} = p \quad \textit{and} \quad P\{\xi_t = -1\} = q := 1 - p \quad \textit{where} \quad p = \frac{1-d}{u-d},$$

$$\mathcal{F}_t = \sigma(\xi_1, \xi_2, \dots, \xi_t).$$

Theorem 3.3.8. ***Martingale Representation Theorem.*** *If $(Y_t)_{0\le t\le T}$ is a martingale relative to the natural filtration $(\mathcal{F}_t)_{0\le t\le T}$, then $(Y_t - Y_0)_{0\le t\le T}$ is a martingale transform of the Stock price martingale $(S_t)_{0\le t\le T}$; that is, there exists a predictable sequence $(\beta_t)_{1\le t\le T}$ such that for each $t = 1, 2, \dots, T$,*

$$Y_t = Y_0 + \sum_{j=1}^{t} \beta_j (S_j - S_{j-1}).$$

Proof. Fix a scenario $\omega = \omega_1\omega_2 \dots \omega_T \in \Omega$ and let $G_t(\omega)$ be the set of all scenarios whose first t entries are $\omega_1\omega_2 \dots \omega_t$. Since the sequence $(Y_t)_{0\le t\le T}$ is a martingale, it follows that for every $t < T$ and every $\omega \in \Omega$,

$$\mathbb{E}[Y_{t+1}1_{G_t(\omega)}] = \mathbb{E}[Y_t 1_{G_t(\omega)}]. \tag{1}$$

Moreover, since the sequence $(Y_t)_{0\le t\le T}$ is adapted to the natural filtration, the value $Y_{t+1}(\omega) - Y_t(\omega)$ depends on the scenario $\omega = \omega_1\omega_2 \dots \omega_T$ only through its first $t + 1$ entries. Thus, equation (1) implies that, for each t and each ω,

$$pY_{t+1}(\omega_1\omega_2 \dots \omega_t+) + qY_{t+1}(\omega_1\omega_2 \dots \omega_t-) = Y_t(\omega_1\omega_2 \dots \omega_t). \tag{2}$$

Equation (2) also holds if Y_{t+1} and Y_t are replaced respectively by S_{t+1} and S_t because $(S_t)_{0\le t\le T}$ is also a martingale. Solving both equations for $-q/p$ leads to the relation:

$$\frac{Y_{t+1}(\omega_1\omega_2 \dots \omega_t+) - Y_t(\omega_1\omega_2 \dots \omega_t)}{Y_{t+1}(\omega_1\omega_2 \dots \omega_t-) - Y_t(\omega_1\omega_2 \dots \omega_t)} = -\frac{q}{p} = \frac{S_{t+1}(\omega_1\omega_2 \dots \omega_t+) - S_t(\omega_1\omega_2 \dots \omega_t)}{S_{t+1}(\omega_1\omega_2 \dots \omega_t-) - S_t(\omega_1\omega_2 \dots \omega_t)},$$

which in turn implies that

$$\begin{aligned}\frac{Y_{t+1}(\omega_1\omega_2 \dots \omega_t+) - Y_t(\omega_1\omega_2 \dots \omega_t)}{S_{t+1}(\omega_1\omega_2 \dots \omega_t+) - S_t(\omega_1\omega_2 \dots \omega_t)} &= \frac{Y_{t+1}(\omega_1\omega_2 \dots \omega_t-) - Y_t(\omega_1\omega_2 \dots \omega_t)}{S_{t+1}(\omega_1\omega_2 \dots \omega_t-) - S_t(\omega_1\omega_2 \dots \omega_t)} \\ &:= \beta_t(\omega_1\omega_2 \dots \omega_t).\end{aligned} \tag{3}$$

Remark that the common value of the fractions on the two sides of equation (3) depends only on $\omega_1\omega_2 \dots \omega_t$, so the definition of β_t is valid, which proves Theorem 3.3.8. ☐

Acknowledgments

It is my pleasure to thank my mentor, Jakob Wellington, for his invaluable guidance throughout the writing of this paper. I would also like to extend my thanks to Professor Rudenko for conducting apprentice lectures, especially in group theory. Special thanks to Professor Lawler for delivering lectures on probability and analysis. Finally, my sincere appreciation to Professor May for organizing the Math REU, which allowed me to gain valuable insights into the experience of conducting mathematical research for the first time.

4. Bibliography

References

[1] Lawler, Gregory F. *Introduction to Stochastic Processes*. Chapman and Hall/CRC, 1995.

[2] Durrett, Rick. *Probability: Theory and Examples*. 5th ed. Cambridge: Cambridge University Press, 2019.

[3] Lalley, Steven P. *Conditional Expectation*. University of Chicago, 2015. http://galton.uchicago.edu/~lalley/Courses/383/ConditionalExpectation.pdf

[4] Zitkovic, Gordan. "Lecture Notes Page." *University of Texas at Austin.* Accessed June 23, 2024. https://web.ma.utexas.edu/users/gordanz/lecture_notes_page.html.

[5] Kwok, Yue Kuen. *An Introduction to the Mathematics of Financial Derivatives.* Elsevier, 2008. https://www.sciencedirect.com/book/9780123846822/an-introduction-to-the-mathematics-of-financial-derivatives.

[6] Harvard Business School. *Online Business Insights.*

[7] Lalley, Steven P. "Lectures 3 and 4: Martingales." *University of Chicago.* http://galton.uchicago.edu/~lalley/Courses/390/Lecture3.pdf

[8] Lalley, Steven P. "The Fundamental Theorem of Arbitrage Pricing." *University of Chicago.* http://galton.uchicago.edu/~lalley/Courses/390/Lecture1.pdf