Gambling with Information Theory

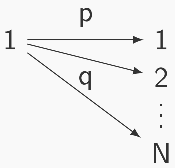
Govert Verkes

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 $p \gg q$.



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

$$1 \longrightarrow 1$$

$$\vdots$$

$$N \longrightarrow N$$

$$V_N = 2^N V_0$$

V_N : Capital after N bets

V_0 : Starting capital

Gambler with private wire

- Communication channel transmitting results
- Noiseless channel

$$1 \longrightarrow 1$$

$$\vdots \qquad\qquad V_N = 2^N V_0$$

$$N \longrightarrow N$$

V_N : Capital after N bets

V_0 : Starting capital

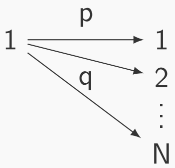
Exponential rate of growth

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

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?

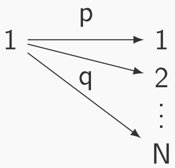
p : probability of correct transmission

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

ℓ : 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

$$S(X) = b(X)o(X)$$

$$b(i) : \text{fraction of gambler's wealth on horse } i$$

$$o(i) : o(i)\text{-for-1 odds on horse } i$$

$$m : \text{number of horses}$$

- Wealth after N races (fraction)

$$S_n = \prod_{i=1}^{n} S(X_i)$$

Horse races doubling rate

Doubling rate

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

p_i : probability that horse i wins

Horse races doubling rate

Doubling rate

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

p_i : probability that horse i wins

Justification

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

$$S_n = \prod_{i=1}^{n} S(X_i) \qquad S_n = 2^{nW(\mathbf{b},\mathbf{p})}$$

Horse races doubling rate

- Maximize doubling rate

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

Horse races doubling rate

- Maximize doubling rate

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

$$\mathbf{b} = \mathbf{p}$$

Horse races doubling rate

- Maximize doubling rate

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

$$\mathbf{b} = \mathbf{p}$$

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

Example (CT 6.1.1)

- 3 horses with 3-for-1 odds

$$p_1 = \frac{1}{2},\ p_2 = p_3 = \frac{1}{4}$$
$$o_1 = o_2 = 3$$

how would you bet?

Example (CT 6.1.1)

- 3 horses with 3-for-1 odds

$$p_1 = \frac{1}{2},\ p_2 = p_3 = \frac{1}{4}$$
$$o_1 = o_2 = 3$$

how would you bet?

$$\sum p_i \log o_i - H(\mathbf{p}) = \log 3 - H(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}) = 0.085$$

$$S_n = 2^{n0.085} = (1.06)^n$$

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

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

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

Gambling with side information

- We have prior information Y
- Conditional doubling rate

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

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

Gambling with side information

- We have prior information Y
- Conditional doubling rate

$$
\begin{aligned}
W^*(X) &= \sum p_i \log o_i - H(\mathbf{p}) \\
W^*(X|Y) &= \max_{\mathbf{b}(x|y)} \sum_{x,y} p(x,y) \log b(x|y) o(x) \\
&= \sum p_i \log o_i - H(X|Y)
\end{aligned}
$$

Gambling with side information

- We have prior information Y
- Conditional doubling rate

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

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

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

- 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

$$\begin{aligned}\Pr(X_1 &= x_1, X_2 = x_2, \ldots, X_n = x_n) \\ &= \Pr(X_{1+t} = x_1, X_{2+t} = x_2, \ldots, X_{n+t} = x_n)\end{aligned}$$

Stochastic processes properties

- Markov

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

- Stationary

$$\begin{aligned}\Pr(X_1 &= x_1, X_2 = x_2, \ldots, X_n = x_n) \\ &= \Pr(X_{1+t} = x_1, X_{2+t} = x_2, \ldots, X_{n+t} = x_n)\end{aligned}$$

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

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

Examples

- X is i.i.d

$$H(X_1, X_2, \ldots, X_n) = nH(X_1)$$
$$H(\mathcal{X}) = H(X_1)$$

Examples

- X is i.i.d

$$\begin{aligned} H(X_1, X_2, \ldots, X_n) &= nH(X_1) \\ H(\mathcal{X}) &= H(X_1) \end{aligned}$$

- 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

$$\{X_n\} : \text{Sequence of horse race outcomes}$$

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

$$W = \log m - H(\mathcal{X})S_n \quad =$$

Stock market

$X = (X_1, X_2, \dots, X_n)$: Stock vector

$b = (b_1, b_2, \dots, b_n)$: Investment vector (portfolio)

$S = \mathbf{b}^t\mathbf{X}$: Money gained after one day

$X \sim F(\mathbf{x})$: joint distribution of vector prices

$$W(\mathbf{b}, F) = \int \log \mathbf{b}^t\mathbf{x}\; dF(\mathbf{x})$$

$$W*(F) = \max_{\mathbf{b}} W(\mathbf{b}, F)$$

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?