Gambling and Information Theory

Giulio Bertoli

UNIVERSITEIT VAN AMSTERDAM

December 17, 2014

Overview

Some Facts

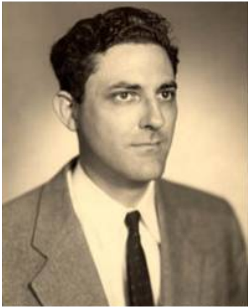
- Shannon (1948): definitions/concepts based on coding
- In following years: information without coding?
- J. L. Kelly (1956): paper "A new interpretation of information rate" on Bell Sys. Tech. Journal*

* B. S. T. J., 35 (4): 917-926, Mar., 1956

John Larry Kelly

- 1923 Corsicana TX
- 1953 - PhD in Physics, then Bell Labs
- 1956 - Kelly Gambling
- 1961 - Speech Synthesis
- 1965 NY †

- Remarkable character:
  gunslinger, stuntman pilot...
- Never profited of his
  findings on gambling (Shannon did!)



Kelly Gambling

Let's bet!

Take a single horse race with three horses, with probability of winning $\left(\frac{1}{6}, \frac{1}{2}, \frac{1}{3}\right)$ respectively.

You can bet any fraction of your capital on any horse and place simultaneous bets, but you must bet all of it.

How would you bet?

Kelly Gambling

Now, let's take the case where every Saturday there's such a horse race.

How does your betting strategy change?

Kelly Gambling

Now, let's take the case where every Saturday there's such a horse race.

How does your betting strategy change?

If you ALWAYS bet on horse 2, you'll go broke!

Most intuitive way: bet according to probabilities.

Kelly Gambling

Let's formalize this, follow Kelly's article (1956).

- Gambler with private wire: channel transmits results on binary bet BEFORE they become public.
    - Noisless binary channel
    - Noisy binary channel
- General case

Gambler with private wire - Noiseless

Gambler sure of winning → bets all his money.

Consider 2-for-1 bet. After N bets, he's got $V_N = 2^N$ times his initial money V_0.

Define the *exponential rate of growth*:

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

In our case, $G = 1$.

Gambler with private wire - Noisy

This time, there's probability of error p (correct transmission with probability $q = 1 - p$).

If gambler bets all his money every time, he will be broke for N large enough!

He should bet a fraction, f, of his money. We have:

$$V_N = (1 + f)^W (1 - f)^L V_0$$

Gambler with private wire - Noisy

Compute G using $V_N = (1+f)^W (1-f)^L V_0$:

$$\begin{aligned} G &= \lim_{N\to\infty} \left[\log \left(\frac{(1+f)^W (1-f)^L V_0}{V_0} \right) \right] \\ &= \lim_{N\to\infty} \left(\frac{W}{N} \log(1+f) + \frac{L}{N} \log(1-f) \right) \\ &= q \log(1+f) + p \log(1-f) \end{aligned}$$

Want money? Maximize G!

Gambler with private wire - Noisy

Maximize G w.r.t. f, using concavity of log or Lagrange multipliers. You get the relations

$$1 + f = 2q$$
$$1 - f = 2p$$

Which give you:

$$G_{max} = 1 + p \log p + q \log q$$

General case

Notation

Consider case where we bet on input symbols, representing outcome of chance events.

Now channel has several inputs x with probability of transmission $p(x)$ and several outputs y with probability of reception $q(y)$. The joint probability is $p(x, y)$.

Let's call $b(x|y)$ the fraction of the gambler's capital that he decides to bet on x after he receives y.

o_x are the odds paid to the gambler for a 1-dollar bet if x wins.

Horse Races with no channel

But first, let's consider no channel at all. Then we simply have a horse race of which we know nothing except the probabilities.

What is G?

Use now $V_N = \prod b(x) o(x)^W V_0$:

$$G = \lim_{N \to \infty} \frac{1}{N} \log \frac{V_N}{V_0}$$
$$= \sum p(x) \log[b(x) o(x)]$$

Horse Races

Again, seek to maximize G. Does Kelly gambling work? YES! (Theorem 6.1.2 in CT, *Kelly gambling is log-optimal*)

$$\begin{aligned} G &= \sum p(x)\log[b(x)o_x] \\ &= \sum p(x)\log[\frac{b(x)}{p(x)}p(x)o_x] \\ &= \sum p(x)\log[o_x] - H(p) - D(p||b) \\ &\leq \sum p(x)\log[o_x] - H(p) \end{aligned}$$

Where equality holds iff $p = b$. QED

Interpretation of result

Take fair horse race, where $\sum \frac{1}{o_x} = 1$. The bookie's estimate is given by $r_x = 1/o_x$, seen as probability distribution. We note:

$$\begin{aligned} G &= \sum p(x) \log[b(x) o_x] \\ &= \sum p(x) \log[\frac{b(x)}{p(x)} \frac{p(x)}{r(x)}] \\ &= D(p||r) - D(p||b) \end{aligned}$$

This means that we can make money only if our estimate (entropy distance) is better (less) than the bookie's!

Horse Races - with channel

Back to case with channel. Consider the most general case with odds $\sum \frac{1}{o_x} = 1$. Now we have:

$$\begin{aligned} G_{max} &= \sum_{x,y} p(x,y) \log[b(x|y) o_x] \\ &= \sum_{x,y} p(x,y) \log[b(x|y)] + \sum_{x} p(x) \log o_x \\ &= \sum_{x} p(x) \log o_x - H(X|Y) \end{aligned}$$

Where in the last line we maximize setting $b(x) = p(x)$.

Mutual Information

Compare this to case without channel. There $G = \sum_x p(x) \log o_x - H(X)$. This results in Theorem 6.2.1 of CT:

The increase in G due to side information Y for a horse race X is given by the mutual information $I(X;Y)$.

Proof: just compare previously obtained results!

$$\begin{aligned}
\Delta G &= G_{\textit{with side info}} - G_{\textit{without side info}} \\
&= \sum_x p(x) \log o_x - H(X|Y) - \left(\sum_x p(x) \log o_x - H(X) \right) \\
&= H(X) - H(X|Y) = I(X;Y)
\end{aligned}$$

QED

Example: 6.15 of CT

Let X be the winner of a fair horse race ($o_x = 1/p(x)$). $b(x)$ is the bet on horse x as usual. What is the optimal growth rate G?

Example: 6.15 of CT

Let X be the winner of a fair horse race ($o_x = 1/p(x)$). $b(x)$ is the bet on horse x as usual. What is the optimal growth rate G?

$$\begin{aligned} G &= \sum p(x) \log[b(x) o_x] \\ &= \sum p(x) \log[1] \\ &= 0 \end{aligned}$$

Example: 6.15 of CT

Suppose now we know that $Y = 1$ if $X = 1, 2$, and $Y = 0$ otherwise. What is then G?

Example: 6.15 of CT

Suppose now we know that $Y = 1$ if $X = 1, 2$, and $Y = 0$ otherwise. What is then G?

$$\begin{aligned} G = 0 + I(X;Y) &= H(Y) - H(Y|X) \\ &= H(Y) \\ &= H(p(1) + p(2)) \end{aligned}$$

Summing up & Outlook

- Gambling and Inf Theory have a lot in common
- If there's no track take, Kelly gambling is the way
- The maximum exponential rate of growth G is larger than it would have been with no channel by an amount equal to $I(X;Y)$.
- This was first glimpse of subfield; nowadays applied to stock market.