An introduction to game-theoretic probability from statistical viewpoint

Akimichi Takemura
(joint with M.Kumon, K.Takeuchi and K.Miyabe)

University of Tokyo

May 14, 2013
RPTC2013

Contents of this talk

1. List of Tokyo papers
2. Background on game-theoretic probability (GTP)
3. Introduction to game-theoretic probability using a coin-tossing game
4. Bayesian Skeptic for the coin-tossing game
5. Non-negative martingales and likelihood ratios

List of Tokyo papers (in http://www.probabilityandfinance.com/)

1. "On a simple strategy weakly forcing the strong law of large numbers in the bounded forecasting game", Kumon and Takemura. *Ann. Inst. Stat. Math.*, **60**, 801–812. 2008.
2. "Game theoretic derivation of discrete distributions and discrete pricing formulas", Takemura and Taiji Suzuki. *J. Japan Stat. Soc.*, **37**, 87–104. 2007.
3. "Capital process and optimality properties of Bayesian Skeptic in the fair and biased coin games", Kumon, Takemura and Takeuchi. *Stochastic Analysis and Applications*, **26**, 1161–1180. 2008.
4. "Game-theoretic versions of strong law of large numbers for unbounded variables", Kumon, Takemura and A.Takeuchi. *Stochastics*, **79**, 449–468. 2007.
5. "Implications of contrarian and one-sided strategies for the fair-coin game", Yasunori Horikoshi and Takemura. *Stochastic Processes and their Applications*, **118**, 2125–2142. 2008.
6. "A new formulation of asset trading games in continuous time with essential forcing of variation exponent", Takeuchi, Kumon and Takemura. *Bernoulli*, **15**, 1243–1258. 2009.

List of Tokyo papers

7. "Multistep Bayesian strategy in coin-tossing games and its application to asset trading games in continuous time", Takeuchi, Kumon and Takemura. *Stochastic Analysis and Applications*, **28**, 842–861. 2010.
8. "The generality of the zero-one laws", by Takemura, V.Vovk and G.Shafer. *Ann. Inst. Stat. Math.*, **63**, 873-886. 2011.
9. "New procedures for testing whether stock price processes are martingales", Takeuchi, Takemura and Kumon. *Computational Economics*, **37**, No.1, 67–88. 2010.
10. "Sequential optimizing strategy in multi-dimensional bounded forecasting games", Kumon, Takemura and Takeuchi. *Stochastic Processes and their Applications*, **121**, 155–183. 2011.
11. "Sequential optimizing investing strategy with neural networks", Ryo Adachi and A.Takemura. *Expert Systems With Applications*. **38**, 12991–12998. 2011.
12. "Approximations and asymptotics of upper hedging prices in multinomial models", by Ryuichi Nakajima, Masayuki Kumon, A.Takemura and Kei Takeuchi. *Japan Journal of Industrial and Applied Mathematics*, **25**, 1–21. 2012.

List of Tokyo papers

13. "Convergence of random series and the rate of convergence of strong law of large numbers in game-theoretic probability", by Kenshi Miyabe and A.Takemura. *Stochastic Processes and their Applications*, **122**, 1–30. 2012.
14. "Bayesian logistic betting strategy against probability forecasting", *Stochastic Analysis and Applications*, **31**, 214–234. Masayuki Kumon, Jing Li, A.Takemura and Kei Takeuchi. 2013.
15. "The law of the iterated logarithm in game-theoretic probability with quadratic and stronger hedges", *Stochastic Processes and their Applications*. Kenshi Miyabe and A.Takemura. 2013.

Background on game-theoretic probability (GTP)

- Kolmogorov's Grundbegriffe (1933) established measure theoretic probability. It justifies mathematical operations such as limiting operations.
- On this firm ground, probability theory found applications in many fields.
- Axiomatic construction: probability is not defined by itself, like "points" or "lines". This actually broadened the applicability of probability theory.
- "Probability is just the Lebesgue measure", K.Ito, 1944.
- On the other hand, foundational arguments, such as Richard von Mises's collectives, have been almost forgotten by probabilists.
- Kolmogorov himself was somewhat hesitant:

  $\rightarrow$ proposal of Kolmogorov complexity

Shafer and Vovk (2001)

- Shafer and Vovk (2001) "Probability and Finance, It's Only a Game!" appeared.
- Vladimir Vovk (PhD, 1988, Moscow State U) is one of the last students of Kolmogorov.
- Around 2003, Takeuchi started to tell me that the book is very interesting. I gave a course on GTP for studying the book.
- In my opinion, at present it is the only alternative framework to measure-theoretic probability.
- Important theorems, such as the strong law of large numbers (SLLN), central limit theorem (CLT), the law of the iterated logarithm (LIL), can be proved in game-theoretic probability without requiring measure theory.

Strength and weakness of game-theoretic probability (GTP)

Strength

- Some clever proofs are very short. For example, even high school students can understand game-theoretic proof of SLLN.
- Black-Scholes formula and CLT are equivalent. In Shafer and Vovk, CLT and the Black-Scholes formula are proved "simultaneously". Their proof shows that these are equivalent. (They do not use characteristic functions, but use the heat equation.)
- In GTP, the set of measure-zero is often more explicitly treated, by an explicit betting strategy with its capital diverging to $+\infty$ on the set.
- Probability is not assume a priori. A game is assumed. Under the game, the players are forced to act probabilistically. (Why stock prices look random?)

Strength and weakness of GTP

Weakness

- Some proofs are, of course, almost the same in measure-theoretic probability and GTP.
- Some simple notions under usual probability, such as independence, identical distribution, are not easy to formulate. (GTP inherently assumes martingale.)
- In 2001 book, continuous stochastic processes were treated by nonstandard analysis, which was probably not very convincing to many people.
- This difficulty was overcome based on the idea in "A new formulation of asset trading games in continuous time ..." by Takeuchi, Kumon and Takemura, *Bernoulli*, 2009, and completely generalized in "Continuous-time trading and the emergence of probability" by Vladimir Vovk, *Finance and Stochastics*, 2012.

Introduction to GTP by a coin-tossing game

- Complete information game between players (two players version)
  - Skeptic (statistician, investor) bets on some outcome.
  - Reality (nature, market) decides the outcome.
- Skeptic $\to$ Reality $\to$ S $\to$ R $\to$.
  They play in turn.
- One round: (Skeptic's turn, Reality's turn) in this order
- $n = 1, 2, \ldots$ denote rounds.
- Skeptic's initial capital: $\mathcal{K}_0 = 1$
- At each round, Skeptic first announces how much he bets: $M_n \in \mathbb{R}$. M_n can be any real number and can be arbitrarily small. Negative M_n allowed (short selling).

Introduction to GTP by a coin-tossing game

- After knowing M_n, Reality chooses the outcome $x_n = 0$ or $x_n = 1$.
- Payoff to Skeptic : $M_n(x_n - p)$, where the "price" $0 < p < 1$ of the "ticket" is given before the game. p is the success probability or the "risk neutral probability".
- Skeptic's capital changes as $\mathcal{K}_n = \mathcal{K}_{n-1} + M_n(x_n - p)$.

In summary:

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

Introduction to GTP by a coin-tossing game

- Reality can choose the sign of $x_n - p$ as the opposite of the sign of M_n. Therefore Reality can always decrease Skeptic's capital.
- No-win situation for Skeptic?
- But then Reality is forced to observe SLLN!

Theorem *There exists Skeptic's strategy $\mathcal{P}$. (He can announce $\mathcal{P}$ even before the start of the game.) If Skeptic uses $\mathcal{P}$, then he is never bankrupt and whenever Reality violates*

$$\lim_{n\to\infty} \frac{1}{n}(x_1 + \cdots + x_n) = p,$$

then

$$\lim_{n\to\infty} \mathcal{K}_n = \infty.$$

Introduction to GTP by a coin-tossing game

- "In the coin-tossing game there exists a non-negative martingale which succeeds on the set"

$$\{\ x_1 x_2 \dots \ \mid \frac{x_1 + \cdots + x_n}{n} \not\to p\} \subset \{0,1\}^\infty \tag{1}$$

- We say that in the coin-tossing game "Skeptic can force SLLN".
- Reality can also have strategies (not fully explored yet).
- Bounded forecasting game: (1) is still true even if Reality can choose any real number in $[0,1]$, and $\{0,1\}^\infty$ is replaced by $[0,1]^\infty$.

Coin-tossing game with the third player

Complete information game between three players.

- Forecaster decides the price of the ticket
- Skeptic bets on the outcome.
- Reality decides the outcome.

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

Coin-tossing game with the third player

- In this game Skeptic can force

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

- He can also force

$$\sum_n p_n < \infty \;\Leftrightarrow\; \sum_n x_n < \infty$$

- Forecaster can also have strategies.

Bayesian Skeptic for a coin-tossing game (without Forecaster)

- Although the proof of SLLN in Shafer and Vovk (2001) is short, we gave an alternative proof (just for a coin-tossing game) based on Bayesian Skeptic (in *Stochastic Analysis and Applications*, 2008).
- We found that the strategy was already discussed in Jean Ville (1939 ls) "Étude critique de la notion de collectif" (English translation by G.Shafer).
- Kullback-Leibler divergence very naturally comes out from our strategy. So Ville might have known KL-divergence.

Bayesian Skeptic for a coin-tossing game (without Forecaster)

- We suppose that Skeptic uses a strategy based on a beta prior distribution for p

$$p \sim p^{\alpha-1}(1-p)^{\beta-1}/B(\alpha,\beta),$$

where α, β are prior numbers of heads and tails.

- Then his prediction of success probability for the n-th round is

$$\hat{p}_n = \frac{\text{``Number of heads up to } n-1\text{''} + \alpha}{n-1+\alpha+\beta}.$$

- Consider Skeptics strategy

$$\mathcal{P} : M_n = \mathcal{K}_{n-1}\frac{\hat{p}_n - p}{p(1-p)}$$

- In the following we let $1 = \alpha = \beta$ for notational simplicity (uniform prior).

Bayesian Skeptic for a coin-tossing game

- If Skeptic uses this $\mathcal{P}$, then his capital at time n is explicitly given as

$$\mathcal{K}_n = \frac{h_n! t_n!}{(n+1)! p^{h_n}(1-p)^{t_n}} = \frac{\int_0^1 p^{h_n}(1-p)^{t_n} dp}{p^{h_n}(1-p)^{t_n}}, \tag{2}$$

where $h_n = x_1 + \cdots + x_n$ ($\#$ of heads), and $t_n = n - h_n$.

- Proof is easy by induction
- This is a likelihood ratio of Bayes marginal distribution and the binomial distribution with the risk neutral probability p.
- In general, a capital process $\mathcal{K}_n$ is "always" a likelihood ratio.
- LR process is a non-negative martingale process.

KL divergence and capital process

- Stirling's formula for $x!$

$$\log x! = \left(x + \frac{1}{2}\right) \log x - x + O(1) = x \log x - x + O(\log x)$$

- Asymptotic behavior of $\log \mathcal{K}_n$

$$\begin{aligned} \log \mathcal{K}_n &= \log h_n! + \log t_n! - \log(n+1)! - h_n \log p - t_n \log(1-p) \\ &= h_n \log h_n + t_n \log t_n - n \log n - (h_n + t_n - n) \\ &\qquad - h_n \log p - t_n \log(1-p) + O(\log n) \\ &= h_n \log \frac{h_n}{np} + t_n \log \frac{t_n}{n(1-p)} + O(\log n). \end{aligned}$$

KL divergence and capital process

- The sum of the first two terms is the KL divergence.
- Hence
$$\log \mathcal{K}_n = nD\left(\frac{h_n}{n}\|p\right) + O(\log n).$$
- If h_n/n deviates from p, then Skeptic's capital $\mathcal{K}_n$ grows exponentially with the rate $D\left(\frac{h_n}{n}\|p\right)$.
- This is the "large deviation principle".

Non-negative martingales and likelihood ratios

As a standard textbook material, it can be easily checked that in the measure-theoretic framework the following two things are equivalent.

1. Non-negative martingales with expected value 1.
2. Likelihood ratios

Martingale $\Rightarrow$ LR

- Let $\mathcal{F}_n$, $n = 0, 1, 2, \ldots$ be a filtration.
- Let $\mathcal{F} = \mathcal{F}_\infty$ be the smallest σ-field containing them.
- Fix a probability measure P on $\mathcal{F}$ and let $\mathcal{K}_n$, $n = 0, 1, 2, \ldots$ be a non-negative martingale under P with $E(\mathcal{K}_n) = 1$, $\forall n$.
- Define Q_n on $\mathcal{F}_n$ by

$$Q_n(A) = \int_A \mathcal{K}_n dP, \quad A \in \mathcal{F}_n.$$

Non-negative martingales and likelihood ratios

- Then it is an easy exercise to show that Q_n's are a consistent (i.e. $Q_n(A) = Q_{n+1}(A), \ \forall A \in \mathcal{F}_n$) family of distributions and $\mathcal{K}_n$ is the likelihood ratio: $\mathcal{K}_n = dQ_n/dP_n$.

LR $\Rightarrow$ Martingale

- Let $Q_1, Q_2, \dots$ be a consistent family of probability distributions on $\mathcal{F}_n$, $n = 0, 1, \dots,$ such that each Q_n is absolutely continuous with P.
- Define
$$\mathcal{K}_n = \frac{dQ_n}{dP}.$$
- Then
$$E(\mathcal{K}_n) = \int_\Omega \frac{dQ_n}{dP} dP = \int_\Omega dQ_n = Q_n(\Omega) = 1.$$
- Furthermore it can be easily shown that $E(\mathcal{K}_{n+1}|\mathcal{F}_n) = \mathcal{K}_n$ and this is equivalent to the consistency condition.

Non-negative martingales and likelihood ratio (GTP)

- From GTP, the capital process $\mathcal{K}_n \geq 0$ is a non-negative martingale with expected value 1 under any risk neutral probability measure.
- However not all non-negative measure-theoretic martingales with expected value 1 can be realized as a capital process. It depends on how rich is the move space of Skeptic, i.e., what kind of strategies are allowed to Skeptic.
- If the game is "complete", such as the coin-tossing game, then the converse is true.

A sequential test can be constructed from betting

- Let $\mathcal{K}_n$ be a non-negative martingale with $E(\mathcal{K}_n) = 1$.
- By Markov inequality

$$P(\sup_n \mathcal{K}_n \geq 1/\alpha) \leq \alpha.$$

- Hence a sequential testing procedure with the level of significance α is constructed by rejecting the null hypothesis as soon as $\mathcal{K}_n \geq 1/\alpha$.
- Suppose that the data generating process for $X_1, X_2, \dots,$ is given as a null hypothesis. If you are allowed to bet on $X_1, X_2, \dots$ and if you can multiply your capital 20-fold, then the null hypothesis is rejected with the significance level of 5%.
- See for example, "New procedures for testing whether stock price processes are martingales" in *Computational Economics*, 2010.

A sequential test can be constructed from betting

- In this sequential setting, betting strategies need not be formal or fully specified. Any betting is OK as long as the future observations are never used (of course).
- On the other hand, when we obtain a batch sample of size n, we often have to be careful that we should have decided to use a particular procedure before seeing the actual data.
- This "hindsight" effect even exists in the maximized likelihood. Then various information criteria are needed to take the hindsight effect into consideration.
- Compared to the standard batch sample setting, use of betting in a sequential test can be more informal.

Summary of the talk

- I have discussed background for game-theoretic probability.
- I did some mathematics of Bayesian betting strategy for coin-tossing games.
- I tried to explain why likelihood ratio appears as the capital process.
- I indicated that capital process (i.e. LR) can be used as a measure of departure from the null hypothesis, leading to a simple sequential test.