Probability and Finance
Its only a game!

G. Shafer and V. Vovk
(Wiley-Interscience, 2001)

1 Introduction

p2 Two fundamental ideas:
(1) **Principle of pricing by dynamical hedging**: when simple gambles are combined over time to produce a more complex gamble, prices for the simple gambles determine the prices for the complex one.
(2) **Hypothesis of Impossibility of a gambling system**: There is no way to select gambles to avoid bankruptcy while to give a reasonable chance of making us rich.

Probability becomes game-theoretic as soon as we treat the expected values as prices in a game.
Defining a probability measure means recommendation of prices for uncertain payoffs on the sample space.
p3 In contrast to the measure-theoretical framework, the game-theoretical framework can model open processes = processes open to influences we cannot model even probabilistically.

Game considered: Player I = Skeptic bets on what will happen and Player II = Reality decides what will happen. The moves are made alternatively.

Protocol for a probability game
The protocol for a probability game specifies moves available to each player:
(1) For Reality it is the sample space Ω (but sample = path, so this should better be called the path space). A cylinder set is called a *situation*.
Skeptics's bets do not affect what is possible in the world, although Reality may consider them in deciding what to do next.
The game is called *terminating* if every path is finite; the game has a *finite horizon*.
Real valued function on Ω is called simply a variable (to avoid the word 'random'). Any subset of Ω is called an *event*.
(2) For Skeptic each move is a gamble, defined by a price to be paid immediately and by a payoff that depends on Reality's following move. Skeptic can borrow money freely without interest.
Skeptic can combine (linearly average) available gambles.

p11 The protocol is *symmetric* if Skeptic is allowed to take either side of any available gamble: if he can buy the payoff x at the price m, he can also sell x at m (i.e., buying $-x$ at $-m$). In this case the available gambles make a linear space.
A *strategy* $\mathcal{P}$ is a plan for how to gamble in each nonterminal situation. For $\mathcal{P}$ and a situation t, $K^{\mathcal{P}}(t)$ is the capital of Skeptic

p12 **Upper and lower prices**
A strategy simulates a transaction satisfactorily if it produces at least as good net payoff $x - \alpha$. We say $\mathcal{P}$ simulates buying x at α satisfactorily, if

$$K^{\mathcal{P}}(\xi) \geq x(\xi) - \alpha \tag{1.1}$$

for every path ξ in Ω.

$$\overline{E}x = \inf\{\alpha \,|\, \text{there is some strategy } \mathcal{P} \text{ such that } K^{\mathcal{P}} \geq x - \alpha\} \tag{1.2}$$

is called the *upper price* of x or the *cost* of x. It is the lowest price Skeptic can buy x (the lowest price of x Skeptic can hedge)..

$$\underline{E}x = \sup\{\alpha \,|\, \text{there is some strategy } \mathcal{P} \text{ such that } K^{\mathcal{P}} \geq \alpha - x\} \tag{1.3}$$

is called the *lower price* of x or the *scrap value* of x (the highest price at which Skeptic can sell x).

$$\underline{E}x = -\overline{E}(-x). \tag{1.4}$$

$\overline{E}_t x$ is the upper price in the situation t, etc.

Hedging
We say $\mathcal{P}$ hedges the obligation y if for every path ξ

$$K^{\mathcal{P}}(\xi) \geq y(\xi). \tag{1.5}$$

We say $\mathcal{P}$ hedges selling x at price α, if $\mathcal{P}$ hedges $x - \alpha$. Thus, $\overline{E}x$ is the lowest selling price for x Skeptic can hedge.
(Notice that these definitions may not make sense if the game is not terminating.)

Game theoretic expectation and variance
p14 If Skeptic cannot make money for sure, then the protocol is said to be *coherent*.[1]
In this case

$$\underline{E}_t x \leq \overline{E}_t x \tag{1.6}$$

[1] For example, the range of x is $[-c, c]$ and the mean m is outside this range, then $M(x - m) > 0$ by appropriate choice of M independent of x.

for every x and they are zero for the variable whose value is always 0 for any path.
If $\overline{E}_t x = \underline{E}_t x$, the common value is called the *(exact) price*, and is designated by $E_t x$. This is the expected price.
variance is defined as

$$\overline{V}_t x = \overline{E}_t (x - E_t x)^2, \tag{1.7}$$

(upper variance), etc.

Interpretative hypothesis
The Principle of 'no risk no gain' is not a mathematical statement; it is neither an axiom nor a theorem. An event is unlikely if it violates this principle.

Game theoretic probability
Upper probability $\overline{P}$ of an event is the upper price of its indicator, $\overline{P}E = \overline{E}\chi_E$, where χ_E is the indicator of the event E, etc.

p126 The upper probability of an event E measures the degree to which a strategy for betting on E can multiply one's capital without risk of bakruptcy:

$$\begin{aligned} \overline{P}E \;=\; & \{\alpha \mid \text{ there is a strategy that begins with } \alpha \\ & \text{and ends up with at least 1 if } E \text{ happens and at least 0 otherwise}\} \end{aligned} \tag{1.8}$$

If we assume that the protocol is coherent (p14), then

$$0 \le \underline{P} \le \overline{P} \le 1, \tag{1.9}$$

and for some event E

$$\underline{P}E = 1 - \overline{P}E^c. \tag{1.10}$$

Suppose $\overline{P}E = 0.001$. Skeptic can buy this at 0.001. If E happens $\chi_E = 1$, so the gain is 1000. Since there is no possibility of bankruptcy because $\chi \ge 0$, this is unlikely. Hence, E is unlikely.
If $\underline{P}E = 0.999$, then $\overline{P}E^c = 0.001$, very unlikely, so E is likely.
We say an event is *practically unlikely*, if Skeptic's capital never goes negative, but increases without bound. E is almost sure if E^c is practically impossible.

p19 Game theoretical probability concept accommodates both objective and subjective concepts. The concepts may be classified by the ultimate authority for the price:
(1) The statistical regularity is the authority for objective probability. We expect events assigned small upper probabilities not to happen, and we expect prices to be reflected in average values.
(2) Belief is the authority for subjective probability. In this case low probability implies that he thinks an event will not happen, i.e., he is willing to

bet heavily against it. In this ‘neosubjectivist conception,’ the principle of fundamental interpretative hypothesis is irrelevant.
(3) The market for financial security is the authority for the market game, and the probability concept is about hedging of market risks.
p46 Game theory is a mathematical account of potentiality: it analyzes what players can do, so game theoretical probability is a measure of potentiality.

p23-4 **Pricing options**
Example. Stock price today is \$8. Option x: can buy the stock tomorrow at \$8.
(a) Tomorrow, it may go to \$5 or \$10.
$Ex = 6$.
(b) Tomorrow, it may go to \$5, \$8, or \$10.
$\overline{E} = 6$ and $\underline{E} = 0$, so no pricing without extra conditions.

2 Historical

Kolmogorov's axioms p40-41.
p45 After *Grundbegriffe*, probability was mathematics.
p43 Kolmogorov regarded his axioms as axioms for a frequentist concept of probability: LLN and “If PE is very small, one can be practically certain that this event would not occur at all if the experiment is realized only once.”

p44 **Cournot's bridge**: Events with zero probability cannot happen. This bridges probability theory and the physical world.
Once this is accepted, LLN is a theorem.

Von Mises collectives: Wald proposed in 1937 that we permit any rule for
p48 selecting a subsequence that can be expressed in a formal logic. Church used effective computability to support Wald. This marked the end of the intense discussion of collectives.

Ville's strengthening of collectives
Ville showed that that probability theory requires even more irregularity from a sequence than is required by the von Mises-Wald collective.
For example, the law of iterated logarithm requires such; no complexity oscil-
p48 lation, so to speak. One can construct collective whose subsequences always give the frequency of 1 to approach to 1/2 from above. In this case one's capital can increase without bound without any risk of bankruptcy. That is, there is a gambling system that can detect deviations from randomness that

cannot be detected by subsequence selection rules.
Ville's proposal is to add his 'universal gambling' in the subsequence choice rules proposed by Wald.

Kolmogorov complexity
Kolmogorov's idea was fully developed by Martin-Löf, showing the existence of the universal test. In 1971 Schnorr[2] showed that Ville's type tests could also give infinite random sequences.
Skeptic cannot become infinitely rich
=
Reality's moves will pass any computable test based on Forecaster's prices
=
Reality's moves will be random in the sense of Schnorr.
p51 Greater use of Kolmogorov complexity will be required to develop game theoretical probability theory.

Martingale
Let L be a real function on the history of gambling. If

$$E(L(x_1, \cdots, x_n) \mid x_1, \cdots, x_{n-1}) = L(x_1, \cdots, x_{n-1}), \tag{2.1}$$

we say L is *martingale*. Ville demonstrated that:
Event E is probability 1 iff there is a nonnegative martingale that diverges to infinity if E fails. Ville came almost to showing

$$P(E) = \inf\{L_0 \mid \liminf_{n\to\infty} L_n \geq \chi_E\} \tag{2.2}$$

Here L_0 is the initial value of L.

p55 **Impossibility of gambling system**
To elaborate the assertion that a gambling system cannot succeed in a fair game:
1. Any subsequence from an iid sequence has the same joint probability distribution,
2. If the expected gain is zero for a game, then no strategy to change it exist.
3. Player's gain in a fair game is a random variable with mean zero.
4. If a player follows a system that does not risk bankruptcy, then the odds are at least $K-1$ to 1 against his multiplying his stake by K.
p56 The last point is expressed by Ville in the form of *Doob's inequality*:

$$P\{\sup_n L_n \geq \lambda\} \leq \frac{1}{\lambda} \tag{2.3}$$

[2] C. P. Schnorr, "A unified approach to the definition of random sequences," Math. Systems Theory **5** 246 (1971) and in German in LNM 218 (1971).

where L is a nonnegative martingale.

p60 **Historical sources of game theoretical probability**
p30 (1) Principle of dynamic hedging due to Pascal
(2) Bridge due to Cournot
(3) Impossibility of gambling strategy due to Ville
(4) Dawid's *prequential principle* that testing should use only what actually happens.

3 Strong Law of Large Numbers

Fair coin game
i. Skeptic begins with an initial capital, say \$1.
ii. He bets by specifying some number $M \in \boldsymbol{R}$.
iii. Skeptic obtains $\$M$ if H appears, and loses $\$M$ if not.
p64 Thus, the protocol reads:
$K_0 = 1$
FOR $n = 1, 2, \cdots$
    Skeptic announces $M(n) \in \boldsymbol{R}$
    Reality announces $x(n) \in \{-1, 1\}$ (T or H)
    $K_n = K_{n-1} + M(n)x(n)$.
Skeptic wins, if (1) $K_n > 0$ for all n and (2)

$$\frac{1}{n}\sum x(n) \to 0 \quad \text{or} \quad K_n \to \infty \tag{3.1}$$

P 3.1 Game-theoretic strong law of large numbers. Skeptic has a winning strategy in the fair coin game.
We say an event happens *almost surely* if Skeptic has a winning strategy: $K_n \to \infty$ keeping $K_n > 0$ for all n.

The winning convention may be said in the following collateral way:
Skeptic tries to keep $K > 0$.
Reality tries to keep $K < \infty$. If Skeptic fails, he loses.
If both fails, then Reality wins.
Therefore, under this collateral description, Skeptic wins only if the LLN holds.

To prove this a more general game is considered:

Bounded forecasting game
The only difference is that the domain of x_n is $[-1, 1]$.

Here, *forecast* is the empirical average of x_n to be zero.
P 3.2 Skeptic has a winning strategy in the bounded forecasting game with forecasts set to zero.

We say the situation s precedes t if s is a prefix of t (we write $s \sqsubset t$). Let Ω^* be the set of all finite sequences of numbers taken from $[-1, 1]$. A real-valued function on Ω^* is called a *process*. Any process $\mathcal{P}$ is interpreted as a *strategy* for Skeptic. $\mathcal{P}(s)$ is the number M Skeptic chooses under situation s. Thus we may write

$$K(tx) = K(t) + \mathcal{P}(t)x. \tag{3.2}$$

Here tx is a concatenation of t and x (in this case an extension of t by one
p67 number $x \in [-1, 1]$). Notice that K is a linear function(al) of $\mathcal{P}$. That is, we may assume a linear portfolio of gambles.

Weakly forcing
We say a strategy $\mathcal{P}$ forces an event E if $K(t)$ bounded from below (by -1) grows without bound if $t \notin E$. If $\sup_n K(t_n) = \infty$, then we say $\mathcal{P}$ forces E *weakly*.

L3.1. If Skeptic can weakly force E, then he can force E.
[Demo] Let $\mathcal{P}$ be the weakly forcing strategy. We make a strategy $\mathcal{P}^C$ that allows an optional stopping as soon as the total capital reaches C.

$$\mathcal{P}^C(s) = \begin{cases} \mathcal{P}(s) \text{ if } K^{\mathcal{P}}(t) < C \text{ for all } t \sqsubset s, \\ 0, \text{ otherwise.} \end{cases} \tag{3.3}$$

Now, we can make such strategy for indefinitely large C. Make a new strategy $\mathcal{Q}$

$$\mathcal{Q} = \sum_{k=1}^{\infty} 2^{-k}\mathcal{P}^{2^k}. \tag{3.4}$$

$K^{\mathcal{Q}} = \infty$ but never below -1. So this $\mathcal{Q}$ is the forcing strategy.

Thus, we have only to find a weakly forcing strategy.

L3.2. If Skeptic can weakly force each of a sequence $E_1, E_2, \cdots$ of events, then he can weakly force $\cap E_k$.
[Demo] Let $\mathcal{P}^k$ be a weakly forcing strategy for E_k. This implies that $1 + K^{\mathcal{P}_k}$ is non-negative. Let us study the upper bound of $1 + K^{\mathcal{P}^k}(x_n)$.[3] For $n = 1$ M_1 is at most 1 to avoid bankruptcy. Therefore, $1 + K(x_1)$ is at most 2. Then, M_2 can be at most 2, and $1 + K(x_2)$ is at most 4. This way we can conclude that for any k

$$1 + K^{\mathcal{P}_k}(x_n) \le 2^n. \tag{3.5}$$

[3] x_n the length n prefix of x.

This implies also

$$|\mathcal{P}_k(x_n)| \le 2^n. \tag{3.6}$$

Therefore,

$$\mathcal{Q} = \sum_k 2^{-k}\mathcal{P}_k \tag{3.7}$$

is well-defined for any finite path. Since $\mathcal{P}_k$ weakly forces E_k, so is Q for all k (i.e., if one of E_k is violated, then Q grows indefinitely). Thus, $\mathcal{Q}$ is the desired strategy to weakly force the common set $\cap E_k$.

Proof of game theoretic strong LLN for coins
L3.3. For any $\epsilon > 0$ Skeptic can weakly force

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{\infty} x_i \le \epsilon \tag{3.8}$$

and

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^{\infty} x_i \ge -\epsilon. \tag{3.9}$$

[Demo] Let $\mathcal{P}$ be the strategy that always choose ϵ times the current capital $K^{\mathcal{P}}(x_n) + 1 = (K^{\mathcal{P}}(x_{n-1}) + 1)(1 + \epsilon x(n))$.[4] Since the initial capital is unity,

$$1 + K^{\mathcal{P}}(x_n) = \prod_{i=1}^{n}(1 + \epsilon x(i)). \tag{3.10}$$

If $K^{\mathcal{P}}(x)$ is bounded, then there is a constant D such that

$$\sum \log(1 + \epsilon x(i)) \le D, \tag{3.11}$$

where D may depend on x. Notice that $\log(1+t) \ge t - t^2$ for $t \ge -1/2$. Therefore, if we assume $\epsilon < 1/2$, then (3.11) implies

$$\epsilon \sum x(i) - \epsilon^2 \sum x(i)^2 \le D. \tag{3.12}$$

That is, since $x(i)^2 \le 1$,

$$\epsilon \sum x(i) - n\epsilon^2 \le D \tag{3.13}$$

or

$$\epsilon \sum x(i) \le D + n\epsilon^2. \tag{3.14}$$

Therefore, we conclude

$$\frac{1}{n}\sum x(n) \le \frac{D}{\epsilon n} + \epsilon. \tag{3.15}$$

[4] $x_n = x_{n-1}x(n)$.

That is, for sufficiently large n the average cannot be larger than ϵ. Replacing ϵ with $-\epsilon$, a lower bound can also be forced.□

To complete the proofpf LLN, we use L3.2. Choose $\epsilon = 2^{-k}$.

In the above $x(i)$ are all in $[-1, 1]$ and with zero price. We may choose different prices $m(i)$. Then,

$$\frac{1}{n}\sum(x(i) - m(i)) \to 0. \tag{3.16}$$

This price may be set freely before Skeptic places his bet M. Therefore, it is natural to introduce the *forecaster* who sets the price.

p70 **Bounded forecasting game (full but bounded version)**
Here $C > 0$ is a parameter; Players are Forecaster, Skeptic, and Reality.
Protocol:
$K_0 = 1$
FOR $n = 1, 2, \cdots$
  Forecaster announces $m(n) \in [-C, C]$,
  Skeptic announces $M(n) \in \boldsymbol{R}$,
  Reality announces $x(n) \in [-C, C]$.
  $K_n = K_{n-1} + M(n)(x(n) - m(n))$.
Skeptic wins if (1) $K_n > 0$ for all n and (2)

$$\frac{1}{n}\sum(x(n) - m(n)) \to 0 \quad \text{or} \quad K_n \to \infty \tag{3.17}$$

P3.3. Skeptic has a winning strategy in the bounded forecasting game.□

p71 Notice that when Forecaster announces $m(n)$,

$$\overline{E}_t x(n) = \underline{E}_t x(n) = m(n). \tag{3.18}$$

Here, $\overline{E}x$ is defined by $\inf_\alpha\{\alpha : \delta K(n) \geq M_n(x(n) - \alpha)\}$, where $\delta K(n) = K_n - K_{n-1}$. This implies $M_n m(n) \leq M_n \alpha$. We have both $M_n > 0$ and < 0, so the smallest α is obtained whan $M > 0$ as $\alpha = m(n)$. Similarly we obtain $\underline{E}x(n) = m(n)$.

Comparison with measure theoretical strong law
p71 In measure theoretical probability, the strong law for sequence x presupposes the joint probability distribution and the price is the conditional expected value of $x(n)$ given x_{n-1}. Therefore, the game theoretical formulation has less assumptions.

Computation of strategies

All the strategies so far discussed are computable, so
P3.5. Skeptic has a computable winning strategy in the bounded forecasting game.

Then, a natural question is the computational complexity of the move for the n-th step.

When LLN is not satisfied, we could ask how fast Skeptic can increase his capital. This is a game-theoretical large deviation principle.
P 3.6. Skeptic has a computable winning strategy in the bounded forecasting game with an unbounded K such that

$$\limsup_{n\to\infty} \frac{\log K(x_n)}{n} > 0 \tag{3.19}$$

□

p74 It might be interesting to study the trade-off between computational efficiency and the rate of capital increase.

4 Strong Law of Large Numbers (Unbounded Case)

Unbounded game
In this case variance must be bounded, so Forecaster must specify not only the price for $x(n)$ but also the price for $(x(n) - m(n))^2$.
Protocol:
$K_0 = 1$
FOR $n = 1, 2, \cdots$:
  Forecaster announces $m(n) \in \boldsymbol{R}$ and $v(n) \geq 0$.
  Skeptic announces $M(n) \in \boldsymbol{R}$ and $V(n) \geq 0$.
  Reality announces $x(n) \in \boldsymbol{R}$.
  The capital process is computed by

$$K_n = K_{n-1} + M_n(x(n) - m(n)) + V_n[(x(n) - m(n))^2 - v(n)]. \tag{4.1}$$

The winning rule is the same as before.
In this game, a path is defined by $v(1)x(1)v(2)x(2)\cdots$.

What happens may be summarized as
P4.1.
1. Skeptic has a winning strategy, if $\sum[v(n)/n^2] < \infty$.
2, Reality has a winning strategy, if $\sum[v(n)/n^2] = \infty$.

1 of P4.2 generalizes P3.2, because $V(n) = 0$ already gives a winning strategy if $x(n)$ are bounded.

Theorem 4.1.
1. Skeptic can force

$$\sum[v(n)/n^2] < \infty \Rightarrow \frac{1}{n}\sum(x(n) - m(n)) \to 0. \tag{4.2}$$

2. Reality can force

$$\sum[v(n)/n^2] = \infty \Rightarrow \frac{1}{n}\sum(x(n) - m(n)) \not\to 0. \tag{4.3}$$

p80 If Reality has (4.3) as her goal, then Skeptic cannot have any winning strategy.

$v(n)$ is an upper price of $(x(n) - m(n))^2$, because $V(n)$ is nonnegative.

We may set all $m(n) = 0$ without any loss of generality.

Capital process and supermartingale
L4.1. The capital processes form a convex cone: if S^i are capital processes defined by (4.1), then $\sum c_i S^i$ ($c_i \in [0, \infty)$) is also a capital process.□
This should be obvious, since K is linear in M and V.

p82 *Supermartigale*: A process of the form $T = S - B$, where B is an increasing process, is called a supermartingale process. The capital process S is also a supermartingale.
B may be understood as a cumulative money that has been thrown away.
If for a supermartingale process T,

$$S(x_n) - S(x_{n-1}) \geq T(x_n) - T(x_{n-1}) \tag{4.4}$$

S is called a *bounding capital process* for the supermartingale T, and the strategy that produces S is called a *bounding strategy.*
L4.2. Supermartingales form a convex cone. □

Minimal bounding strategy:
Let $\|\delta T\|$ for a supermartingale T be defined as

$$\begin{aligned} \|\delta T\|(svx) \quad &= \quad \inf\{\|(M, V)\| : V \text{ is nonnegative}, \\ &\qquad \text{and } T(svx) - T(s) \leq Mx + V(x^2 - v) \text{ for all } x \in \boldsymbol{R}\} \end{aligned} \tag{4.5}$$

for every s and every nonnegative number v. $\|(M,V)\| = \sqrt{M^2+V^2}$. $\|\delta T\|$ is a *deterministic process* because it does not depend on x. If the infimum is actually attainable, then $\|\delta T\|(s) = \|(M,V)\|(s)$. Such a strategy (M,V) is called a minimal bounding strategy for T.
L4.3. Every supermartingale has a minimal bounding strategy.
[Demo] Fixing s and v we show the infimum is attained. For each x the set

$$\mu(x) = \{(M,V) : V \text{ is nonnegative, and } T(svx) - T(s) \le Mx + V(x^2 - v) \text{ for all } \} \tag{4.6}$$

is a closed set. A supermartingale process has at least one bounding strategy by definition, so $\mu = \cap\mu(x) \neq \emptyset$. Therefore, μ is a nonempty closed set, so it must have an element closest to the origin.□

L4.4 Let T^k be a sequence of nonnegative supermartingale processes and $c_k \ge 0$. Write

$$T = \sum_{k=1}^{\infty} c_k T^k. \tag{4.7}$$

If the initial value is finite, and $\sum c_k \|\delta T^k\| < \infty$, then, T is a nonnegative supermartingale and

$$\|\delta T\| \le \sum_k c_k \|\delta T^k\|. \tag{4.8}$$

[Demo] To show that T is well defined (then T is obviously nonnegative supermartingale), we have only to show the finiteness of T for all time. T_0 is assumed to be finite. Suppose T_n is finite. Then,

$$T^k(svx) - T^k(s) \le M^k(s)x + V^k(s)(x^2 - v) \tag{4.9}$$

p84 and the RHS is finite, so is T^k_{n+1}. It is also easy to see that

$$\|\delta T\| \le \sum c_k \|\delta T^k\|. \tag{4.10}$$

Therefore, the increment is always bounded. □

Proof of Theorem 4.1
L4.5. If T is a nonnegative supermartingale, then T_n converges (i.e, has a finite limit) almost surely.
[Here, as usual, 'almost surely' means that we can construct a nonnegative supermartingale (*witnessing supermartingale*) T^* that diverges on all paths that fail to converge. That is, an unlikely event that actually happens can exit (can be constgructed).]
[Demo essentially following Doob] Choose two positive rational numbers $a < b$. We name the times at which T exits (a,b): σ_1 is the first up-exiting time. $\tau_1 > \sigma_1$ is the first down-exiting time after σ_1, etc. Formally, $\tau_0 = 0$ and

$$\sigma_k = \min\{i > \tau_{k-1} : T_i > b\}, \quad \tau_k = \min\{i > \sigma_k : T_i < a\}. \tag{4.11}$$

Let $\mathcal{P}$ be a minimal bounding strategy for T. We construct a strategy to gamble only between the a-downcrossing times and b-upcrossing times $(\tau_k, \sigma_k]$:

$$\mathcal{P}_n^{a,b} = \begin{cases} \mathcal{P}_n, \text{ if } \exists k : \tau_{k-1} < n \le \sigma_k, \\ (0,0), \text{ otherwise} \end{cases} \tag{4.12}$$

Let $T^{a,b}$ be a nonnegative martingale defined as $T_0 + K^{P^{a,b}}$.

$$\|\delta T^{a,b}\| \le \|\delta T\|. \tag{4.13}$$

(They agree or LHS = 0.) Also, if the T_n oscillation lasts forever, then

$$T_n^{a,b} \to \infty. \tag{4.14}$$

This is because, $\mathcal{P}^{a,b}$ is the strategy to increase T by $b-a$ for sure. Collecting all the rational intervals, make T^* as

$$T^* = \frac{1}{2}T + \sum_{k=1}^{\infty} 2^{-k-1} T^{a_k, b_k} \tag{4.15}$$

L4.4 says that T^* is a nonnegative supermartingale with

$$\|\delta T^*\| \le \|\delta T\|. \tag{4.16}$$

If T_n does not converge (if it grows unboundedly, Skeptic wins), there must be a minimum amplitude of oscillation of T_n. Therefore, T_n^* tends to infinity. □

Let T be a supermartingale, and A an increasing predictable process. Then $U = T + A$ is called *semimartigale*. A is called a *compensator* for U.

L4.6. Let U be a semimartingale with a compensator A. If $A_\infty < \infty$, then U_n converges.
[Demo] Set $T = U - A$, which is a supermartingale. Define a nonnegative supermartingale T^C ($C \in \boldsymbol{N}^*$) by $T_0^C = C$ and by the increment at time n

$$\delta T^C(n) = \begin{cases} \delta T(n), \text{ if } A(n) \le C, \\ 0, \text{ otherwise} \end{cases} \tag{4.17}$$

(That is, until the increasing $A(n)$ hits C, $T \ge 0$ irrespective of U, because it is nonnegative.) Following (4.15), define $(T^C)^*$. $\|\delta(T^C)^*\| \le \|\delta T^C\|$. Define

$$R = \sum 2^{-C}(T^C)^*. \tag{4.18}$$

p85 Since $\|\delta T^C\| \le \|\delta T\|$ (this is from the definition), $\|\delta R\| \le \|\delta T\|$. Then, R, which is a nonnegative supermartingale according to L4.4, witnesses the divergence of U. □

L4.7. Let S be a supermartingale, and S^2 be a semimartingale with compensator A. If $A_\infty < \infty$, then S_n converges almost surely.
[Demo] If A is a compensator of S^2, then it is a compensator of $(S+1)^2$ as well. [$(S+1)^2 - A = (S^2 - A) + 2S + 1$; $(S^2 - A)$ and S are supermartingale. Then, use L4.2.] Thanks to L4.6, S^2 and $(S+1)^2$ both converge almost surely. Since $2S = (S+1)^2 - S^2 - 1$, this implies the convergence of S. □

Completion of proof of Theorem 4.1
To conclude the proof that Skeptic has a winning strategy, we make a capital process

$$S_n = \sum^n \frac{x(k)}{k} \tag{4.19}$$

and an increasing deterministic process

$$A_n = \sum^n \frac{v(k)}{k^2}. \tag{4.20}$$

$S^2 - A$ is a capital process. This means (p82) that $(S^2-A)_n - (S^2-A)_{n-1}$ may be written as $Mx(n) + V(x(n)^2 - v(n))$, M and V depends on xi_{n-1}. This can be see explicitly by computation. Therefore, S^2 is a semimartingale with A as its compensator. Thus, L4.7 implies that if A converges, then S converges almost surely. Kronecker's lemma:[5]

$$\sum \frac{x(k)}{k} \text{ converges } \Rightarrow \frac{1}{n}\sum x(k) = 0 \tag{4.22}$$

concludes the proof. □
p87 The nonnegative supermartigale that witnesses the strong law of large numbers can be constructed explicitly (P4.2).

Reality's winning strategy
Kolmogorov devised a randomized strategy for Reality:

---

[5]

$$\begin{aligned}
\sum^m a_k/k &= (1/n)\sum^n a_k + \sum^{n-1}((1/k) - (1/n-1))a_k, \\
&= (1/n)\sum^n a_k + (1/(n-1) - 1/n)\sum^{n-1} a_k + \sum^{n-2}((1/k) - (1/n-2))a_k,
\end{aligned} \tag{4.21}$$

etc.

if $v(n) < n^2$,

$$x(n) = \begin{pmatrix} n \\ -n \\ 0 \end{pmatrix} \text{ with probability } \begin{pmatrix} v(n)/(2n^2) \\ v(n)/(2n^2) \\ 1 - v(n)/n^2 \end{pmatrix}, \tag{4.23}$$

if $v(n) \geq n^2$,

$$x(n) = \begin{pmatrix} \sqrt{v(n)} \\ -\sqrt{v(n)} \end{pmatrix} \text{ with probability } \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}. \tag{4.24}$$

But the argument is measure theoretical.

Martingale strong law

5 Martin's Theorem

Consider a perfect information game with Player I and II, who alternate moves.

p95 **Finite horizon game**
Zermelo proved that if a game has a finite horizon, one of the players has a winning strategy. This can be proved by a backward induction.

General case
A game path may be expressed as

$$a_1 b_1 a_2 b_2 \cdots, \tag{5.1}$$

where a_i is the moves of I, and b_i that of II. Generally, we may write a path as c_i, and interpret that for odd i the moves are by I, etc.
p96 Let Γ be the totality of paths. Let $E \subset \Gamma$. $\mathcal{G}(E)$ denotes the games that Player I has wins, if a path is in E; otherwise II wins. If I or II has a winning strategy, we say $\mathcal{G}(E)$ is *determined*.

The *quasi-Borel* subsets of Γ form the smallest class of subsets of Γ that contains all open sets and is closed under
(1) complementation
(2) countable union
(3) Open-separated union = difference.

(That is, roughly if a set is written as a countable union of open sets and its comlements, then the set is quasi-Borel.)
Martin's Theorem 1990. If $E \subset \Gamma$ is quasi-Borel, then $\mathcal{G}(E)$ is determined.

Axiom of determinacy [Mycielski and Steinhaus 1962]
Every perfect information game $\mathcal{G}(E)$ on $\boldsymbol{N}^*$ is determined.

p98 AD + DC (= dependent choice) is more regular in a certain ways than AC:
Mycielski-Świerczkowski: Any set of real numbers is Lebesgue measurable.

Davis: Any uncountable set of real numbers contains a perfect set.

This implies the continuum hypothesis, because a perfect set has a cardinality of the continuum.

6 Law of Iterated Logarithm

p99

Khinchin's law claims that the convergence to the limit dictated by the LLN is almost surely oscillatory with maximal deviations on both sides asymptotically close to $\sqrt{\log\log n}/\sqrt{2n}$: for the fair coin tossing process the number of heads y_n up to the nth trial obeys

$$\limsup_{n\to\infty} \frac{y_n/n - 1/2}{\sqrt{\frac{\log\log n}{2n}}} = 1, \quad \liminf_{n\to\infty} \frac{y_n/n - 1/2}{\sqrt{\frac{\log\log n}{2n}}} = -1. \tag{6.1}$$

p122 **Predictable unbounded forecasting game**
Protocol:
$K_0 = 1$,
FOR $n = 1, 2, \cdots$
  Forecaster announces $m(n) \in \boldsymbol{R}$, $c(n) \geq 0$, and $v(n) \geq 0$.
  Skeptic announces $M(n) \in \boldsymbol{R}$ and $V(n) \in \boldsymbol{R}$.
  Reality announces $x(n)$ such that $|x(n) - m(n)| \leq c(n)$.
  The capital process is computed by

$$K_n = K_{n-1} + M_n(x(n) - m(n)) + V_n[(x(n) - m(n))^2 - v(n)]. \tag{6.2}$$

The winning rule is the same as before.

Let

$$A_n = \sum_{k=1}^{n} v_k. \tag{6.3}$$

Theorem 5.1 In the *predictable* unbounded forecasting game, if $A_n \to \infty$ and

$$c_n = o\left[\sqrt{\frac{A_n}{\log\log A_n}}\right], \tag{6.4}$$

then Skeptic can force

$$\limsup_{n\to\infty} \frac{\sum_{k=1}^{n}(x(k) - m(k))}{\sqrt{\log\log A_n}} = 1. \tag{6.5}$$

□

The ordinary unbounded forecasting game is not strong enough to claim the equality above (Skeptic's advantage is not enough):

Theorem 5.2 In the unbounded forecasting game, if $A_n \to \infty$ and

$$|x(n) - m(n)| = o\left[\sqrt{\frac{A_n}{\log\log A_n}}\right], \tag{6.6}$$

then Skeptic can force

$$\limsup_{n\to\infty} \frac{\sum_{k=1}^{n}(x(k) - m(k))}{\sqrt{\log\log A_n}} \le 1. \tag{6.7}$$

□

7 Weak Laws

The game theoretic version of weak laws reads P replaced by $\overline{P}$ in the ordinary measure-theoretical framework.

p124 **Finite horizon fair coin game**

Reality's move is H or T or ± 1. We make a game up to time N (finite horizon).

Protocol:

$K_0 = \alpha \ (> 0)$

FOR $n = 1, 2, \cdots, N$

  Skeptic announces $M(n) \in \boldsymbol{R}$

  Reality announces $x(n) \in \{-1, 1\}$.

$K_n = K_{n-1} + M(n)x(n)$.
If K_n is never negative and either $K_N \geq 1$ or $|S_N/N| < \epsilon$ (> 0), then Skeptic wins, where

$$S_n = \sum_{k}^{n} x(k). \tag{7.1}$$

Bernouilli without probability
Bernoulli's theorem says if N is sufficiently large Skeptic has a winning strategy.

We must make a nonnegative martingale that witnesses the weak law.
L6.1. L_n defined as follows is a nonnegative martingale with $L_0 = 1$:

$$L_n = \frac{S_n^2 + N - n}{N}. \tag{7.2}$$

[Demo] By a simple calculation the increment reads

$$\delta(S_n^2 - n) = 2S_{n-1}x(n) + x^2(n) - 1 = 2S_{n-1}x(n) \tag{7.3}$$

Therefore, $S_n^2 - n$ is a capital process; martingale. Therefore, L_n is also martingale. Since $S_n^2 \geq 0$ obviously $L_n \geq 0$.□

P6.1 [Bernouilli's theorem]. Skeptic has a winning strategy in the finite horizon fair coin game, if $N \geq 1/\alpha\epsilon^2$.
[Demo] Let us choose the strategy so that $M_n = \alpha S_{n-1}$:

$$K_n = K_{n-1} + 2\alpha S_{n-1} x_n. \tag{7.4}$$

Then, according to L6.1 $K_n = \alpha L_n$. Therefore, $K_N = \alpha S_N^2/N$. If this is 1 or more, Skeptic wins. If not, then since $1/N < \alpha\epsilon^2$

$$\alpha S_N^2/N < 1, \Rightarrow \ S_N^2/N^2 < \epsilon^2. \tag{7.5}$$

And Skeptic still wins. □

Bernouilli's theorem with game-theoretical probability
p126 The upper probability of an event E measures the degree to which a strategy for betting on E can multiply one's capital without risk of bankruptcy:

$$\begin{aligned} \overline{P}E \ = \ & \{\alpha \,|\, \text{ there is a strategy that begins with } \alpha \\ & \text{and ends up with at least 1 if } E \text{ happens and at least 0 otherwise}\}. \end{aligned} \tag{7.6}$$

According to P6.1 Skeptic has a strategy to increase α to be 1 if the path does not satisfy (7.5). That is, if $\alpha > 1/N\epsilon^2$, then (7.5) does not happen:

$$\overline{P}(|S_N/N| \geq \epsilon) < 1/N\epsilon^2. \tag{7.7}$$

Since $\underline{P} \leq \overline{P}$, from this we have

$$\underline{P}(|S_N/N| \geq \epsilon) < 1/N\epsilon^2. \tag{7.8}$$

Martingale (measure theoretical)

[I] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes* nd Ed. (North-Holland, 1989)
[P] P. Protter, *Stochastic Integration and Differential Equations, a new approach* (Springer, 1990).

Martingale
Let $X = (X_t)_{t \in T}$ be a stochastic process. It is called a *martingale* (resp., supermartingale, submartingale) wrt $(\mathcal{F}_t)$, if
(i) X_t is integrable for each $t \in T$,
(ii) X_t is $\mathcal{F}_t$ adapted,
(iii) $E(X_t \,|\, \mathcal{F}_s) = X_s$ (resp., $\leq$, $\geq$) for $s < t$.

The sample path of a continuous martingale is of a.e. unbounded variation unless it is constant everywhere [P64] (due to Fisk).

Let X be martingale, and f a deterministic process. Then, $Y = f \cdot X$ defined by

$$\begin{aligned} Y_0 &= X_0, & (7.9) \\ Y_n &= Y_{n-1} + f_n \cdot (X_n - X_{n-1}) & (7.10) \end{aligned}$$

is also martingale. This is the definition of the *stochastic integral* wrt the martingale [I55, P50].
Optional stopping. If $\tau(\omega)$, $\sigma(\omega)$ are stopping times. If $\sigma \leq \tau$, then $E(X_\tau \,|\, \mathcal{F}_\sigma) = X_\sigma$ [I26].

Submartingale inequality [Doob]. Let X be submartingale. Then for any $\lambda > 0$

$$\lambda P(\max_{0 \leq n \leq N} X_n \geq \lambda) \leq E(X_N \,|\, \max_{0 \leq n \leq N} X_n \geq \lambda) \leq E(X_N^+) \leq E(|X_N|), \tag{7.11}$$

and

$$\lambda P(\min_{0\le n\le N} X_n \le -\lambda) \le -E(X_0)+E(X_N \mid \min_{0\le n\le N} X_n \le -\lambda) \le E(|X_0|)+E(|X_N|). \tag{7.12}$$

[Demo] Let n be the time that $X \ge \lambda$ occurs for the first time. This is a stopping time. Therefore, $E(X_N|\mathcal{F}_n) \le X_n$ or $E(X_N) \le E(X_n)$.

$$\begin{aligned} E(X_n) &= E(X_n\,;\, \max_{0\le n\le N} X_n \ge \lambda) + E(X_n\,;\, \max_{0\le n\le N} X_n < \lambda), && (7.13)\\ &= E(X_n\,;\, \max_{0\le n\le N} X_n \ge \lambda) + E(X_N\,;\, \max_{0\le n\le N} X_n < \lambda), && (7.14)\\ &\ge \lambda P(\max_{0\le n\le N} X_n \ge \lambda) + E(X_N\,;\, \max_{0\le n\le N} X_n < \lambda). && (7.15)\end{aligned}$$

Here $E(A\,;\,B)$ is the average of A on B (i.e., $E(A\chi_B)$). $\square$

(generalized) Doob-Kolmogorov's inequality
Let X be martingale. Then $|X|^p$ is a submartingale. If $E(|X-n|^p) < \infty$, then for $p \ge 1$

$$P(\max_{0\le n\le N} X_n \ge \lambda) \le E(|X_N|^p)/\lambda^p. \tag{7.16}$$

If $p > 1$,

$$E(\max_{0\le n\le N} |X_n|^p) \le [p/(1-p)]^p E(|X_N|^p). \tag{7.17}$$

(7.16) immediately follows from the inequality above. $\square$

If X is submartingale such that $\sup_n E(X_n^+) < \infty$, then $X_\infty = \lim_{n\to\infty}$ exists almost surely and is integrable.$\square$ [I30]

Doob-Meyer decomposition of submartingale. Let X be submartingale (with some integrability condition). Then, $X = M + A$, where M is martingale, and A is predictable and increasing. The decomposition is unique [I34] ([Funaki]).
[Demo] A is defined by

$$\begin{aligned} A_0 &= 0, && (7.18)\\ A_n &= A_{n-1} + E(X_n - X_{n-1} \mid \mathcal{F}_{n-1}). && (7.19)\end{aligned}$$

$M_n = X_n - A_n$ is a martingale. If we have two decompositions, then $M_n - M_n'$ is deterministic, so

$$M_n - M_n' = E(M_n - M_n' \mid \mathcal{F}_{n-1}) = M_{n-1} - M_{n-1}' = \cdots = X(0) - X(0) = 0. \tag{7.20}$$

That is, the decomposition is unique. $\square$

There are continuous time versions.

Semimartingale.

Roughly put, if $I_X(f) : f \to f \cdot X$ is continuous (in probability) for any predictable process,[6] X[7] is called a *semimartingale*.

If X is of finite total variation, then X is a semimartingale.

A square integrable martingale is a semimartingale.

This implies

(i) A locally square integrable cadlag martingale[8] is a semimartingale.

(ii) A local semimartingale with continuous paths is a semimartingale.

(iii) The Wiener process is a semimartingale.

Let A and M be $\mathcal{F}_t$-adapted stochastic processes; A be of bounded variation and right-continuous, and M be (square integrable) martingale. Then

$$X(t) = X(0) + M(t) + A(t) \tag{7.21}$$

is called a *decomposable process* [P48]. If X is cadlag, then decomposability = being a semimartingale [P88].

A decomposable process is a semimartingale.

A Lévy process is a semimartingale.

Stochastic integration preserves the property of being a semimartingale [P55].

[6] [P43], this means that, roughly, when a change occurs, it is measurable with respect to the event just up to the jump.

[7] X must be 'cadlag' (right continuous [i.e., $\lim_{x\downarrow x_0} f(x) = f(x_0)$] and with left limits.

[8] local martingale is enough [P33].