ECE253/CSE208 Introduction to Information Theory

Lecture 9: Gambling and Information Theory

Dr. Yu Zhang

ECE Department
University of California, Santa Cruz

- Chap 6 of *Elements of Information Theory (2nd Edition) by Thomas Cover & Joy Thomas*
- `https://en.wikipedia.org/wiki/Gambling_and_information_theory`

Theory of Gambling and Information Theory

- Information measures tell us how to take the best guess in the face of partial information.
- In that sense, information theory is a formal expression of the theory of gambling.
- The sum of the growth rate and the entropy rate is a constant.
- Financial value of side information $=$ the mutual information between the horse race and the side information.
- Horse race is a special case of stock market investment; see Chapter 16.

Doubling Rate

Gambling on a horse race:

- m: number of horses.
- p_i: the probability of horse i winning.
- b_i: the proportion of wealth bet on horse i.
- o_i: the payoff (odds) (e.g., $o_i = 2$ if horse i winning pays double the amount bet).
- $S(X) = b(X)o(X)$: the wealth relative; i.e., the factor by which the gambler's wealth grows if horse X wins.
- $S_n = \prod_{i=1}^{n} S(X_i)$: gambler's wealth after n repeated races.

Doubling rate in gambling on a horse race is

$$\boxed{W(\mathbf{b}, \mathbf{p}) = \mathrm{E}[\log_2 S(X)] = \sum_{i=1}^{m} p_i \log_2(b_i o_i).}$$

Optimum Doubling Rate

Theorem

Let the race outcomes $X_1, X_2, \ldots$ be i.i.d. $\sim p(x)$. Then, the wealth of the gambler using betting strategy $\mathbf{b}$ grows exponentially at rate $W(\mathbf{b}, \mathbf{p})$; i.e., $S_n \doteq 2^{nW(\mathbf{b},\mathbf{p})}$.

Proof: $\frac{1}{n} \log S_n = \frac{1}{n} \sum_{i=1}^{n} \log S(X_i) \to \mathrm{E}[\log S(X)]$ in probability.

Portfolio optimization:

$$\max_{\mathbf{b}} \quad W(\mathbf{b}, \mathbf{p}) = \sum_{i=1}^{m} p_i \log_2(b_i o_i) \tag{1a}$$

$$\text{s.t.} \quad b_i \geq 0,\ i = 1, 2, \ldots, m \tag{1b}$$

$$\sum_i b_i = 1 \tag{1c}$$

This is a convex optimization problem: maximize a concave function over a probability simplex.

Proportional (Kelly) Gambling

- Solution 1 – Lagrangian relaxation: $L(\mathbf{b}, \lambda) = \sum_{i=1}^{m} p_i \ln(b_i o_i) + \lambda(\sum b_i - 1)$
  Set the gradient to be zero: $\frac{\partial L}{\partial b_i} = \frac{p_i}{b_i} + \lambda = 0 \implies b_i = -\frac{p_i}{\lambda}$.
  Constraint $\sum b_i = 1 \implies \lambda = -1 \implies \mathbf{b}^* = \mathbf{p}$. Hence, the best strategy for the gambler is to bet on each horse in proportion to its probability of winning.
  Plugging $\mathbf{b}^*$ into the objective, we get the optimum doubling rate:

$$W^*(\mathbf{p}) = \sum p_i \log(o_i) - H(\mathbf{p}).$$

- Solution 2 – Information inequality:

$$W(\mathbf{b}, \mathbf{p}) = \sum_i p_i \log b_i o_i = \sum_i p_i \log\left(\frac{b_i}{p_i} p_i o_i\right) \tag{2}$$

$$= \sum_i p_i \log o_i - H(\mathbf{p}) - D(\mathbf{p}||\mathbf{b}) \tag{3}$$

$$\leq \sum_i p_i \log o_i - H(\mathbf{p}) \tag{4}$$

  with equality iff $\mathbf{b} = \mathbf{p}$.

Fair Odds

Consider a special case when odds are fair with respect to some distribution. That is, there is no track take (money deducted from each pool for track revenue and taxes) and $\sum \frac{1}{o_i} = 1$. Define $r_i = \frac{1}{o_i}$ that is the bookie's estimate of the win probabilities. Then,

$$W(\mathbf{b}, \mathbf{p}) = \sum p_i \log(b_i o_i) = \sum p_i \log\left(\frac{b_i}{p_i}\frac{p_i}{r_i}\right) = D(\mathbf{p}||\mathbf{r}) - D(\mathbf{p}||\mathbf{b}).$$

Interpretations:

1. The doubling rate is the difference between the distance of the bookie's estimate from the true distribution and the distance of the gambler's estimate from the true distribution.
2. The gambler can make money only if his estimate (as expressed by $\mathbf{b}$) is better than the bookie's.
3. For m-for-1 odds (i.e., $r_i = \frac{1}{m}$), the optimum doubling rate is

$$W^*(\mathbf{p}) = D\left(\mathbf{p}||\frac{1}{m}\right) = \log m - H(\mathbf{p}).$$

Conservation Theorem

Theorem

For uniform fair odds, we have $W^*(\mathbf{p}) + H(\mathbf{p}) = \log m$.

- Every bit of entropy decrease doubles the gambler's wealth.
- Low entropy races are the most profitable.

Partial Investment

If a gambler keeps some of wealth as cash.

- $b(0)$: the proportion of wealth held out as cash.
- $b(1), b(2), \cdots, b(m)$: the proportions bet on different horses.
- The wealth relative becomes $S(X) = b(0) + b(X)o(X)$.

Now, proportional gambling is not necessarily the optimum strategy, which may depend on the odds. We have the following 3 cases:

1. **Fair odds** ($\sum \frac{1}{\sigma_i} = 1$): the option of withholding cash does not change the analysis. Proportional betting is optimal.
2. **Superfair odds** ($\sum \frac{1}{\sigma_i} < 1$): the odds are even better than fair odds (seldom in real life), always want to put all wealth into the race rather than leave it as cash. The optimum strategy is proportional betting.
3. **Subfair odds** ($\sum \frac{1}{\sigma_i} > 1$): More representative of real life. The optimal strategy has a simple "water-filling" form that can be found using KKT conditions.

Gambling with Side Information

- Suppose the gambler has some information that is relevant to the outcome of the gamble; e.g., the past performance of the horses.
- What is the value of this side information: the financial value of such information can be defined as the increase in wealth.
- Let horse $X \in \{1, 2, \ldots, m\}$ win the race with probability $p(x)$ and pay odds of $o(x)$ for 1. Let (X, Y) have joint PMF $p(x, y)$, where Y denotes the side information.
- Conditional betting $b(x|y)$: the proportion of wealth bet on horse x when y is observed. Hence, $b(x|y) \geq 0$ and $\sum_x b(x|y) = 1$.
- The optimal unconditional and conditional doubling rates:

$$W^*(X) = \max_{\mathbf{b}(x)} \sum_x p(x) \log[b(x)o(x)] \tag{5}$$

$$W^*(X|Y) = \max_{\mathbf{b}(x|y)} \sum_{x,y} p(x, y) \log[b(x|y)o(x)] \tag{6}$$

- Wealth increase: $\Delta W = W^*(X|Y) - W^*(X)$.

Gambling with Side Information (Cont'd)

Theorem (Increase in doubling rate is equal to the mutual information)

The increase ΔW in doubling rate due to side information Y for a horse race X is

$$\Delta W = I(X;Y).$$

Proof: It can be shown that $b^*(x|y) = p(x|y) = \arg\max W^*(X|Y) \implies$

$$W^*(X|Y) = \max_{\mathbf{b}(x|y)} \sum_{x,y} p(x,y) \log[b(x|y)o(x)] \tag{7}$$

$$= \sum_{x,y} p(x,y) \log[p(x|y)o(x)] \tag{8}$$

$$= \sum_{x} p(x) \log o(x) - H(X|Y). \tag{9}$$

$$W^*(X) = \sum_{x} p(x) \log o(x) - H(X). \tag{10}$$

$$\Delta W = W^*(X|Y) - W^*(X) = H(X) - H(X|Y) = I(X;Y). \tag{11}$$

Thank You!

Email: <zhangy@ucsc.edu>

Homepage: https://people.ucsc.edu/~yzhan419/