Probability without Measure!

Mark Saroufim

University of California San Diego

msaroufi@cs.ucsd.edu

February 18, 2014

Overview

A gambler's perspective

- This was the day before probability theory was even a field in mathematics, a field without foundations. Pascal and Fermat simply wanted to win a ton of money betting on horses and wanted to first see what it meant for a game to be fair.

A gambler's perspective

- This was the day before probability theory was even a field in mathematics, a field without foundations. Pascal and Fermat simply wanted to win a ton of money betting on horses and wanted to first see what it meant for a game to be fair.
  - If I pay a\$
  - You pay a\$
  - The winner gets paid $2a$\$

A gambler's perspective

- This was the day before probability theory was even a field in mathematics, a field without foundations. Pascal and Fermat simply wanted to win a ton of money betting on horses and wanted to first see what it meant for a game to be fair.
    - If I pay a\$
    - You pay a\$
    - The winner gets paid $2a$\$
- This is what is referred to as *inter alia* (equal terms)
- $P[E] =$ how much money you're willing to put on a game where you could win 1\$

Looking at the real world

Bernoulli was the first to suggest that probability can be measured from observation

$$P\{|y/N - p| < \epsilon\} > 1 - \delta$$

Now it seems that there could be a more mathematical treatment of probability..

Kolmogorov's axioms

- The axioms and definitions below relate a set Ω called the sample space and the set of subsets of Ω, $\mathcal{F}$. Every element in $E \in \mathcal{F}$ is called an event
  1. If $E, F \in \mathcal{F}$ then $E \cup F, E \cap F, E \setminus F \in \mathcal{F}$. Or more concisely we say that $\mathcal{F}$ is a field of sets.
  2. $\Omega \subset F$ which with the first axiom means that $\mathcal{F}$ is an algebra of sets
  3. Every set $E \in \mathcal{F}$ is assigned a probability which is a non-negative real value using the function $P : E \to [0, 1]$
  4. $P[\Omega] = 1$
  5. If $E \cap F = \Phi$ then $P[E \cap F] = P[E] + P[F]$, more generally we get what is called the union bound when E and F are not disjoint then $P[E \cap F] \leq P[E] + P[F]$
  6. If $\cap_{n=1}^{\infty} E_n = \Phi$ where $E_n \subseteq E_{n-1} \cdots \subseteq E_1$ we have that $\lim_{n\to\infty} P[E_n] = 0$. This axiom with axiom 2 allows us to call $\mathcal{F}$ a σ-algebra
- A random variable x is then understood as a mapping from the size of elements of $\mathcal{F}$ with respect to the probability measure P

Some Set Theory

Kolmogorov's 6th axiom needs the axiom of choice

Some Set Theory

Kolmogorov's 6th axiom needs the axiom of choice

Definition (ZFC Axiom of Choice)

It is possible to pick one cookie from each jar even if there is an infinite numbers of jars

Some Set Theory

Kolmogorov's 6th axiom needs the axiom of choice

Definition (ZFC Axiom of Choice)

It is possible to pick one cookie from each jar even if there is an infinite numbers of jars

Weird things happen: Can divide a ball into two balls of equal volume

Some Set Theory

Kolmogorov's 6th axiom needs the axiom of choice

Definition (ZFC Axiom of Choice)

It is possible to pick one cookie from each jar even if there is an infinite numbers of jars

Weird things happen: Can divide a ball into two balls of equal volume

Definition (Axiom of Determinacy)

Every 2 player game with perfect information where two players pick natural numbers at every turn is already determined.

Some Set Theory

Kolmogorov's 6th axiom needs the axiom of choice

Definition (ZFC Axiom of Choice)

It is possible to pick one cookie from each jar even if there is an infinite numbers of jars

Weird things happen: Can divide a ball into two balls of equal volume

Definition (Axiom of Determinacy)

Every 2 player game with perfect information where two players pick natural numbers at every turn is already determined.

Weird things also happen: We get that there is no such thing as non-measurable sets

Sequential Learning

- Von Mises was the first to propose that probability could find its foundations in games
- Given a bit string 001111
- Predict the odds of a 1 (number shouldn't change much if we look at a subsequence called a collective)

Sequential Learning

- Von Mises was the first to propose that probability could find its foundations in games
- Given a bit string 001111
- Predict the odds of a 1 (number shouldn't change much if we look at a subsequence called a collective)
- Fortunately we have a method of quantifying how difficult it is to predict the next bit in a string: Kolmogorov complexity!

Martingales

Originally Martingales are a gambling strategy that can guarantee a win of 1$ given an infinite supply of money

$$\alpha + 2\alpha + \cdots + 2^i\alpha$$

stop as soon as you win once

Martingales

Originally Martingales are a gambling strategy that can guarantee a win of 1$ given an infinite supply of money

$$\alpha + 2\alpha + \cdots + 2^i\alpha$$

stop as soon as you win once

Definition (Martingale)

Given a sequence of outcomes $x_1, \ldots, x_n$ we call a capital process L if

$$E[L(x_1, \ldots, x_n) \mid x_1, \ldots, x_{n-1}] = L(x_1, \ldots, x_n)$$

Martingales

Originally Martingales are a gambling strategy that can guarantee a win of 1$ given an infinite supply of money

$$\alpha + 2\alpha + \cdots + 2^i\alpha$$

stop as soon as you win once

Definition (Martingale)

Given a sequence of outcomes $x_1, \ldots, x_n$ we call a capital process L if

$$E[L(x_1, \ldots, x_n) \mid x_1, \ldots, x_{n-1}] = L(x_1, \ldots, x_n)$$

$L(E) \to \infty$ if E has probability 0 (more on this next slide)
Now we define the probability of an event E as

$$P(E) = \inf\{L_0 \mid \lim_{n\to\infty} L_n \geq I\}$$

Martingales

Theorem (Doob's inequality)

$$P[\sup_n L(x_1, \ldots, x_n) \geq \lambda] \leq \frac{1}{\lambda}$$

Look familiar?

Martingales

Theorem (Doob's inequality)

$$P[\sup_n L(x_1, \dots, x_n) \geq \lambda] \leq \frac{1}{\lambda}$$

Look familiar?

Markov's inequality!

$$P[x \geq \lambda] \leq \frac{Ex}{\lambda}$$

Martingales

Theorem (Doob's inequality)

$$P[\sup_n L(x_1, \ldots, x_n) \geq \lambda] \leq \frac{1}{\lambda}$$

Look familiar?

Markov's inequality!

$$P[x \geq \lambda] \leq \frac{Ex}{\lambda}$$

Other Chernoeff bounds can be derived in this way as well.

Bounded Fair Coin Game

We're now ready to setup the first game between what we the skeptic and nature

Bounded Fair Coin Game

We're now ready to setup the first game between what we the skeptic and nature

- K_i is the skeptic's capital at time i

Bounded Fair Coin Game

We're now ready to setup the first game between what we the skeptic and nature

- K_i is the skeptic's capital at time i
- M_n is the amount of tickets that the skeptic purchases

Bounded Fair Coin Game

We're now ready to setup the first game between what we the skeptic and nature

- K_i is the skeptic's capital at time i
- M_n is the amount of tickets that the skeptic purchases
- x_n is the value of a ticket (determined by nature)

Bounded Fair Coin Game

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

Theorem

There exists a winning strategy for skeptic but let's formally define what we mean by winning

Winning conditions

We claim that the skeptic wins if $K_n > 0 \forall n$ and if one two things happen, either

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} x_i = 0$$

Winning conditions

We claim that the skeptic wins if $K_n > 0 \forall n$ and if one two things happen, either

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} x_i = 0$$

or

$$\lim_{n\to\infty} K_n = \infty$$

Winning conditions

We claim that the skeptic wins if $K_n > 0 \forall n$ and if one two things happen, either

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} x_i = 0$$

or

$$\lim_{n\to\infty} K_n = \infty$$

- The first condition has two different interpretations in the finance litterature is called a self financing strategy because it means that the skeptic can remain in the game forever without ever having to borrow money. Also it means that skeptic wins if nature is forced to play randomly.

Winning conditions

We claim that the skeptic wins if $K_n > 0 \forall n$ and if one two things happen, either

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} x_i = 0$$

or

$$\lim_{n\to\infty} K_n = \infty$$

- The first condition has two different interpretations in the finance litterature is called a self financing strategy because it means that the skeptic can remain in the game forever without ever having to borrow money. Also it means that skeptic wins if nature is forced to play randomly.
- The second makes sense, you win if you become infinitely rich but since this is unlikely this condition embodies the infinitary hypothesis which says that there is no strategy that avoids bankruptcy that guarantees that skeptic become infinitely rich.

Winning conditions

We claim that the skeptic wins if $K_n > 0 \forall n$ and if one two things happen, either

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} x_i = 0$$

or

$$\lim_{n\to\infty} K_n = \infty$$

- The first condition has two different interpretations in the finance litterature is called a self financing strategy because it means that the skeptic can remain in the game forever without ever having to borrow money. Also it means that skeptic wins if nature is forced to play randomly.
- The second makes sense, you win if you become infinitely rich but since this is unlikely this condition embodies the infinitary hypothesis which says that there is no strategy that avoids bankruptcy that guarantees that skeptic become infinitely rich.

Back to the Fair Coin Game

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

Theorem

There exists a winning strategy for skeptic

Back to the Fair Coin Game

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

Theorem

There exists a winning strategy for skeptic

Law of Large Numbers.

Skeptic bets ϵ on heads, this forces nature not to play heads often or else skeptic will become infinitely rich. So nature will start playing tails, when that happens skeptic puts an ϵ on tails. $\square$

Bounded Fair Coin Game

What if $x_n \in [-1, 1]$ instead of $\{-1, 1\}$?

Players: Skeptic, Reality

Protocol:

$\mathcal{K}_0 = 1.$

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

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

Reality announces $x_n \in [-1, 1]$.

$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n x_n.$

Theorem

There exists a winning strategy for skeptic

Proof of Bounded Fair Coin Game

We will need some terminology to tackle this problem we define a real valued function on Ω called P which is a strategy that takes situations $s = x_1, x_2, \ldots, x_n$ and decides the number of tickets to buy $P(s)$.

Proof of Bounded Fair Coin Game

We will need some terminology to tackle this problem we define a real valued function on Ω called P which is a strategy that takes situations $s = x_1, x_2, \ldots, x_n$ and decides the number of tickets to buy $P(s)$.

$$K^P(x_1x_2\ldots x_n) = K^P(x_1x_2\ldots x_{n-1}) + P(x_1x_2\ldots x_n)x_n$$

Proof of Bounded Fair Coin Game

We will need some terminology to tackle this problem we define a real valued function on Ω called P which is a strategy that takes situations $s = x_1, x_2, \ldots, x_n$ and decides the number of tickets to buy $P(s)$.

$$K^P(x_1 x_2 \ldots x_n) = K^P(x_1 x_2 \ldots x_{n-1}) + P(x_1 x_2 \ldots x_n) x_n$$

Definition

Skeptic forces an event E if $K^P(s) = \infty \forall s \in E^c$

Proof of Bounded Fair Coin Game

Lemma

The skeptic can force

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

and

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

Proof of Bounded Fair Coin Game

Proof.

take 1 as the starting capital

$$1 + K^P(x_1 x_2 \dots x_n) = (1 + K^P(x_1 \dots x_{n-1}))(1 + \epsilon x_n) = \prod_{i=1}^{n}(1 + \epsilon x_i) < C$$

where C is a constant so take the log on both sides

$$\sum_{i=1}^{n} \ln(1 + \epsilon x_i) \leq D$$

Now use $\ln(1 + t) \geq t - t^2$ when $t \geq -1/2$

$$\frac{1}{n}\sum_{i=1}^{n} x_i \leq \frac{D}{\epsilon n} + \epsilon$$

and we get the top part of the lemma. Replace by $-\epsilon$ to get the second

Bounded Forecast games

Somebody has got to be setting the prices, let a forecaster announce price of ticket at iteration n as m_n

Parameter: $C > 0$
Players: Forecaster, Skeptic, Reality
Protocol:

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

Theorem

There exists a winning strategy for skeptic by reduction to the bounded fair coin game.

Bounded Forecast games

Somebody has got to be setting the prices, let a forecaster announce price of ticket at iteration n as m_n

Parameter: $C > 0$
Players: Forecaster, Skeptic, Reality
Protocol:

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

Theorem

There exists a winning strategy for skeptic by reduction to the bounded fair coin game.

Proof.

First divide all prices by C to normalize prices to $[-1, 1]$ then set $m_n = 0$ and we recover the previous game. Note we also need to change the first condition to $\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n}(x_i - m_i) = 0$ □

Measure theoretic law of large numbers

Assuming X_i are i.i.d random variables with mean μ and variance σ^2 we define $A_n = \frac{X_1+X_2+\cdots+X_n}{n}$ then $E[A_n] = \frac{n\mu}{n} = \mu$ and similarly $Var[A_n] = \frac{n\sigma^2}{n^2} = \sigma^2/n$. By Chebyshev's inequality we get the weak law of large numbers

$$P(\ |A_n - \mu| \geq \epsilon) \leq Var[A_n]/\epsilon^2 = \frac{\sigma^2}{n\epsilon^2}$$

Measure theoretic law of large numbers

Assuming X_i are i.i.d random variables with mean μ and variance σ^2 we define $A_n = \frac{X_1+X_2+\cdots+X_n}{n}$ then $E[A_n] = \frac{n\mu}{n} = \mu$ and similarly $Var[A_n] = \frac{n\sigma^2}{n^2} = \sigma^2/n$. By Chebyshev's inequality we get the weak law of large numbers

$$P(\ |A_n - \mu| \geq \epsilon) \leq Var[A_n]/\epsilon^2 = \frac{\sigma^2}{n\epsilon^2}$$

To prove Chebyshev's we define $A = \{w \in \Omega \mid |X(w)| \geq \alpha\}$. $X(w) \geq \alpha I_A(w)$. Take the expectation on both sides to get $E(|X|) \geq \alpha E(I_A) = \alpha P(A)$

Measure theoretic law of large numbers

Assuming X_i are i.i.d random variables with mean μ and variance σ^2 we define $A_n = \frac{X_1+X_2+\cdots+X_n}{n}$ then $E[A_n] = \frac{n\mu}{n} = \mu$ and similarly $Var[A_n] = \frac{n\sigma^2}{n^2} = \sigma^2/n$. By Chebyshev's inequality we get the weak law of large numbers

$$P(\ |A_n - \mu| \geq \epsilon) \leq Var[A_n]/\epsilon^2 = \frac{\sigma^2}{n\epsilon^2}$$

To prove Chebyshev's we define $A = \{w \in \Omega \mid |X(w)| \geq \alpha\}$.
$X(w) \geq \alpha I_A(w)$. Take the expectation on both sides to get
$E(|X|) \geq \alpha E(I_A) = \alpha P(A)$
In game theoretic proof we don't need i.i.d assumption

Measure theoretic law of large numbers

Assuming X_i are i.i.d random variables with mean μ and variance σ^2 we define $A_n = \frac{X_1+X_2+\cdots+X_n}{n}$ then $E[A_n] = \frac{n\mu}{n} = \mu$ and similarly $Var[A_n] = \frac{n\sigma^2}{n^2} = \sigma^2/n$. By Chebyshev's inequality we get the weak law of large numbers

$$P(\ |A_n - \mu| \geq \epsilon) \leq Var[A_n]/\epsilon^2 = \frac{\sigma^2}{n\epsilon^2}$$

To prove Chebyshev's we define $A = \{w \in \Omega \mid |X(w)| \geq \alpha\}$.
$X(w) \geq \alpha I_A(w)$. Take the expectation on both sides to get
$E(|X|) \geq \alpha E(I_A) = \alpha P(A)$
In game theoretic proof we don't need i.i.d assumption we don't even to assume a distribution exists!

Unbounded game

Players: Forecaster, Skeptic, Reality

Protocol:

$\mathcal{K}_0 := 1.$

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

Forecaster announces $m_n \in \mathbb{R}$ and $v_n \geq 0$.

Skeptic announces $M_n \in \mathbb{R}$ and $V_n \geq 0$.

Reality announces $x_n \in \mathbb{R}$.

$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(x_n - m_n) + V_n((x_n - m_n)^2 - v_n).$

Unbounded game

Players: Forecaster, Skeptic, Reality
Protocol:

$\mathcal{K}_0 := 1$.

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

Forecaster announces $m_n \in \mathbb{R}$ and $v_n \geq 0$.

Skeptic announces $M_n \in \mathbb{R}$ and $V_n \geq 0$.

Reality announces $x_n \in \mathbb{R}$.

$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(x_n - m_n) + V_n((x_n - m_n)^2 - v_n)$.

Theorem

If $\sum_{n=1}^{\infty} \frac{v_n}{n^2} < \infty$ *then the skeptic has a winning strategy*

Unbounded game

Players: Forecaster, Skeptic, Reality
Protocol:

$\mathcal{K}_0 := 1$.
FOR $n = 1, 2, \ldots$:
Forecaster announces $m_n \in \mathbb{R}$ and $v_n \geq 0$.
Skeptic announces $M_n \in \mathbb{R}$ and $V_n \geq 0$.
Reality announces $x_n \in \mathbb{R}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + M_n(x_n - m_n) + V_n((x_n - m_n)^2 - v_n)$.

Theorem

If $\sum_{n=1}^{\infty} \frac{v_n}{n^2} < \infty$ *then the skeptic has a winning strategy*

Proof.

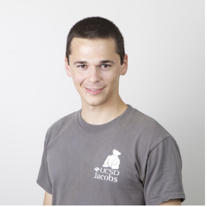
Similar in nature to proof of the bounded fair coin game. Main idea is that the skeptic's capital is a supermartingale (a sequence that decreases in expectation) $\square$

What about an application

Suppose you're a clever young guy/gal who wants to make money off of these ideas

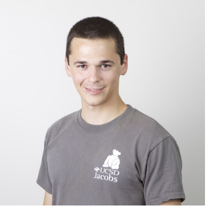
What about an application

Suppose you're a clever young guy/gal who wants to make money off of these ideas



What about an application

Suppose you're a clever young guy/gal who wants to make money off of these ideas



A natural next step is to make an infinitely large amount of money off the stock market

Unfortunately it seems that its difficult to have consistently better returns than the market and we will prove this. We make two assumptions that transaction costs are neglible (not as controversial as it sounds) and that the capital of a specific investor isn't too big relative to the market.

Securities Market Protocol

Parameters: $\mathcal{K}_0 > 0$, natural number $K > 1$
Players: Opening Market, Investor, Skeptic, Closing Market
Protocol:

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

Opening Market selects $m_n \in [0,1]^K$ such that $\sum_{k=1}^K m_n^k = 1$.
Investor selects $g_n \in \mathbb{R}^K$.
Skeptic selects $h_n \in \mathbb{R}^K$.
Closing Market selects $x_n \in [-1, \infty)^K$ such that $m_n \cdot x_n = 0$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + h_n \cdot x_n$.

Securities Market Protocol

Parameters: $\mathcal{K}_0 > 0$, natural number $K > 1$
Players: Opening Market, Investor, Skeptic, Closing Market
Protocol:

FOR $n = 1, 2, \ldots$:
  Opening Market selects $m_n \in [0,1]^K$ such that $\sum_{k=1}^{K} m_n^k = 1$.
  Investor selects $g_n \in \mathbb{R}^K$.
  Skeptic selects $h_n \in \mathbb{R}^K$.
  Closing Market selects $x_n \in [-1, \infty)^K$ such that $m_n \cdot x_n = 0$.
  $\mathcal{K}_n := \mathcal{K}_{n-1} + h_n \cdot x_n$.

Proof.

Maybe next time, Finance theory might need its own talk :) $\square$

References

Shafer and Vovk (2001)
Probability and Finance It's only a Game!

Ramon Van Handel
Stochastic Calculus

Peter Clark
All I ever needed to know from Set Theory

Let's think about how this could change machine learning, talk to me and let's write a paper about it!