It is not a priori quite clear to the author why exactly the standard framework of measure-theoretic probability (and financial calculus, in particular) is defined as such. This article aims to derive it from common sense and not pull any insights from higher intelligence.

1. Available information

A fundamental concept underlying financial mathematics is the mathematical model of *available information* — the way we encode what we know and what we don't know about a real-world event or process.

1.1. Setup

Imagine a probability experiment that can be "run" and on which "observations" can be made.

1. A single run fixes (determines) *everything* that can be observed about the experiment.
2. A *run* is commonly denoted by ω, and the set of *all* runs is commonly denoted by Ω.
3. Probabilities are assigned to sensible subsets of Ω, forming a σ-algebra $\mathcal{F} \subset \mathcal{P}(\Omega)$ and a probability measure $\mathbb{P} : \mathcal{F} \to [0, 1]$.
4. In practice, we do not know exactly which run $\omega \in \Omega$ occurred. Instead,
   a. we can only know if the run is *among whole sets* A of runs;
   b. we can usually only "look" at the experiment, and *observe* values of projections $X : \Omega \to E$ onto other sets E.

1.1.1. Events

Such a set of runs A, that we can know if ω lies in, is called *an event* and is only allowed to be in $\mathcal{F}$. Thus it is in the domain of $\mathbb{P}$ and has a well-defined *probability* $\mathbb{P}(A)$. The model does not really talk about whether a run is in a set not in $\mathcal{F}$.

1.1.2. Random variables

If we observe a particular value $e = X(\omega)$ through a projection $X : \Omega \to E$, then we know that the run that happened lies in $X^{-1}(\{e\}) \subset \Omega$. But we're only allowed to know if ω is in $\mathcal{F}$-sets, so the only (sets of) values $A \subset E$ (including the case $A = \{e\}$) we are allowed to observe from X must have measurable preimages $X^{-1}(A) \in \mathcal{F}$. Thus to model observables we use *measurable* functions $X : \Omega \to E$ into measure spaces E, and we call X a *random variable*.

In this sense, $\mathcal{F}$ is a restriction on the nature of all random variables, containing the information *that is allowed to be known* and revealed by any observables.

2. Experiments that happen in time

Imagine we're observing an experiment that is laid out as a process over time $\mathbb{N}$.

2.1. Atomicity of history

A single run ω of the experiment contains the whole history of the process. That is, from a mathematical point of view, the future is predetermined.

As time evolves in the real world, though, we can only observe what is *currently* happening, and not the whole history, even though in theory it already "happened". Many runs of the experiment may have the same observable properties up to a particular time $n \in \mathbb{N}$, thus we only know that the run ω we're observing lies in a (probably large) measurable subset of Ω containing runs of the same observed history up to now but different futures.

For any n, let $\mathcal{F}_n \subset \mathcal{F}$ contain all maximal sets of events that share a common observable property up to time n or earlier. That is, $\mathcal{F}_n$ contains only, and all, sets of runs that a particular run can be determined to lie in, by observing it up to time n. Then trivially $\mathcal{F}_n \subset \mathcal{F}_m$ for all $n \leq m$.

If two runs ω and ω' differ in no observable way up to time n, they are not $\mathcal{F}_n$-separable.

Thus in order to model slowly-appearing history over Ω (recall, Ω is a set of atoms ω containing entire individual histories) we introduce a nested sequence of σ-algebras in $\mathcal{F}$

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots \subset \mathcal{F}_n \subset \ldots$$

called a *filtration* over Ω. It is a way to *lay out in time* the otherwise indivisible experiment runs ω without actually chopping them in time pieces. So at each step n we're allowed to know only which event in $\mathcal{F}_n$ has occurred and nothing more specific (like which run ω has happened).

2.2. Random processes

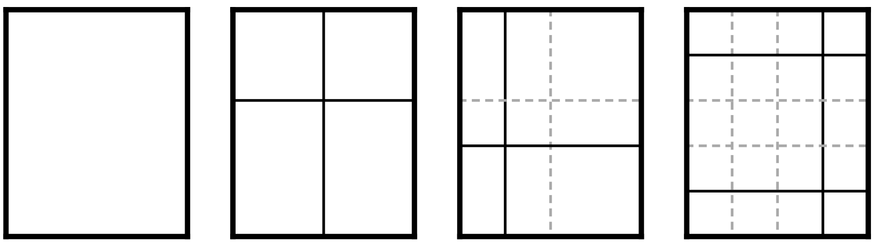
A sequence of real-world observations is then modelled by

- projections $X_n : \Omega \to E$ that distinguish only properties observable up to time n,
- i.e. $X_n(\omega) \neq X_n(\omega')$ (with both values in disjoint measurable sets in E) is only allowed if ω and ω' differ *in an observable way up to time* n, so that X_n does not accidentally reveal information about ω that is not supposed to be known at time n.
- That is, the preimages of observable (sets of) values in E must be maximal w.r.t. observables up to time n,
- i.e. $X_n^{-1}(A)$ must be in $\mathcal{F}_n$ for all measurable $A \subset E$.

Or, X_n must be $\mathcal{F}_n$-measurable for all n. Such a sequence $X_\bullet$ is called a *stochastic process.*

2.3. Filtration example diagram

Schematically we can imagine a filtration on a rectangular Ω like this:



$\mathcal{F}_0 = \{\emptyset, \Omega\}$ $\quad$ $\mathcal{F}_1$ $\quad$ $\mathcal{F}_2$ $\quad$ $\mathcal{F}_3$

A run ω is a point in the rectangle. Each of the four rectangles represents the same Ω but with a different σ-algebra on it. The σ-algebras $\mathcal{F}_n$ contain all rectangles outlined in each case (both dashed and non-dashed[1]):

1. In the beginning, we know nothing (except that the run *happened*), so no subrectangles.
2. In $\mathcal{F}_1$, the model allows to observe which of the four regions ω lies in, but nothing more.
3. In $\mathcal{F}_2$, we're allowed to observe which of the 9 (smaller!) rectangles it is in, and so on.

[1] non-dashed borders indicate the newly introduced measurable sets that weren't measurable in the previous step

2.4. Example

For an experiment with 3 consecutive coin flips we can put

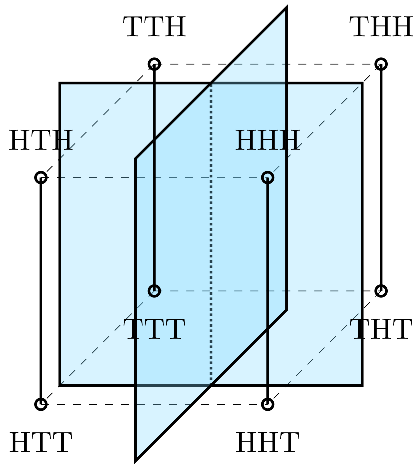
$$\Omega = \{\text{Heads}, \text{Tails}\}^3,$$

a discrete cube in 3D.

$\mathcal{F}_1$ contains only the front and back halves (apart from the whole cube and the empty set), $\mathcal{F}_2$ is generated by the four vertical edges, and $\mathcal{F}_3$ — by all 8 corners of the cube.

If $X_i : \Omega \to \{\text{Heads}, \text{Tails}\}$ is the outcome of the i th coin flip, then X_i is allowed to (though does not necessarily have to) differ only on different sets from $\mathcal{F}_i$, which it does.

Again, keep in mind that the entire runs of the experiment are fully encoded in the actual corners. The random process X_1, X_2, X_3 is only a sequence revealing smaller and smaller subsets in Ω around the true ω.



$\mathcal{F}_2$ is generated by the four vertical edges, i.e. the outcomes of the first two flips. Both X_1 and X_2 are $\mathcal{F}_2$-measurable, though neither of the two on its own separates the entire $\mathcal{F}_2$, only the vector (X_1, X_2) does.

3. Stopping times

A *stopping time* for a filtration $\mathcal{F}_\bullet$ is a $\mathbb{N}$-valued random variable, equivalently a random boolean sequence, equivalently a sequence of events.

4. Conditional expectation

If $\mathcal{G} \subset \mathcal{F}$ are σ-algebras on Ω, and $X : \Omega \to E$ is a function, then $\mathcal{G}$-measurability is a stronger restriction on X than $\mathcal{F}$-measurability because $X^{-1}(A) \in \mathcal{G}$ is a stronger condition than $X^{-1}(A) \in \mathcal{F}$. That is, an event $A \in E$ is allowed to reveal less information on the run ω in the $\mathcal{G}$ case than in the $\mathcal{F}$ case, just because $\mathcal{G}$ contains less information than $\mathcal{F}$.

Thus an $\mathcal{F}$-measurable X need not be $\mathcal{G}$-measurable.

Assume that $\mathcal{G}$ is finitely generated by random variables

$$\mathcal{G} = \sigma(Y, Z, \text{whatever})$$

5. Discrete-time market model

This section follows the first chapter of the book "Introduction to Stochastic Calculus Applied to Finance" (Lamberton & Lapeyre), whose explanations I found missing or unsatisfactory so I enhance the presentation with as much motivation as possible.

The market price vector is a stochastic process

$$S_i,\ i = 0, \ldots$$

with values in $\mathbb{R}^{1+d}$ adapted to a filtration

$$\mathcal{F}_i,\ i = 0, \ldots$$

of the market state known at time i. We consider $d \in \mathbb{N}$, $\mathcal{F}_\bullet$ and $S_\bullet$ fixed from now on.

The number $1 + d$ reflects that we track d risky assets and one riskless asset (the time-value of money) in the first component of $\mathbb{R}^{1+d}$.

Definition 5.1: A *trading strategy* is any $\mathbb{R}^{1+d}$-valued *predictable* sequence

$$\phi_i,\ i = 0, \ldots$$

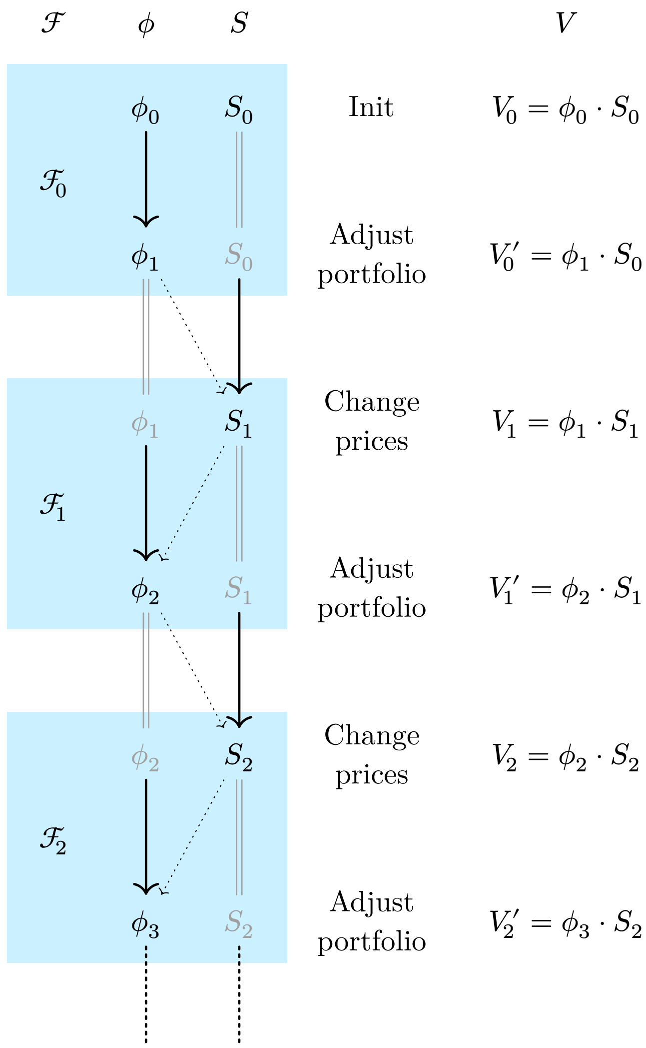
with respect to that filtration.

Predictable means that ϕ_{i+1} is $\mathcal{F}_i$-measurable, i.e. the information used (revealed) by the *next* portfolio ϕ_{i+1} is restricted to the information provided by the current market state $\mathcal{F}_i$, which is only guaranteed to contain the current stock prices (because S_i is $\mathcal{F}_i$-measurable) and past stock prices (because $\mathcal{F}_i \subset \mathcal{F}_{i+1}$).

Definition 5.2: The *value of the portfolio* defined by a strategy ϕ_i is

$$V_i(\phi) := \phi_i \cdot S_i.$$

Now a slight confusion may arise as to what is the time order of the quantities S_i and ϕ_i. They both use the same index i so we might be under the impression that the time evolution is described by a sequence of pairs (ϕ_i, S_i). While not false, this may be deceiving when we define *self-financing* strategies. Consider the diagram on the right. Thus we put (non-standard notation)



Here, in each timebox $\mathcal{F}_i$, first prices change and then the portfolio is adjusted, so we have an alternating sequence of price and portfolio changes.
The strategy ϕ is then self-financing iff the portfolio value doesn't change within a time-box ($V_i = V_i'$), but only responds to stock price changes $S_i \to S_{i+1}$.

Diagonal dotted arrows indicate the causal relationship between the time-evolution of the strategy and the stock prices in the process $(\phi_0, S_0), (\phi_1, S_1), \ldots$.

Definition 5.3: The *adjusted value of the portfolio* at time i after the new strategy ϕ_{i+1} is applied is

$$V_i'(\phi) \coloneqq \phi_{i+1} \cdot S_i$$

It is still $\mathcal{F}_i$-measurable.

Notation 5.4: Denote the *portfolio adjustment* (again, non-standard terminology) performed at time i by

$$\Delta\phi_i \coloneqq \phi_{i+1} - \phi_i.$$

(consider the diagram above to convince yourself that $\Delta\phi_i$ is $\mathcal{F}_i$-measurable)

It is a vector of the sells/buys of each stock performed at time i. That is,

$\Delta\phi_i^j$ is the amount of stock j that is sold (if $\Delta\phi_i < 0$) or bought (if $\Delta\phi_i > 0$) at time i.

Notation 5.5: Denote the *price changes* that happened at time $n+1$ by

$$\Delta S_{n+1} \coloneqq S_{n+1} - S_n.$$

Remark 5.6: $\Delta\phi_n$ was defined as the *forward* difference (between times $n+1$ and n), while the ΔS_n was defined as the *past* difference (between times n and $n-1$), so that both $\Delta\phi_n$ and ΔS_n are $\mathcal{F}_n$-measurable.

5.1. Self-financing strategies

Definition 5.1.1: A strategy ϕ_i is *self-financing* if the value of the portfolio stays the same after the adjustment:

$$V_i(\phi) = V_i'(\phi) \qquad \text{for all } i,$$

i.e. $\phi_{i+1} \cdot S_i = \phi_i \cdot S_i$ for all i.

Remark 5.1.2: The self-financing condition $V_i = V_i'$ is equivalent to

$$\Delta\phi_i \cdot S_i = 0,$$

i.e. all buys and sells cancel each other in value, i.e. no money is lost or needs to be brought in for the adjustment.

It should be close to mind that even for a strategy that is not self-financing, if we were to put the quantity $\Delta\phi_n \cdot S_n$ into ϕ_n^0, i.e. if we invest (borrow) the surplus (shortage) of sells-buys into (from) the riskless asset, we'd get a self-financing strategy.

Proposition 5.1.1: Any $\mathbb{R}^d$-valued predictable sequence (a "strategy" only on the risky assets)

$$\left(\phi_n^1, ..., \phi_n^d\right) \qquad \text{for} \ \ n = 0, 1, ...$$

is a restriction of a unique self-financing strategy ($\mathbb{R}^{1+d}$-valued)

$$\left(\boldsymbol{\phi_n^0}, \phi_n^1, ..., \phi_n^d\right) \qquad \text{for} \ \ n = 0, 1, ...$$

for any choice of an initial value $V_0(\phi) \in \mathbb{R}$.

In other words, the self-financing strategies restricting to a given $\left(\phi_n^1, ..., \phi_n^d\right)_n$ are a one-parameter family indexed by V_0.

Proof: Fix $\underline{\phi}_n = \left(\phi_n^1, ..., \phi_n^d\right)$ for all $n \geq 0$. Any strategy $\phi_\bullet$ that (on the risky assets) restricts to this process is determined by the choices of bank account amounts $\varphi_\bullet^0$.

The self-financing condition $\Delta\phi_n \cdot S_n = 0$ imposes a restriction on $\varphi_\bullet^0$, though, and reads

$$\underbrace{\left(\phi_{n+1}^0 - \phi_n^0\right)} \cdot S_n^0 + \sum_{i=1}^{d} \left(\phi_1^i - \phi_n^i\right) \cdot S_n^i = 0 \qquad \text{for all } n = 0, 1, ...$$

Here a choice of any element of the sequence $\phi_\bullet^0$ determines the rest, so the self-financing strategies restricting to $\underline{\phi}_n$ are a one-parametric family indexed by (e.g.) ϕ_0^0.

Since $V_0 \equiv \phi_0^0 \cdot S_0 + \sum_{i=1}^d \phi_0^i \cdot S_0^i$ establishes a 1:1 relation between V_0 and ϕ_0^0 (given $\underline{\phi}_0$), we can also consider this parameter to be V_0. □

Proposition 5.1.2: