Gambling with Information Theory

Govert Verkes

University of Amsterdam

January 27, 2016
How do you bet?

Private noisy channel transmitting results while you can still bet, correct transmission($p$) or error in transmission($q$), with $p \gg q$. 

\[
\begin{align*}
1 & \quad 1 \\
p & \quad q \quad 2 \\
& \quad \vdots \\
& \quad N
\end{align*}
\]
## Overview

- Kelly gambling
- Horse races
- Value of side information
- Entropy rate of stochastic processes
- Dependent horse races
John L. Kelly

- John Larry Kelly, Jr. (1923–1965)
- PhD in Physics
- Bell labs
- Shannon (Las Vegas)
- Warren Buffett (Investor)
Gambler with private wire

- Communication channel transmitting results
- Noiseless channel

\[ V_N = 2^N V_0 \]

\[ V_N : \text{Capital after } N \text{ bets} \]

\[ V_0 : \text{Starting capital} \]
Gambler with private wire

- Communication channel transmitting results
- Noiseless channel

\[ V_N = 2^N V_0 \]

\( V_N \): Capital after \( N \) bets
\( V_0 \): Starting capital

Exponential rate of growth

\[ G = \lim_{N \to \infty} \frac{1}{N} \log \frac{V_N}{V_0} \]
Gambler with noisy private wire

**Exponential rate of growth**

\[ G = \lim_{N \to \infty} \frac{1}{N} \log \frac{V_N}{V_0} \]

How would you bet on the received result?

- \( p \) : probability of correct transmission
- \( q \) : probability of error in transmission
Gambler with noisy private wire

**Exponential rate of growth**

\[
G = \lim_{N \to \infty} \frac{1}{N} \log \frac{V_N}{V_0}
\]

How would you bet on the received result?

- \(p\) : probability of correct transmission
- \(q\) : probability of error in transmission
- \(\ell\) : the fraction of gambler’s capital that he bets

\[
V_N = (1 + \ell)^W (1 - \ell)^L V_0
\]
Horse races

- Wealth relative

\[ S(X) = b(X) o(X) \]

- \( b(i) \): fraction of gambler’s wealth on horse \( i \)
- \( o(i) \): \( o(i) \)-for-1 odds on horse \( i \)
- \( m \): number of horses
Horse races

- **Wealth relative**

\[ S(X) = b(X) o(X) \]

- \( b(i) \): fraction of gambler's wealth on horse \( i \)
- \( o(i) \): \( o(i) \)-for-1 odds on horse \( i \)
- \( m \): number of horses

- **Wealth after \( N \) races (fraction)**

\[ S_n = \prod_{i=1}^{n} S(X_i) \]
Horse races doubling rate

**Doubling rate**

\[ W(b, p) = \mathbb{E}[\log S(X)] = \sum_{i=1}^{m} p_i \log b_i o_i \]

\[ p_i : \text{probability that horse } i \text{ wins} \]
Horse races doubling rate

**Doubling rate**

\[ W(b, p) = \mathbb{E}[\log S(X)] = \sum_{i=1}^{m} p_i \log b_i o_i \]

\( p_i \): probability that horse \( i \) wins

**Justification**

\[ \frac{1}{n} \log S_n = \frac{1}{n} \sum_{i=1}^{n} \log S(X_i) \xrightarrow{LLN} \mathbb{E}[\log S(X)] \]

\[ S_n = \prod_{i=1}^{n} S(X_i) \quad S_n = 2^{nW(b,p)} \]
Horse races doubling rate

- Maximize doubling rate

\[
W^*(p) = \max_{b: \sum b_i = 1} W(b, p) = \max_{p: \sum b_i = 1} \sum_{i=1}^{m} p_i \log b_i o_i
\]
Horse races doubling rate

- Maximize doubling rate

\[ W^*(p) = \max_{b: \sum b_i = 1} W(b, p) = \max_{p: \sum b_i = 1} \sum_{i=1}^{m} p_i \log b_i o_i \]

\[ b = p \]
Horse races doubling rate

Maximize doubling rate

\[ W^*(p) = \max_{b: \sum b_i = 1} W(b, p) = \max_{p: \sum b_i = 1} \sum_{i=1}^{m} p_i \log b_o i \]

\[ b = p \]

\[ W^*(b) = \sum p_i \log o_i - H(p) \]
Example (CT 6.1.1)

- 3 horses with 3-for-1 odds

\[ p_1 = \frac{1}{2}, \quad p_2 = p_3 = \frac{1}{4} \]

\[ o_1 = o_2 = 3 \]

how would you bet?
Example (CT 6.1.1)

- 3 horses with 3-for-1 odds

\[ p_1 = \frac{1}{2}, \quad p_2 = p_3 = \frac{1}{4} \]

\[ o_1 = o_2 = 3 \]

how would you bet?

\[ \sum p_i \log o_i - H(p) = \log 3 - H\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right) = 0.085 \]

\[ S_n = 2^{n^{0.085}} = (1.06)^n \]
Example (CT 6.1.1)

- Odds are fair with respect to some distribution

\[
\sum \frac{1}{o_i} = 1 \quad \text{and} \quad r_i = \frac{1}{o_i}
\]
Example (CT 6.1.1)

- Odds are fair with respect to some distribution

\[
\sum \frac{1}{o_i} = 1 \quad \text{and} \quad r_i = \frac{1}{o_i}
\]

\[
W(b, p) = \sum p_i \log \frac{b_i}{p_i r_i}
= D(p\|r) - D(p\|b)
\]
Gambling with side information

- We have prior information $Y$
- Conditional doubling rate

$$W^*(X) = \sum p_i \log o_i - H(p)$$

$$W^*(X|Y) = \max_{b(x|y)} \sum_{x,y} p(x,y) \log b(x|y) o(x)$$
Gambling with side information

- We have prior information $Y$
- Conditional doubling rate

$$W^*(X) = \sum p_i \log o_i - H(p)$$

$$W^*(X|Y) = \max_{b(x|y)} \sum p(x, y) \log b(x|y) o(x)$$

$$= \sum p_i \log o_i - H(X|Y)$$
Gambling with side information

- We have prior information \( Y \)
- Conditional doubling rate

\[
W^*(X) = \sum p_i \log o_i - H(p)
\]

\[
W^*(X|Y) = \max_{b(x|y)} \sum p(x, y) \log b(x|y) o(x)
\]

\[
= \sum p_i \log o_i - H(X|Y)
\]

- Increase in doubling rate

\[
\Delta W = W^*(X|Y) - W^*(X)
\]

\[
= H(X) - H(X|Y) = I(X; Y)
\]
Stochastic processes

- Sequence of random variables

\[ \{X_t\}_{t \in \mathcal{T}} \quad \text{for discrete process } \mathcal{T} = \mathbb{N} \]

\[ Pr(X_1, X_2, \ldots, X_n) \]
Stochastic processes

- Sequence of random variables
  \[
  \{X_t\}_{t \in \mathcal{T}} \quad \text{for discrete process } \mathcal{T} = \mathbb{N}
  \]
  \[
  Pr(X_1, X_2, \ldots, X_n)
  \]

- \( t \in \mathcal{T} \) is more often than not interpreted as time

- Arbitrary dependence
  \[
  Pr(X_{n+1} \mid X_1, X_2, \ldots, X_n)
  \]
Stochastic processes properties

- Markov
  \[ \Pr(X_{n+1} \mid X_1, X_2, \ldots, X_n) = \Pr(X_{n+1} \mid X_n) \]

- Stationary
  \[ \Pr(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \Pr(X_{1+t} = x_1, X_{2+t} = x_2, \ldots, X_{n+t} = x_n) \]
Stochastic processes properties

▶ Markov

\[
Pr(X_{n+1} \mid X_1, X_2, \ldots, X_n) = Pr(X_{n+1} \mid X_n)
\]

▶ Stationary

\[
Pr(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = Pr(X_{1+t} = x_1, X_{2+t} = x_2, \ldots, X_{n+t} = x_n)
\]

Example: Simple random walk

\[
Y = \begin{cases} 
1 & \text{with Pr: } \frac{1}{2} \\
-1 & \text{with Pr: } \frac{1}{2}
\end{cases}
\]

\[
X_n = \sum_{i=1}^{n} Y_i
\]

Stationary? Markov?
Entropy rate

**Definition**

\[
H(X) = \lim_{n \to \infty} \frac{1}{n} H(X_1, X_2, \ldots X_n)
\]
Examples

- $X$ is i.i.d

\[
H(X_1, X_2, \ldots, X_n) = nH(X_1) \\
H(X) = H(X_1)
\]
Examples

- $X$ is i.i.d

$$H(X_1, X_2, \ldots, X_n) = nH(X_1)$$

$$H(X) = H(X_1)$$

- $X$ independent but not identically distributed

$$H(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} H(X_i)$$
Entropy rate, related quantity

Definition

\[ H(\mathcal{X}) = \lim_{n \to \infty} \frac{1}{n} H(X_1, X_2, \ldots, X_n) \]

\[ H'(\mathcal{X}) = \lim_{n \to \infty} H(X_n \mid X_1, X_2, \ldots, X_{n-1}) \]

- For stationary processes

\[ H(\mathcal{X}) = H'(\mathcal{X}) \]
Dependent horse races

- Horse race is dependent on past performance of horses

\[ \{X_n\} : \text{Sequence of horse race outcomes} \]
Dependent horse races

- Horse race is dependent on past performance of horses

\[ \{X_n\} : \text{Sequence of horse race outcomes} \]

\[ W^*(X_n|X_{n-1}, X_{n-2}, \ldots, x_1) = \log m - H(X_n|X_{n-1}, X_{n-2}, \ldots, X_1) \]

\[ W = \log m - H(\mathcal{X})S_n \]
**Stock market**

\[ X = (X_1, X_2, \ldots, X_n) : \text{Stock vector} \]

\[ b = (b_1, b_2, \ldots, b_n) : \text{Investment vector (portfolio)} \]

\[ S = b^tX \quad : \text{Money gained after one day} \]

\[ X \sim F(x) \quad : \text{joint distribution of vector prices} \]

\[ W(b, F) = \int \log b^t \! x \, dF(x) \]

\[ W \ast (F) = \max_b W(b, F) \]
Conclusion

- Optimal betting strategy not always highest expected value
Conclusion

- Optimal betting strategy not always highest expected value
- Proportional betting is the way to go for fair odds
Conclusion

- Optimal betting strategy not always highest expected value
- Proportional betting is the way to go for fair odds
- Stock market interesting ([CT] chapter 15)
References

- Thomas M. Cover, Joy A. Thomas. "Elements of information theory"
- J. L. Kelly, Jr., "A New Interpretation of Information Rate"

Questions?