Information Theory and Networks

Lecture 16: Gambling and Information Theory

Paul Tune

<paul.tune@adelaide.edu.au>

http://www.maths.adelaide.edu.au/matthew.roughan/Lecture_notes/InformationTheory/

School of Mathematical Sciences,
University of Adelaide

September 18, 2013

Part I

Gambling and Information Theory

If fighting is sure to result in victory, then you must fight, even though the ruler forbid it;
If fighting will not result in victory, then you must not fight even at the ruler's bidding.

Sun Tzu, The Art of War, Chapter 10, 23

Section 1

Horse Racing

Fixed-Odds Horse Racing

- Pool of money betting on horses
  - odds: expressed as o-for-1 or $(o-1)$-to-1
  - probability of success by probability of failure
  - assume no track take, no commissions
- What's the best strategy?
  - one-off bet
  - multiple ongoing bets, or *parlayed bets*

Example

- Here, only bet on horse win (not other bets like place etc.)
- Odds are fixed by a bookie
- We use o-for-1 convention

Horse	**Odds**
1	10
2	2
3	20
4	5

Betting Strategies

- One-off bet: all in
  - equivalent: maximizing arithmetic mean
- Parlayed bets: Kelly criterion
  - equivalent: maximizing geometric mean
- What happens with all-in for parlayed bets?
- Note: payout asymmetry most important
- **Make sure your capital survives before it can compound**

Section 2

The Kelly Criterion

Some History

- Developed by J. L. Kelly at Bell Labs; Shannon reviewed
  - Texan tough guy, gunslinger, daredevil pilot and mathematician!
- Wirelines were used to transmit information between bookies
  - application: placing bets on horses

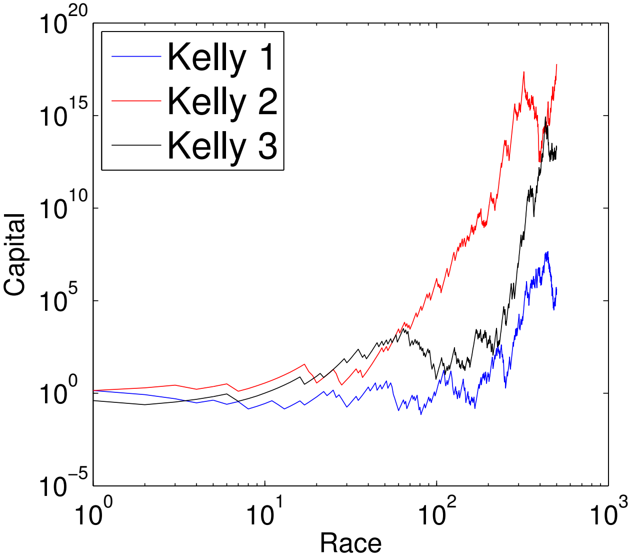
Formulation

- Assume m horses, each with i.i.d. probability of winning p_i
- Assume starting capital $S_0 = 1$
- Odds: o_i, alternative $(1 + r_i)$, r_i the rate of return
- Play for T races
  - allocate b_i fraction of capital on horse i
  - capital at T: $S_T = \prod_{t=1}^{T} \prod_{i=1}^{m} b_i o_i$
- Objective: assuming fully invested, choose allocation $b_i \geq 0$, $\sum_i b_i = 1$ to maximize S_T

Maximising Wealth Growth

- Assume $T \to \infty$
  - maximise $E[\sum_{i=1}^{m} \log b_i o_i]$ subject to constraints
  - doubling rate: $W(\mathbf{b}, \mathbf{p}) := \sum_{i=1}^{m} p_i \log b_i o_i$
- Solution: the Kelly criterion, or log-optimal wealth growth
  - answer: $b_i^\star = p_i$, proportional gambling (for fair odds)
  - solve using standard KKT conditions, or log-sum inequality
- Nature of solution will depend on odds: see [CT91, Exercise 6.2]

Example Run of Kelly's Strategy

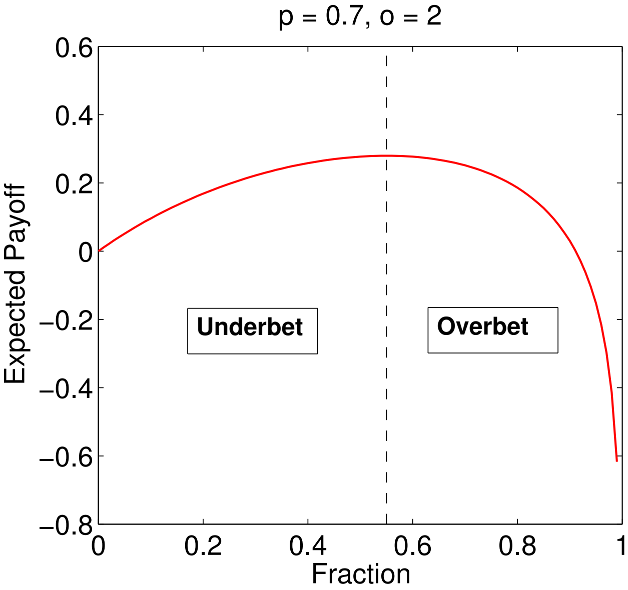


A Simple Bet

- Say a biased coin toss, win if heads, lose if tails
  - heads with probability p, q otherwise
  - each round, add \$1 to bet
- Odds: o-for-1 (remember: win-lose event)
- Kelly solution: $b^\star = \frac{op-q}{o} = \frac{p(o+1)-1}{o}$
  - what does it mean if $o = q/p$?
  - what does it mean when $b^\star < 0$ ($o < q/p$)?
  - what about $b^\star > 1$?
- A simple way to remember (for two events)

$$b^\star = \frac{\text{edge}}{\text{odds}}$$

Simple Bet: Payoff



Simple Bet: Under and Overbetting

- There is no gain in overbetting: growth decreases, risk increases
- Sweet spot: full Kelly for maximum wealth growth
- In practice, partial Kelly more applicable, i.e. αb_i^*
  - with α fraction, only α^2 volatility
  - more robust to error in estimating returns
  - lower wealth growth compared to full Kelly

Section 3

Downsides

Caveats

- Strategy is guaranteed to beat any other strategy on wealth growth
- BUT Strategy is asymptotically optimal: assume playing forever
- No guarantee to win in the short term (or at all), just the best chance
- Psychologically unsettling: imagine capital dropping 60% right before tripling!
  - partial Kelly strategies trade smoothness with growth rate
- Guaranteed not to go to ruin
  - BUT assumes capital infinitely divisible
  - capital could be 10^{-10} but hey, at least not bankrupt!
  - can show $\lim_{T \to \infty} P(S_T > \epsilon) = 0$, for any $\epsilon > 0$
- Assumes know the probability of winning: not true in real life
  - again, half Kelly strategies help: gives a safety margin
  - estimation methods (e.g. maximum entropy, shrinkage)

Criticism from Modern Finance

- Kelly criterion assumes maximizing growth rate exponent
- Called the log-utility function in finance
- **Criticism 1**: not everybody would want to maximise growth rate exponent
  - does not take into account risk-averseness (or "sleep test")
  - definition of risk in finance: volatility
  - different utilities for different folks
- **Criticism 2**: time horizon, as discussed, need very long term
- Counter-argument: not many people want to do with less money
- "Money can't buy you happiness, but love can't get you a Ferrari."

Approximation of the Stock Market

- Suppose m risky assets, each with random "odds" r_i in one investment period
- One asset with return r_0 is deterministic
- Assume starting capital $S_0 = 1$
- The return vector $\mathbf{r}$, with $\boldsymbol{\mu}_{\mathbf{r}} = E[\mathbf{r}]$, $\Sigma = E[(\mathbf{r} - r_0\mathbf{1})(\mathbf{r} - r_0\mathbf{1})^T]$
  - Σ is full rank
  - correlations apply only "spatially"
- Derive the optimal allocation $\mathbf{b}$ to optimise the wealth doubling rate
  - optimise $E[\log(r_0 + \mathbf{b}^T(\mathbf{r} - r_0\mathbf{1})]$
- Assume no constraints on $\mathbf{b}$
- For what return distribution is this allocation optimal?

Further reading I

Thomas M. Cover and Joy A. Thomas, *Elements of information theory*, John Wiley and Sons, 1991.