SIPTA Online School 2020, University of Liverpool Institute for Risk and Uncertainty

Game-theoretic foundations for statistical testing and imprecise probabilities

Remote lectures by **Glenn Shafer**. December 9th and 10th, 2020.

Lecture 3. Basing probability theory on betting.

Reading: Bernoulli's and De Moivre's theorems, Chapter 2 of *Game-Theoretic Foundations for Probability and Finance*, by Glenn Shafer and Vladimir Vovk, Wiley, 2019.

Slides revised February 7, 2021

1. Game-theoretic probability in the 17^{th} century
2. Game-theoretic events, variables, and martingales
3. Game-theoretic expected value and probability
4. For fun: History of martingales

1. Game-theoretic probability in the 17th century

The thing that hath been, it is that which shall be;
and that which is done is that which shall be done:
and there is no new thing under the sun.

Ecclesiastes 1:9

In 1754, Christiaan Huygens explained how to find the "value of an expectation".

What these words meant:

expectation = payoff = random variable

value = expected value = what it's worth

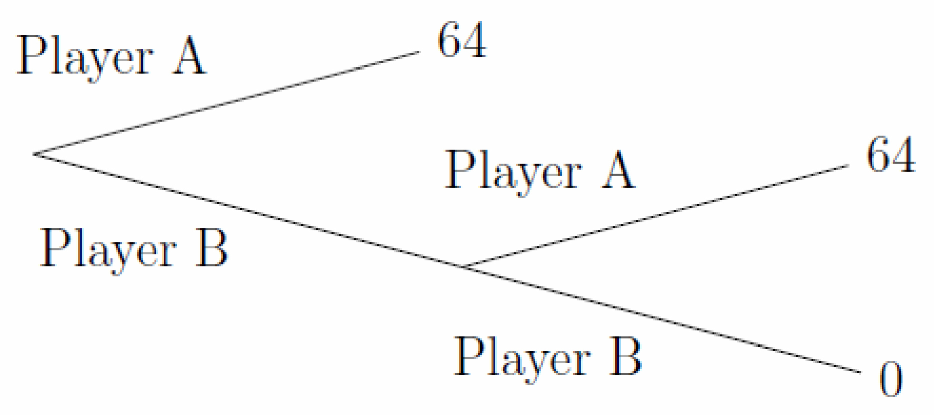
ps. Huygens did not use the word "probability", but "probability" is a special case of "expected value".

Huygens was inspired by Blaise Pascal’s 1654 discussion with Pierre Fermat.

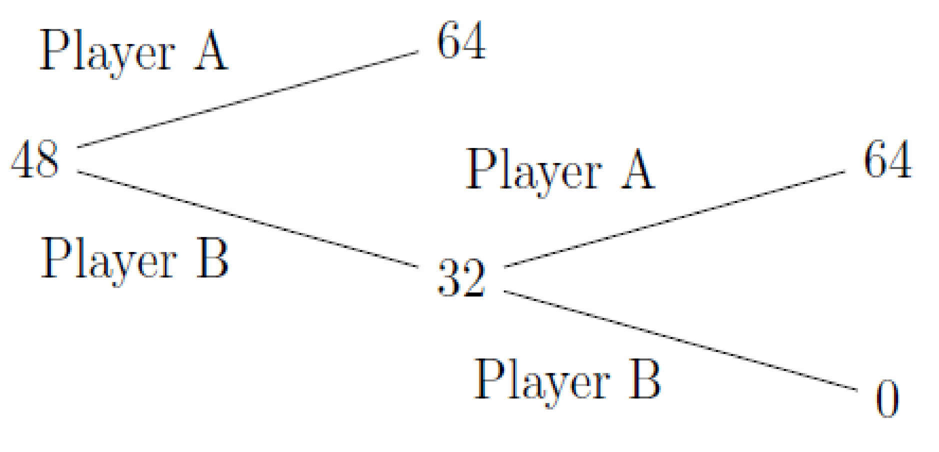
Blaise Pascal
1623-1662

Pierre Fermat
1607-1665

Pascal’s division problem



Pascal’s solution



Pascal's letters to Fermat.

- Written in 1654.
- Published in 1679.

Pascal's *Triangle arithmétique*.

- Written at the end of 1654.
- Published in 1665.
- Rare.

Blaise Pascal
1623 - 1662

Huygens's *De Ratiociniis in ludo aleae*.

- Inspired by 1655 visit to Paris.
- Drafted 1656.
- Published 1657.
- Widely distributed and translated.

Christiaan Huygens
1629 - 1695

Having learned about the Pascal-Fermat correspondence from mathematicians in Paris, Huygens saw an opportunity to use Descartes's algebra.

René Descartes
1596 - 1650

Use equations to express conditions on a number x.

- *Analysis:* Find what x must be if there is a number satisfying the conditions.
- *Synthesis:* Prove that the number found does satisfy the conditions.

Proposition I. If I have the same chance to get a or b it is worth as much to me as $(a + b)/2$.

Consider this fair game:

- We both stake x.
- The winner will give a to the loser.

The analysis:

- If I win, I get $2x - a$.
- If this is equal to b, then $x = (a + b)/2$.

The synthesis:

- Having $(a + b)/2$, I can play with an opponent who stakes the same amount, on the understanding that the winner gives the loser a.
- This gives me equal chances of getting a or b.

Pascal proved the same thing using his two principles and the assumption that the game is one of *pure chance*.

Huygens's argument is purely game-theoretic:

- Does not require that the game be one of pure chance.
- Replaces principle of equal division by willingness of players to play on even terms.

Proposition I. If I have the same chance to get a or b it is worth as much to me as $(a + b)/2$.

Consider this fair game:

- We both stake x.
- The winner will give a to the loser.

The analysis:

- If I win, I get $2x - a$.
- If this is equal to b, then $x = (a + b)/2$.

The synthesis:

- Having $(a + b)/2$, I can play with an opponent who stakes the same amount, on the understanding that the winner gives the loser a.
- This gives me equal chances of getting a or b.

Hans Freudenthal
1905 – 1990

Emphasized the game-theoretic aspect of Huygens's argument.

Proposition III. If I have p chances for a and q chances for b, this is worth $(pa + qb)/(p + q)$.

Steve Stigler, born 1941, called Huygens "the father of the hedge".

Synthetic (constructive) proof:

- Assign each chance to a different player.
- I am one of the $p + q$ players.
- Each of us puts up $(pa + qb)/(p + q)$.
- Winner takes all.

- I make side bet with q opponents; winner gives loser b.
- I make side bet with other $p - 1$ opponents; winner gives loser a.

- This gives me p chances for a and q chances for b.

$$(p + q)\,(pa + qb)/(p + q) - qb - (p - 1)a = a$$

Proposition III. If I have p chances for a and q chances for b, this is worth $(pa + qb)/(p + q)$.

Synthetic (constructive) proof:

- Assign each chance to a different player.
- I am one of the $p + q$ players.
- Each of us puts up $(pa + qb)/(p + q)$.
- Winner takes all.

- I make side bet with q opponents; winner gives loser b.
- I make side bet with other $p – 1$ opponents; winner gives loser a.

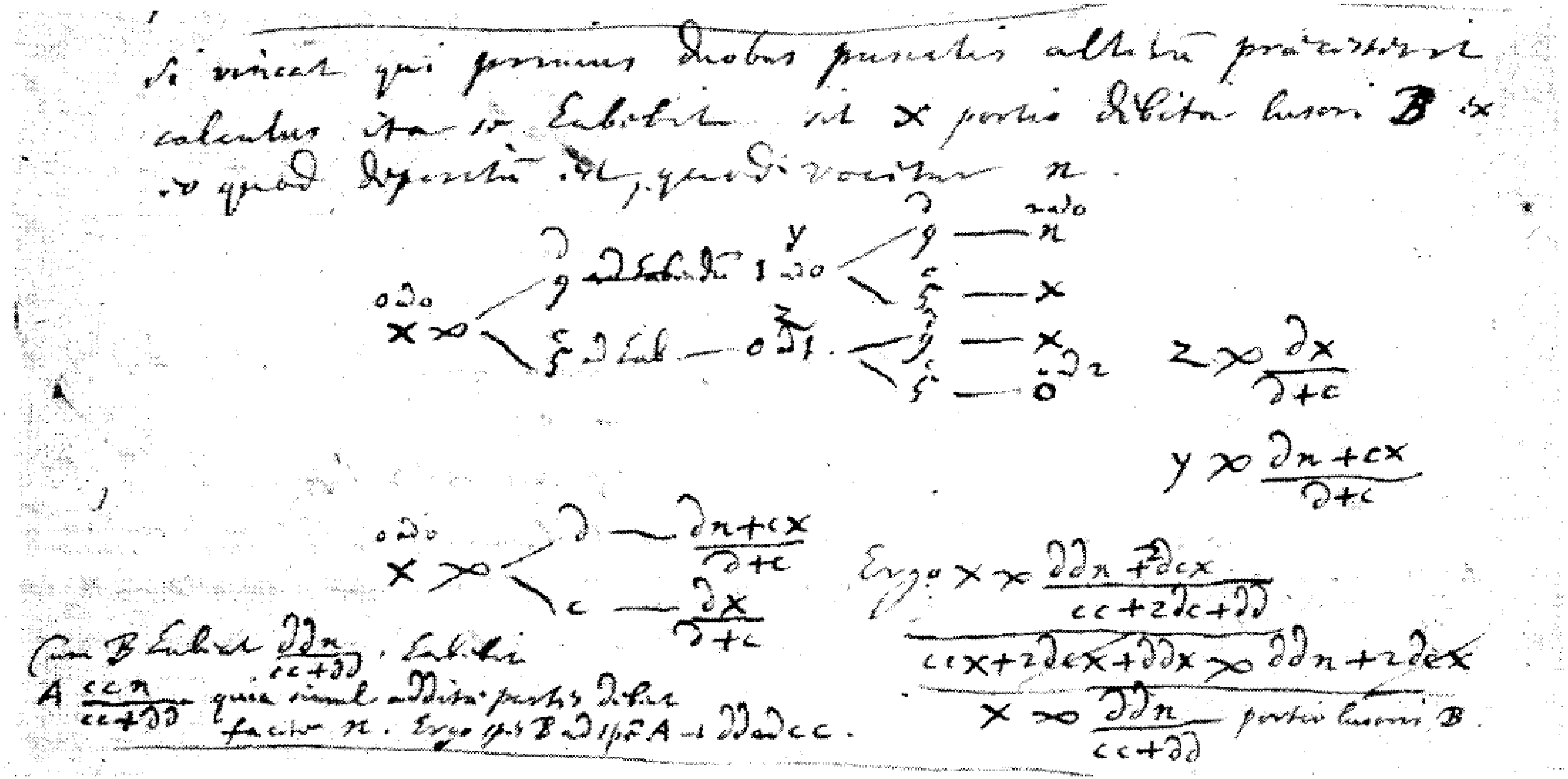
- This gives me p chances for a and q chances for b.

$$(p + q)\,(pa + qb)/(p + q) – qb – (p – 1)a = a$$

Ivo Schneider
Born 1938

Ivo's objection to Huygens:
Not all players are treated the same if one gets to decide what side bets to make.

The first known game tree, in a 1676 manuscript in which Huygens solves one of the exercises at the end of his book.



Huygens's game-theoretic definition of probability

$\mathbf{E}(X)$ = amount you need at beginning to get X at end.

Corollary:

$\mathbf{P}(A)$ = amount you need at beginning to get **1 if *A* happens, 0 otherwise** at end.

X is a random variable.
A is an event.

2. Game-theoretic events, variables, and martingales

To make a name for learning
when other roads are barred,
take something very easy
and make it very hard.

Piet Hein, 1905–1996

I will use these two protocols to explain the language of game-theoretic probability.

Testing a probability p

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
Skeptic announces $M_n \in \mathbb{R}$.
Reality announces $y_n \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p)$.

Probability forecasting

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
Forecaster announces $p_n \in [0, 1]$.
Skeptic announces $M_n \in \mathbb{R}$.
Reality announces $y_n \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p_n)$.

- Requirement that Skeptic keep capital non-negative not built in.
  But he must do so in order to discredit *p* or Forecaster.
- Protocol goes on forever, but Skeptic can stop betting whenever.

Concepts:

- path
- sample space
- event
- variable
- situation
- process
- supermartingale
- martingale
- upper/lower expected value
- upper/lower probability

Testing a probability p

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
  Skeptic announces $M_n \in \mathbb{R}$.
  Reality announces $y_n \in \{0, 1\}$.
  $\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p)$.

Probability forecasting

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
  Forecaster announces $p_n \in [0, 1]$.
  Skeptic announces $M_n \in \mathbb{R}$.
  Reality announces $y_n \in \{0, 1\}$.
  $\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p_n)$.

Testing a probability p

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
Skeptic announces $M_n \in \mathbb{R}$.
Reality announces $y_n \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p)$.

To understand goals Skeptic can achieve…

….consider how his opponent (Reality) might move.

- *Path* = sequence $y_1 y_2 \ldots = \omega$
- *Sample space* = set of all paths = Ω
- Here $\Omega = \{0, 1\}^\infty$.

Testing a probability p

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
Skeptic announces $M_n \in \mathbb{R}$.
Reality announces $y_n \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p)$.

- *Path* $=$ sequence $y_1 y_2 \ldots = \omega$
- *Sample space* $=$ set of all paths $= \Omega$
- Here $\Omega = \{0,1\}^\infty$.

- *Event* $=$ subset of Ω
- *Variable* $=$ real-valued function on Ω
- *Situation* $=$ initial segment of a path, say $s = y_1 \ldots y_n$
- $\Box$ $=$ initial (empty) situation
- *Process* $=$ real-valued function on situations
- *Supermartingale* $=$ capital process for a strategy for Skeptic that uses only information in the protocol

FINITE HORIZON VERSION

Testing a probability p

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, \ldots, N$:
Skeptic announces $M_n \in \mathbb{R}$.
Reality announces $y_n \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p)$.

- *Path* $=$ sequence $y_1 \ldots y_N = \omega$
- *Sample space* $=$ set of all paths $= \Omega$
- Here $\Omega = \{0, 1\}^N$.

These definitions are unchanged

- *Event* $=$ subset of Ω
- *Variable* $=$ real-valued function on Ω
- *Situation* $=$ initial segment of a path, say $s = y_1 \ldots y_n$
- $\Box$ $=$ initial (empty) situation
- *Process* $=$ real-valued function on situations
- *Supermartingale* $=$ capital process for a strategy for Skeptic that uses only information in the protocol

$\mathcal{Martingale}$ = supermartingale $\mathcal{T}$ such that $-\mathcal{T}$ is supermartingale.

Supermartingales that are not martingales arise whenever the protocol

- offers bets on only one side,
- requires Skeptic to keep his capital non-negative, or
- allows Skeptic to bet suboptimally.

Testing the probability $\frac{1}{2}$ for $y = 1$

Skeptic announces M such that $0 \leq M \leq 2$.
Reality announces $y \in \{0, 1\}$.
$\mathcal{K} := 1 + M(y - \frac{1}{2})$.

one-sided

Testing a forecaster over time

$\mathcal{K}_0 := 1$.
FOR $n = 1, 2, \ldots, N$:
Forecaster announces a probability distribution P_n on $\mathcal{Y}$.
Skeptic announces $S_n : \mathcal{Y} \to [0, \infty)$ such that $\mathbf{E}_{P_n}(S_n) \leq \mathcal{K}_{n-1}$.
Reality announces $y_n \in \mathcal{Y}$.
$\mathcal{K}_n := S_n(y_n)$.

suboptimal

Testing a probability $\frac{1}{2}$

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
Skeptic announces $M_n \in \mathbb{R}$.
Reality announces $y_n \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - \frac{1}{2})$.

All supermartingales in this protocol are martingales.

A classical martingale with $p = \frac{1}{2}$.

Most popular betting system in 19th-century casinos: the d'Alembert:

- First bet 1.
- When you lose, increase bet by 1.
- When you win, decrease bet by 1, unless it is already 1.
- Stop when k ahead or out of money.

This betting system is a strategy for Skeptic in the protocol. What is its martingale?

--$10 in pocket.

--Bet $1.

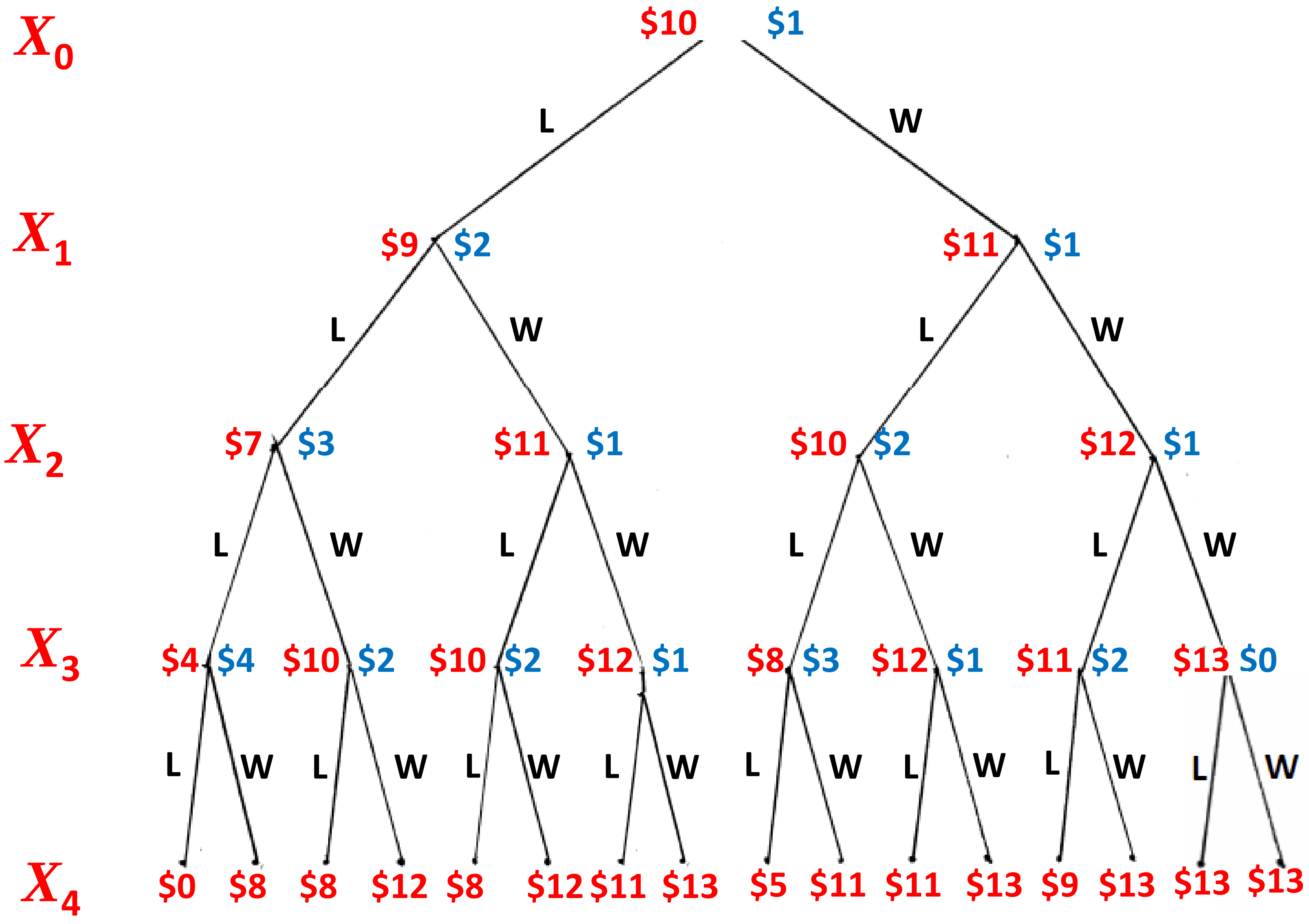
--Lose, increase bet by $1.

--Win, decrease bet by $1, unless already = $1.

--Stop when you gain $3 or go broke.

Current capital in red.
Amount bet in blue.

$X_0, X_1, X_2, X_3, X_4\ldots$ is a martingale.



Originally <u>betting systems</u> (strategies for Skeptic) were called *martingales*.
In 1939, Jean Ville shifted to calling the resulting <u>capital processes</u> martingales.

Stop after 3 rounds.

Testing a probability $\frac{1}{2}$

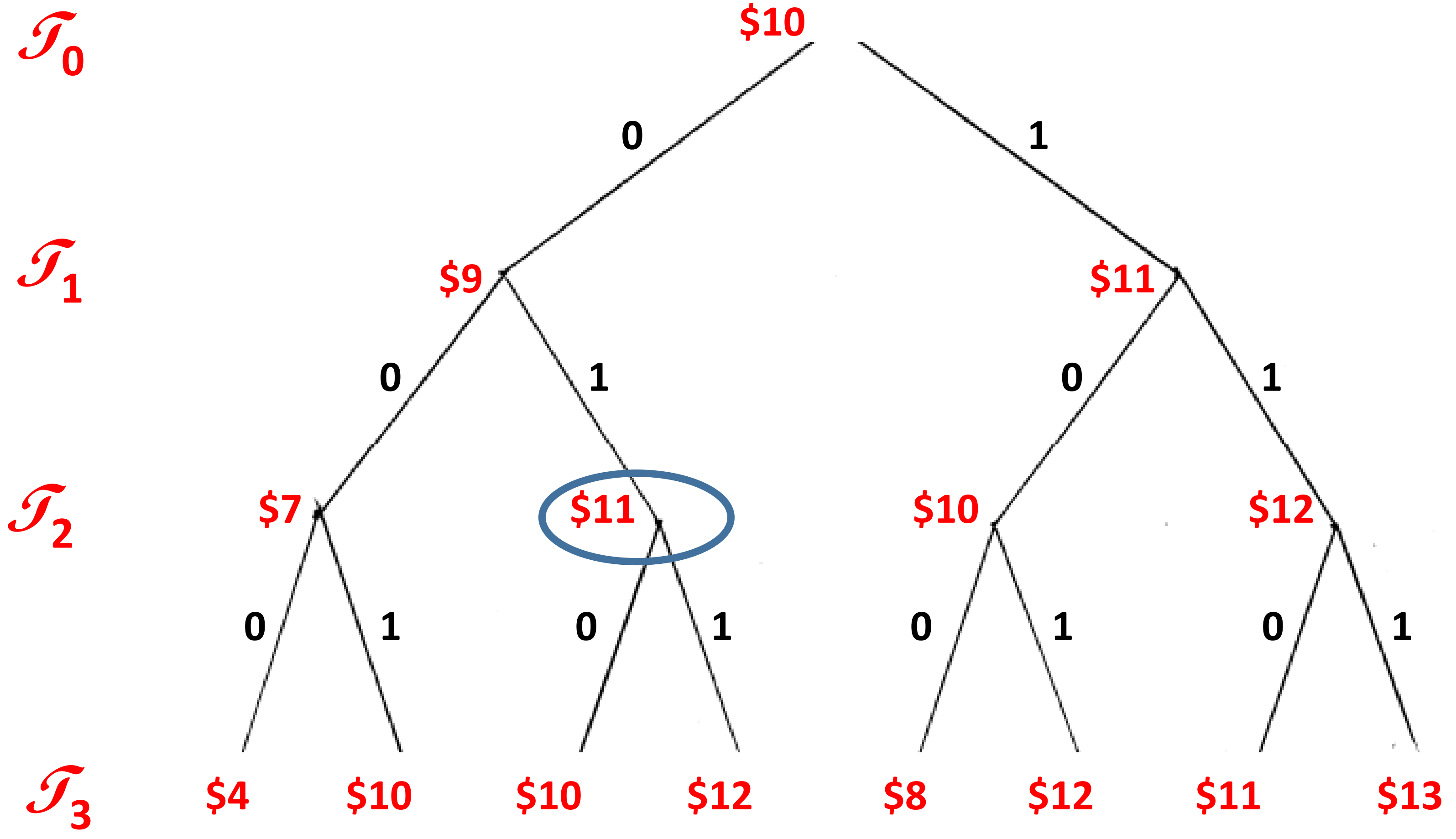
Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.

FOR $n = 1, 2, 3$:

Skeptic announces $M_n \in \mathbb{R}$.

Reality announces $y_n \in \{0, 1\}$.

$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - \frac{1}{2})$.



Think of the martingale as a function $\mathscr{I}$ on the situations. For example,

$\mathscr{I}(01)=11$.

Or think of it as a sequence $\mathscr{I}_0, \mathscr{I}_1, \mathscr{I}_2, \mathscr{I}_3$ of random variables. For example,

$\mathscr{I}_2(010) = \mathscr{I}_2(011)=11$.

Testing a probability p

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \dots$:
Skeptic announces $M_n \in \mathbb{R}$.
Reality announces $y_n \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p)$.

In this protocol, game-theoretic probability resembles standard probability.

Probability forecasting

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \dots$:
Forecaster announces $p_n \in [0, 1]$.
Skeptic announces $M_n \in \mathbb{R}$.
Reality announces $y_n \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p_n)$.

In this protocol, game-theoretic probability already becomes “imprecise”.

In both protocols, all supermartingales are martingales.

Probability forecasting

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
Forecaster announces $p_n \in [0, 1]$.
Skeptic announces $M_n \in \mathbb{R}$.
Reality announces $y_n \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p_n)$.

Here Skeptic has two opponents, Reality & Forecaster.

To understand goals Skeptic can achieve…

….consider how the two might move.

- *Path* = sequence $p_1 y_1 p_2 y_2 \ldots = \omega$
- *Sample space* = set of all paths = Ω
- Here $\Omega = ([0, 1] \times \{0, 1\})^\infty$.

Probability forecasting

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
Forecaster announces $p_n \in [0, 1]$.
Skeptic announces $M_n \in \mathbb{R}$.
Reality announces $y_n \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p_n)$.

- *Path* = sequence $p_1 y_1 p_2 y_2 \ldots = \omega$
- *Sample space* = set of all paths = Ω
- Here $\Omega = ([0,1] \times \{0,1\})^\infty$.

All supermartingales in this protocol are martingales.

- *Event* = subset of Ω
- *Variable* = real-valued function on Ω
- *Betting situation* = initial segment of the form $s = p_1 y_1 \ldots p_{n-1} y_{n-1} p_n$.
- *Settlement situation* = initial segment of the form $s = p_1 y_1 \ldots p_n y_n$.
- *Supermartingale* = capital process for a strategy for Skeptic that uses only information in the protocol; this is a real-valued function on the settlement situations.

Definition of *martingale* given in textbooks:

- A sequence of random variables $X_0, X_1, \ldots$ is a *stochastic process*.

- A stochastic process $X_0, X_1, \ldots$ is a *martingale* if for all n,
$$\mathbf{E}(X_n | X_0, X_1, \ldots, X_{n-1}) = X_{n-1}.$$

This is another way of saying that a martingale is the capital process for a betting strategy using fair bets. Because it says that the expected gain at each time is zero:
$$\mathbf{E}(X_n - X_{n-1} | X_0, X_1, \ldots, X_{n-1}) = 0.$$

But often students of advanced probability study this topic without being told about the betting interpretation.

3. Game-theoretic expected value and probability

It would all come out in good time, I observed,
and in the meantime nothing was to be said,
save that I had come into great expectations
from a mysterious patron.

Charles Dickens, *Great Expectations*, 1861

For simplicity, begin with the protocol that tests a probability p and ends in N rounds.

Testing a probability p

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, \ldots, N$:
  Skeptic announces $M_n \in \mathbb{R}$.
  Reality announces $y_n \in \{0, 1\}$.
  $\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p)$.

Here Huygen's definition applies more or less directly.

From earlier slide

Testing a probability p

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.

FOR $n = 1, \ldots, N$:

Skeptic announces $M_n \in \mathbb{R}$.

Reality announces $y_n \in \{0, 1\}$.

$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p)$.

- A sequence $y_1 \ldots y_N$ is a *path*.
- Set of all possible paths is *sample space*.
- Write Ω for the sample space.
- Write $\omega = y_1 \ldots y_N$.
- $\Omega = \{0, 1\}^\infty$.

- An *event* is a subset of Ω.
- A *variable* is a real-valued function on Ω.
- A *situation* is an initial segment of a path, say $s = y_1 \ldots y_n$.
- Write $\Box$ for the initial (empty) situation.
- A *process* is a real-valued function on the set of situations.
- A *martingale* is the capital process resulting from a strategy for Skeptic.

(All supermartingales are martingales.)

Write $\mathbf{T}$ for the set of all martingales in the protocol.

Definitions. The *expected value* of a variable X is

$$\mathbb{E}(X) := \inf\{\mathcal{T}_0 \mid \mathcal{T} \in \mathbf{T} \text{ and } \mathcal{T}_N \geq X\}.$$

The *probability* of an event A, for which we write $\mathbb{P}(A)$, is $\mathbb{E}(\mathbf{1}_A)$, where

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

Write $\mathbf{T}$ for the set of all martingales in the protocol.

Testing a probability p

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, \ldots, N$:
Skeptic announces $M_n \in \mathbb{R}$.
Reality announces $y_n \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p)$.

Definitions. The *expected value* of a variable X is

$$\mathbb{E}(X) := \inf\{\mathcal{T}_0 \mid \mathcal{T} \in \mathbf{T} \text{ and } \mathcal{T}_N \geq X\}.$$

The *probability* of an event A, for which we write $\mathbb{P}(A)$, is $\mathbb{E}(\mathbf{1}_A)$, where

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

This rigorous definition agrees with Huygens's intuitive definition in this protocol. Because there is always a martingale $\mathcal{T}$ such that $\mathcal{T}_N = X$ exactly, and its initial value $\mathcal{T}_0$ is equal to $\mathbb{E}(X)$.

Write $\mathbf{T}$ for the set of all martingales in the protocol.

Testing a probability p

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, \ldots, N$:
Skeptic announces $M_n \in \mathbb{R}$.
Reality announces $y_n \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p)$.

Definitions. The *expected value* of a variable X is

$$\mathbb{E}(X) := \inf\{\mathcal{T}_0 \mid \mathcal{T} \in \mathbf{T} \text{ and } \mathcal{T}_N \geq X\}.$$

The *probability* of an event A, for which we write $\mathbb{P}(A)$, is $\mathbb{E}(\mathbf{1}_A)$, where

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

Theorem. If ω has k 1s and $N - k$ 0s, then

$$\mathbb{P}(\{\omega\}) = p^k(1-p)^{N-k}.$$

Can you prove it?

So here game-theoretic probability is the same as standard probability.

From page 34 of Game-Theoretic Foundations

Proposition 2.5. *Suppose X is a variable in Protocol 1.10 that is determined in the first N rounds. Then*

$$\mathbb{E}(X) = \sum_{t \in \{0,1\}^N} p^{\#t}(1-p)^{N-\#t} X(t), \tag{2.11}$$

where $\#t$ is the number of 1s in t.

Proof. Consider the process $\mathcal{M}$ defined by $\mathcal{M}_N := X$ and then by backward recursion for $n = N-1, \dots, 0$:

$$\mathcal{M}(y_1 \dots y_n) := p\mathcal{M}(y_1 \dots y_n 1) + (1-p)\mathcal{M}(y_1 \dots y_n 0). \tag{2.12}$$

It follows from (2.12) that

$$\mathcal{M}(y_1 \dots y_n) = \mathcal{M}(y_1 \dots y_{n-1}) + L_n(y_n - p)$$

for $n = 1, \dots, N$, where

$$L_n := \mathcal{M}(y_1 \dots y_{n-1} 1) - \mathcal{M}(y_1 \dots y_{n-1} 0). \tag{2.13}$$

Thus $\mathcal{M}$ is the capital process for the strategy that sets $\mathcal{K}_0 := \mathcal{M}_0$ and makes the move L_n defined by (2.13) on round n. It is clear that $-\mathcal{M}$ is the capital process for the strategy that sets $\mathcal{K}_0 := -\mathcal{M}_0$ and makes the move $-L_n$ on round n. So $\mathcal{M}$ is a martingale.

If also follows from (2.12), by backward induction, that

$$\mathcal{M}(y_1 \dots y_n) = \sum_{t \in \{0,1\}^{N-n}} p^{\#t}(1-p)^{N-n-\#t} X(y_1 \dots y_n t) \tag{2.14}$$

for $n = N-1, \dots, 0$. When $n = 0$, (2.14) reduces to the right-hand side of (2.11). This is the value of $\mathcal{M}_0$, and by Lemma 2.3, it is equal to $\mathbb{E}(X)$. ■

Lemma 2.3. *If $\mathcal{M}$ is a bounded martingale, then $\mathbb{E}(\mathcal{M}_N) = \mathcal{M}_0$.*

$$\overline{\mathbb{E}}(X) := \inf\{\mathcal{T}_0 \mid \mathcal{T} \in \mathbf{T} \text{ and } \mathcal{T}_N \geq X\}$$

Lemma 2.2. *Global upper expectation has the following properties*

1. $\overline{\mathbb{E}}(X_1 + X_2) \leq \overline{\mathbb{E}}(X_1) + \overline{\mathbb{E}}(X_2)$.
2. *If* $c \in \mathbb{R}$, *then* $\overline{\mathbb{E}}(X + c) = \overline{\mathbb{E}}(X) + c$.
3. *If* $c \geq 0$, *then* $\overline{\mathbb{E}}(cX) = c\overline{\mathbb{E}}(X)$.
4. *If* $X_1 \leq X_2$, *then* $\overline{\mathbb{E}}(X_1) \leq \overline{\mathbb{E}}(X_2)$.
5. *If* $X(\omega) = c$ *for all* $\omega \in \Omega$, *then* $\overline{\mathbb{E}}(X) = c$.

Define:

$$\underline{\mathbb{E}}(X) := -\overline{\mathbb{E}}(-X).$$

Prove:

$$\underline{\mathbb{E}}(X) \leq \overline{\mathbb{E}}(X)$$

When $\underline{\mathbb{E}}(X)$ and $\overline{\mathbb{E}}(X)$ are equal, we write $\mathbb{E}(X)$ for their common value and call it X's *global expected value*.

The following relations hold for any event E

$$\overline{\mathbb{P}}(E) + \overline{\mathbb{P}}(E^{\mathsf{c}}) \geq 1,$$

$$\underline{\mathbb{P}}(E) = 1 - \overline{\mathbb{P}}(E^{\mathsf{c}}),$$

$$0 \leq \underline{\mathbb{P}}(E) \leq \overline{\mathbb{P}}(E) \leq 1.$$

When $\underline{\mathbb{E}}(X)$ and $\overline{\mathbb{E}}(X)$ are equal, we write $\mathbb{E}(X)$ for their common value and call it X's *global expected value*.

Some theorems we prove game-theoretically:

Testing a probability p

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, \dots, N$:
  Skeptic announces $M_n \in \mathbb{R}$.
  Reality announces $y_n \in \{0, 1\}$.
  $\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p)$.

Bernoulli's theorem. For any $\epsilon > 0$,

$$\lim_{N\to\infty} \mathbb{P}\left(|\overline{y}_N - p| \geq \epsilon\right) = 0.$$

De Moivre's theorem. When $a < b$,

$$\lim_{N\to\infty} \mathbb{P}\left(a < \frac{\overline{y}_N - p}{\sqrt{p(1-p)/N}} < b\right) = \int_a^b \mathcal{N}_{0,1}(dz).$$

Testing a probability p

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \dots$:
  Skeptic announces $M_n \in \mathbb{R}$.
  Reality announces $y_n \in \{0, 1\}$.
  $\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p)$.

And after we give infinite-horizon definitions:

Borel's law of large numbers.

$$\mathbb{P}(\lim_{n\to\infty} \overline{y}_n = p) = 1$$

Probability forecasting

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.

FOR $n = 1, 2, \dots$:

Forecaster announces $p_n \in [0, 1]$.

Skeptic announces $M_n \in \mathbb{R}$.

Reality announces $y_n \in \{0, 1\}$.

$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(y_n - p_n)$.

More complicated:

- Events and variables may depend on both Forecaster and Reality.
- We do not have probabilities for Forecaster's moves.

We obtain only <u>upper</u> probabilities and <u>upper</u> expected values.

But we can still prove versions of standard theorems. For example:

$$\overline{\mathbb{P}}\left(|\overline{y}_N - \overline{p}_N| \geq \epsilon\right) \leq \frac{1}{\epsilon^2 N}$$

$$\overline{\mathbb{P}}\left(|\overline{y}_N - \overline{p}_N| \geq \epsilon\right) \leq 2\exp\left(-\frac{\epsilon^2 N}{4}\right)$$

$$\mathbb{P}(\lim_{n\to\infty}(\overline{y}_n - \overline{p}_n) = 0) = 1$$

4. For fun: History of martingales

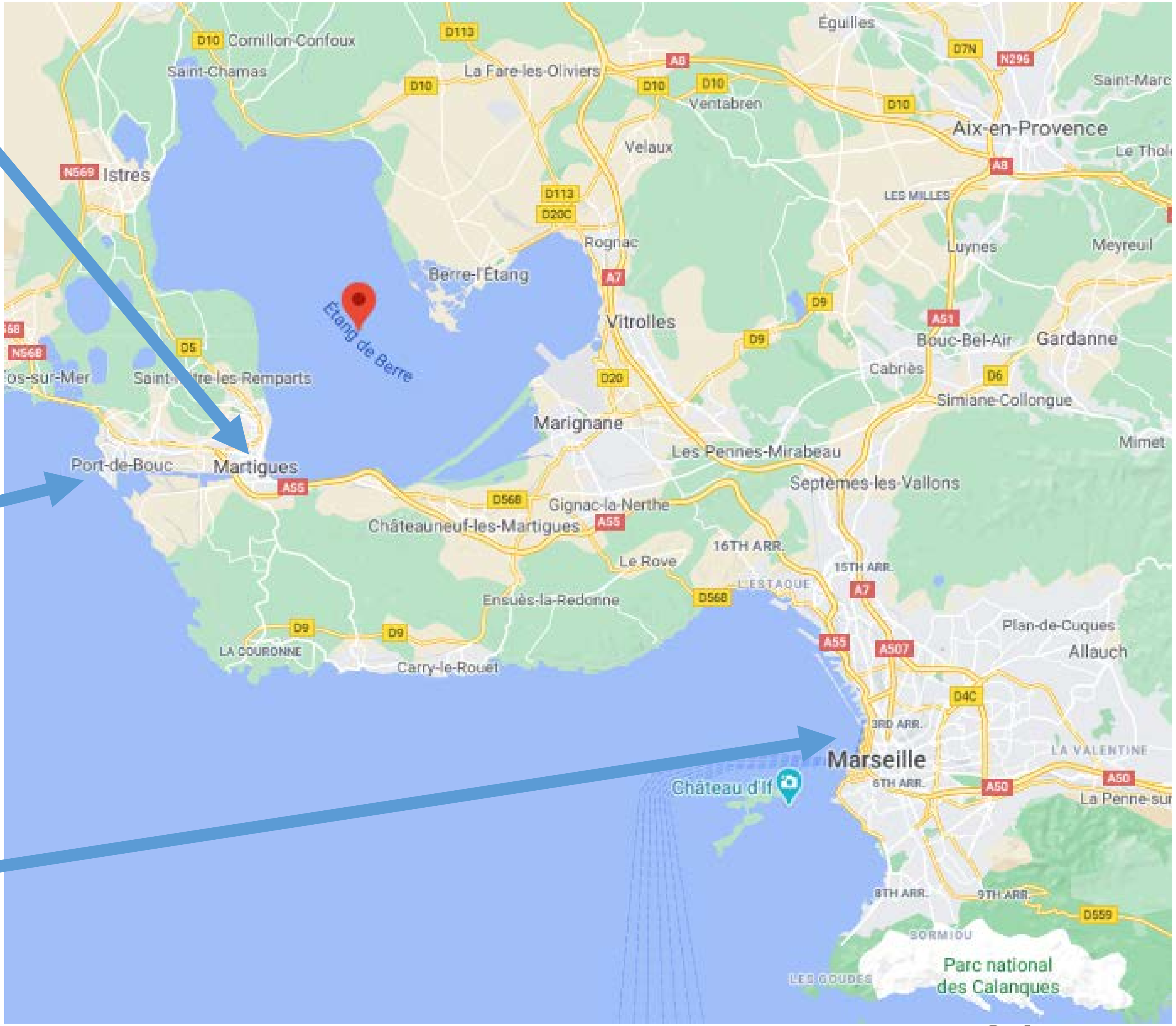
> You have not played as yet? Do not do so; above all avoid a martingale if you do. Play ought not to be an affair of calculation, but of inspiration. I have calculated infallibly, and what has been the effect?
>
> William Thackery, *The Newcomes*, 1854

Between the Mediterranean Sea and a large salt-water lagoon, Martigues had great boat-builders and sailors.

People have lived here for 6,000 years. Founded in 1232, Martigues now has about 50,000 people.

Canal

Marseilles, founded by the Greeks around 600 BC is the second largest city in France.

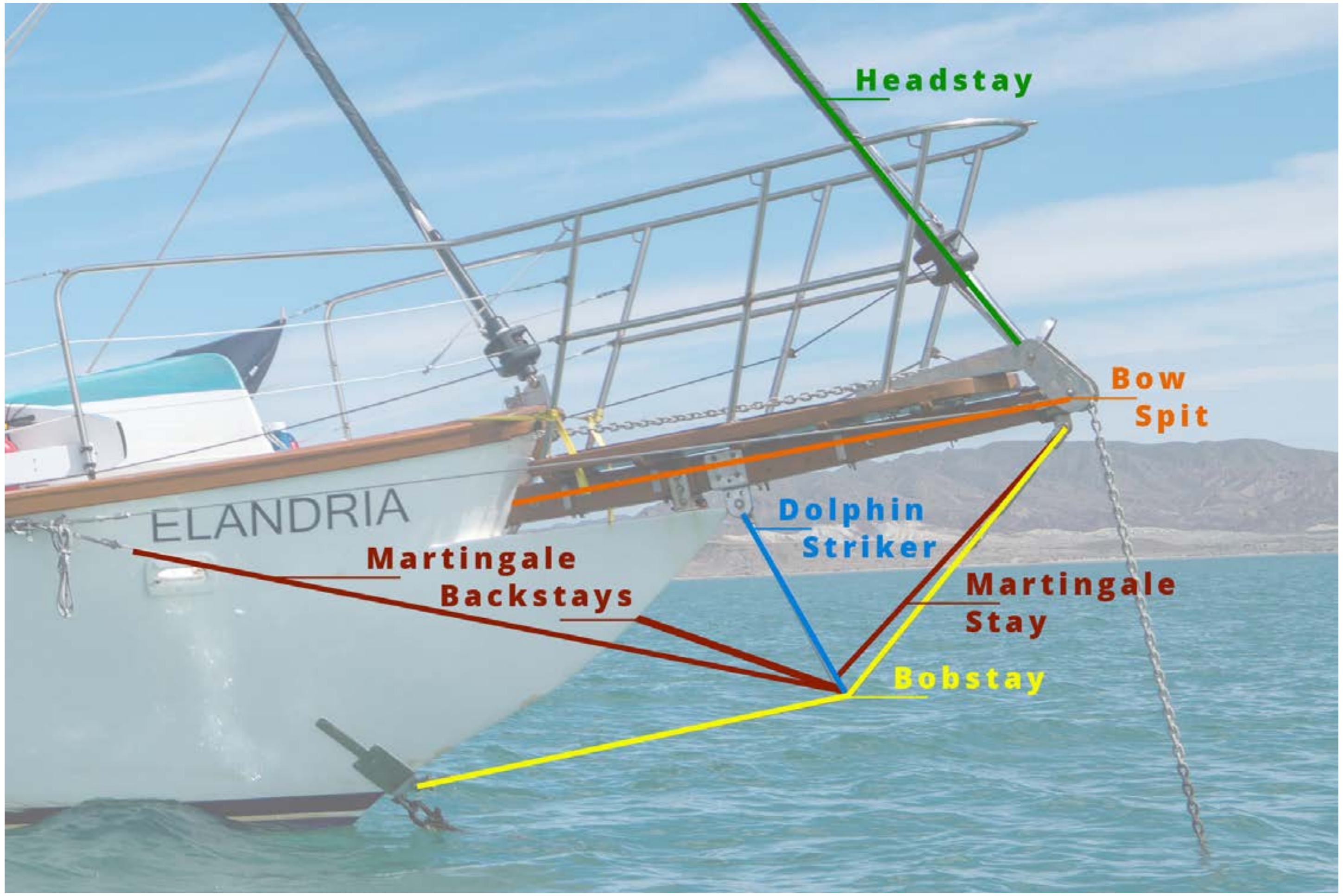


Martingales help hold down foresails.

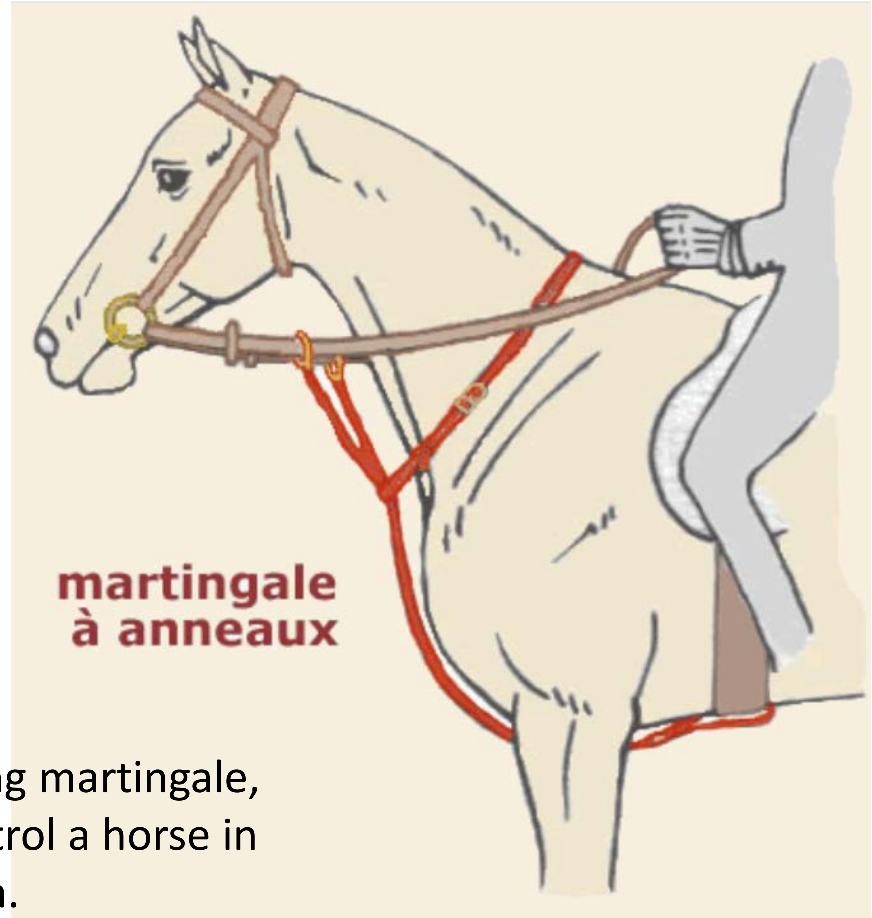
This martingale consists of three stays.

The sails put great upward pressure on the bow spit.

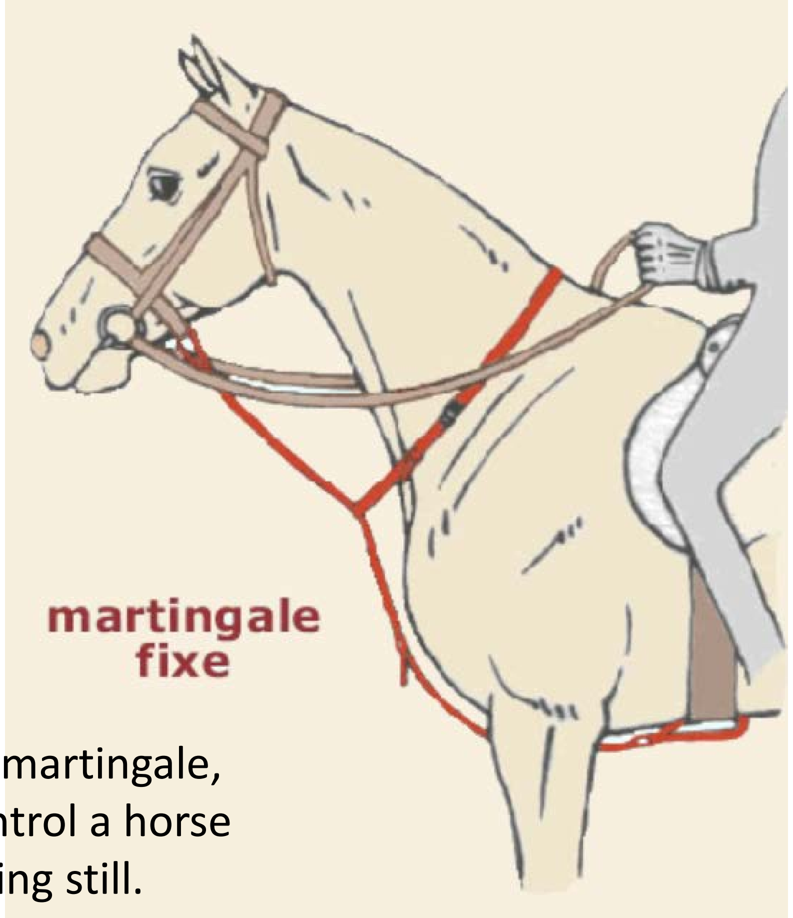
- The martingale resists the upward pull on the bow spit.
- The dolphin striker holds the martingale out.
- The bobstay also helps, but it does not have the martingale's stability.



“Martingale” also became a name for straps used to control a horse, apparently because of the resemblance to the three stays of a sailing martingale.



Running martingale, to control a horse in motion.



Fixed martingale, to control a horse standing still.

A dog collar with two loops is a “martingale”.

The larger loop chokes only when the dog pulls too hard.

Even men’s coats can be “martingales”.

Originally a type of pants that opened in the back.

Inhabitants of Martigues also known for extravagant dress and behavior: behaved ***à la martingale***.

In the 1700s, a method of betting came to be called a ***martingale***: keep doubling your bet until you win.

Bet	Result	Gain	Cumulative
\$1	lose	−\$1	−\$1
\$2	lose	−\$2	−\$3
\$4	lose	−\$4	−\$7
\$8	lose	−\$8	−\$15
\$16	win	\$16	\$1

Double or quits, double or nothing, *quitte ou double*, *doppelt oder nichts*. When you lose a bet, bet twice as much. By winning, you will recover your loss and also make the gain you had first hoped for. This idea for turning loss into gain is surely as ancient as gambling itself.

Repeated doubling following repeated losses can hardly be less ancient. As the losses mount, the redoubling becomes a strategy of desperation, and the gambler is ruined. The earliest known literary enactment of this spectacle is the 13th century fabliau of St. Peter and the Minstrel, which explains why there are no minstrels in hell. There once was a minstrel in hell, we are told. Satan had left this minstrel in charge when he and his devils went to hunt more souls. St. Peter, spying an opportunity, engaged the minstrel in a game of *hazart*, a three-die ancestor of craps: St. Peter's gold against the minstrel's souls. The minstrel gambled away all the souls. Satan, enraged when he returned, expelled the minstrel from hell.[1]

Review: The notion of a martingale

Define a joint probability distribution P for $Y_1, Y_2, \dots$ by giving probability distributions

- for Y_1 and
- for Y_n for each sequence $y_1, \dots, y_{n-1}$.

This makes P a strategy for Forecaster, because the conditional expected value $\mathbf{E}_P(S(Y_n)|y_1, \dots, y_{n-1})$ is well defined whenever S is a measurable function of Y_n or, equivalently, a measurable function of $Y_1, \dots, Y_n$.

Imposing this strategy on Forecaster and removing him from the game, we have a protocol for testing P:

$\mathcal{K}_0 := 1$.
FOR $n = 1, 2, \dots, N$:
  Skeptic announces $S_n : \mathcal{Y} \to (-\infty, \infty)$ such that
    $\mathbf{E}_P(S_n(Y_n)|y_1, \dots, y_{n-1}) = \mathcal{K}_{n-1}$.
  Reality announces $y_n \in \mathcal{Y}$.
  $\mathcal{K}_n := S_n(y_n)$.

Standard probability theory for discrete time is the very special case of game-theoretic probability represented by this protocol.

In the even more special case where a strategy for Skeptic is fixed, Skeptic's capital process $\mathcal{K}_0, \mathcal{K}_1, \dots$ is a martingale in the sense of standard probability theory.

Martingales in measure-theoretic probability.

There are at least three widely used definitions of the notion of a *martingale* in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

1. A *martingale* is a sequence of random variables $M_0, M_1, \dots$ such that

$$\mathbb{E}(M_n | M_1, \dots, M_{n-1}) = M_{n-1}$$

   for $n = 1, 2, \dots$.

2. Fix a sequence $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$ of σ-algebras, all contained in $\mathcal{F}$. A *martingale with respect to this filtration* is a sequence of random variables $M_0, M_1, \dots$ such that

$$\mathbb{E}(M_n | \mathcal{F}_{n-1}) = M_{n-1}$$

   for $n = 1, 2, \dots$.

3. Fix a sequence of random variables $Y_0, Y_1, \dots$. A *martingale with respect to* $Y_0, Y_1, \dots$ is a sequence of random variables $M_0, M_1, \dots$ such that

$$\mathbb{E}(M_n | Y_1, \dots, Y_{n-1}) = M_{n-1}$$

   for $n = 1, 2, \dots$.

Martingales in statistics and other applications of probability

Instead of beginning with probability space and using measure-theoretic conditional expectation, define a joint probability distribution for $Y_0, Y_1, \ldots$ by

- giving a probability distribution for Y_0 and
- giving a probability distribution for Y_n for each sequence $y_0, \ldots, y_{n-1}$.

A *martingale* is a sequence of random variables $M_0, M_1, \ldots$ such that

- M_n is a function of $Y_0, \ldots, Y_n$ and
- $\mathbb{E}(M_n | y_1, \ldots, y_{n-1}) = M_{n-1}(y_0, \ldots, y_{n-1})$.

Notice that after $y_0, \ldots, y_{n-1}$ are observed, the random variable M_n depends on Y_n only.

For a treatment of martingales in standard but non-measure-theoretic probability, see Stewart Ethier's *Doctrine of Chances*, Springer, 2010, Chapter 3.

Review: Heads and tails as an adversarial game

Let's review the game-theoretic definition of probability using the simplest and most familiar example.

The "fair" coin

Skeptic announces $\mathcal{K}_0 > 1$.

FOR $n = 1, 2, \ldots$:

Skeptic announces $M_n \in [-\mathcal{K}_{n-1}, \mathcal{K}_{n-1}]$.

Reality announces $y_n \in \{\mathrm{H}, \mathrm{T}\}$.

$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n \times (1 \text{ if } y_n = H, -1 \text{ if } Y_n = T)$.

The "fair" coin

Skeptic announces $\mathcal{K}_0 > 1$.
FOR $n = 1, 2, \ldots$:
Skeptic announces $M_n \in [-\mathcal{K}_{n-1}, \mathcal{K}_{n-1}]$.
Reality announces $y_n \in \{\mathrm{H}, \mathrm{T}\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n \times (1 \text{ if } y_n = H, -1 \text{ if } Y_n = T)$.

In this infinite-horizon protocol, upper and lower probabilities are equal to each other and to the usual probabilities for coin-tossing.

The protocol forces Skeptic's capital processes (i.e., the supermartingales) to be nonnegative.

The "fair" coin

Skeptic announces $\mathcal{K}_0 > 1$.
FOR $n = 1, 2, \ldots$:
Skeptic announces $M_n \in [-\mathcal{K}_{n-1}, \mathcal{K}_{n-1}]$.
Reality announces $y_n \in \{\mathrm{H}, \mathrm{T}\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n \times (1 \text{ if } y_n = H, -1 \text{ if } Y_n = T)$.

In this protocol,

- the *sample space* Ω is the set of all infinite sequences of Hs and Ts, and
- all measurable subsets of Ω have equal upper and lower probabilities.

Definitions applicable to all protocols where Skeptic can choose initial capital $\mathcal{K}_0$ but must keep capital nonnegative

- *Event:* A subset of Ω, measurable or not, is an event.
- *Upper probability:* The upper probability of an event E is the infimum of all α such that Skeptic has a strategy with $\mathcal{K}_0 = \alpha$ and $\lim_{n\to\infty} \mathcal{K}_n(\omega) \geq 1$ for all $\omega \in E$.
- *Almost sure:* (E is almost sure) means $(\overline{\mathbb{P}}(E^c) = 0)$ or $(\underline{\mathbb{P}}(E) = 1)$.

- *Upper probability:* The upper probability of an event E is the infimum of all α such that Skeptic has a strategy with $\mathcal{K}_0 = \alpha$ and $\lim_{n\to\infty} \mathcal{K}_n(\omega) \geq 1$ for all $\omega \in E$.

- *Almost sure:* (E is almost sure) means ($\overline{\mathbb{P}}(E^c) = 0$).

$$\boxed{\begin{gathered}\text{Skeptic has strategy with } \mathcal{K}_0 = \alpha \text{ and } \lim_{n\to\infty} \mathcal{K}_n(\omega) \geq 1 \text{ for all } \omega \in E.\\ \Longleftrightarrow \\ \text{Skeptic has strategy with } \mathcal{K}_0 = 1 \text{ and } \lim_{n\to\infty} \mathcal{K}_n(\omega) \geq 1/\alpha \text{ for all } \omega \in E.\end{gathered}}$$

Obvious Lemma. E is almost sure if and only if for every $C > 0$ Skeptic has a strategy such that $\mathcal{K}_0 = 1$ and $\lim_{n\to\infty} \mathcal{K}_n(\omega) \geq C$ for all $\omega \notin E$.

Easy Lemma. E is almost sure if and only if for Skeptic has a strategy such that $\mathcal{K}_0 = 1$ and $\lim_{n\to\infty} \mathcal{K}_n(\omega) = \infty$ for all $\omega \notin E$.

Roughly speaking, E is almost sure means Skeptic can get infinitely rich unless E happens.

The "fair" coin

Skeptic announces $\mathcal{K}_0 > 1$.

FOR $n = 1, 2, \ldots$:

Skeptic announces $M_n \in [-\mathcal{K}_{n-1}, \mathcal{K}_{n-1}]$.

Reality announces $y_n \in \{\mathrm{H}, \mathrm{T}\}$.

$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n \times (1 \text{ if } y_n = H, -1 \text{ if } Y_n = T)$.

Theorem. The event

$$\lim_{n\to\infty} \frac{\text{number of Hs in the first } n \text{ rounds}}{n} = \frac{1}{2}$$

is almost sure.

In other words, Skeptic has a strategy such that guarantees, no matter what Reality does, that either he gets infinitely rich or else the fraction of heads converges to 1/2.

In other words, Skeptic has a strategy such that guarantees, no matter what Reality does, that either he gets infinitely rich or else the fraction of heads converges to 1/2.

Question: Why do we put "Reality" in an adversarial role. Is reality malicious?
Answer: Making Reality the opponent merely dramatizes the worst-case nature of our results.

Any existence statement in mathematics asserts the existence of a winning strategy in an adversarial two-person perfect-information game.

- "For every a in A, there is a b in B satisfying C(a,b)." Here the game is that Player I chooses a in A. Then Player II choses b in B. Player II wins if a and b satisfy C(a,b). So the statement in quotes says that Player II has a winning strategy.
- "There exists a in A such that C(a,b) holds for all b in B". Again Player I chooses a in A and then Player II choses b in B. Player I wins if C(a,b) holds. The statement in quotes says that Player I has the winning strategy.

The statement in the red box has the second form.

To simplify, suppose Skeptic announces a betting strategy, then Reality chooses a sequence of heads and tails knowing Skeptic's strategy. Skeptic wins if he gets infinitely rich or half Reality's moves are heads. Skeptic has a winning strategy.