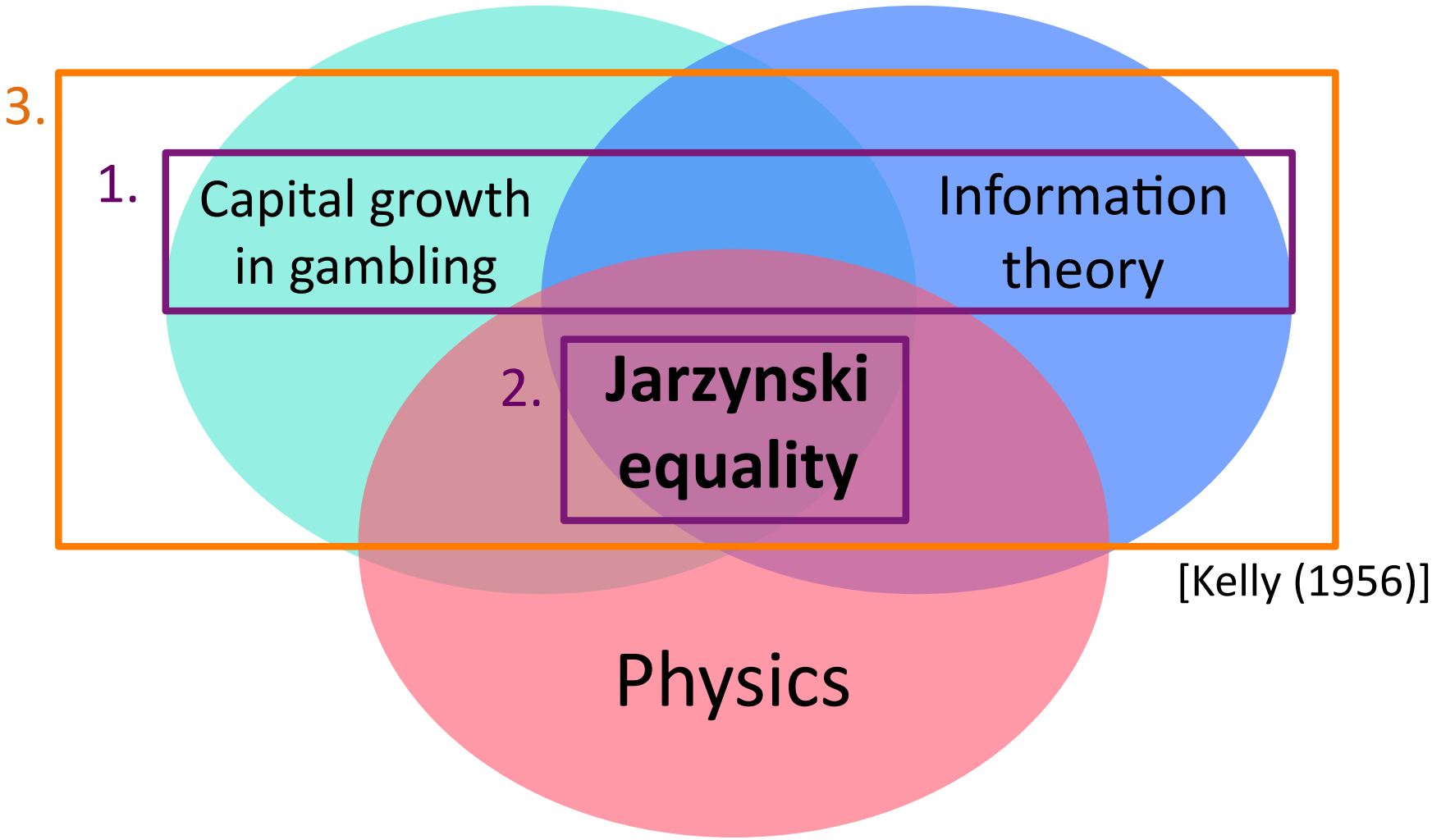
Role of information in gambling

Yuji Hirono [Stony Brook Univ.]

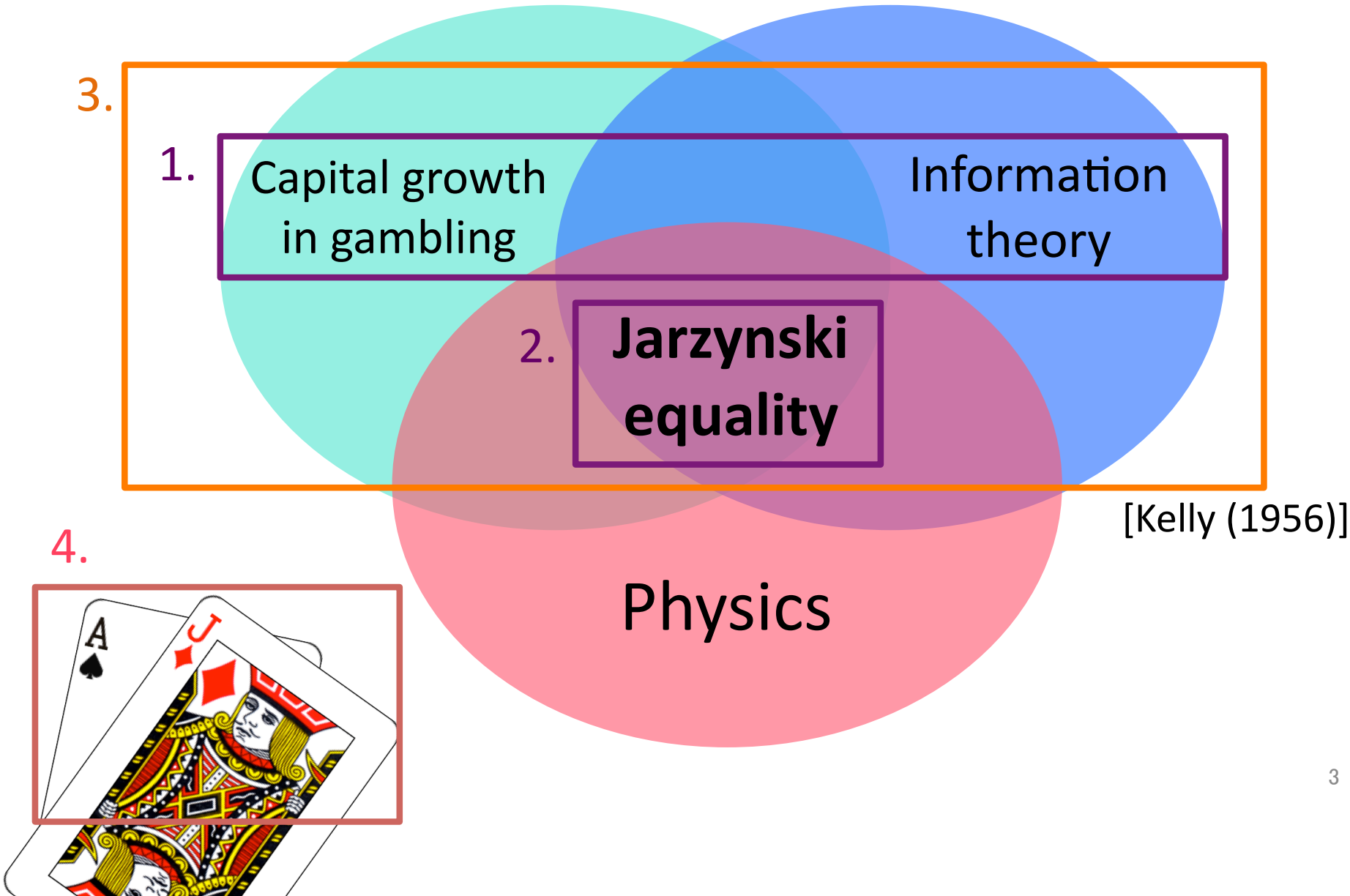
In collaboration with Y. Hidaka [RIKEN]

Why is gambling interesting?



[Kelly (1956)]

Why is gambling interesting?



[Kelly (1956)]

Gambling

- Control parameters
  - Which horse?
  - How much money we bet?
- Results are affected by:
  - Memory effects
  - “Side information”
    - weather, surface, …

Biased coin tossing

head

tail

Win

Lose

$$p \geq 1/2$$

$$1 - p \equiv \bar{p}$$

Biased coin tossing

Capital evolution f : fraction of money to bet

$$M_0 \xrightarrow{\text{Win}} M_0 - fM_0 + 2fM_0 = M_0(1+f)$$

$$M_0 \xrightarrow{\text{Lose}} M_0 - fM_0 = M_0(1-f)$$

$$M_1 = M_0(1 + fy_1)$$

$$y_1 \in \{1, -1\}$$

Biased coin tossing

Capital evolution $\quad f$: fraction of money to bet

$$M_0 \xrightarrow{\text{Win}} M_0 - fM_0 + 2fM_0 = M_0(1+f)$$

$$M_0 \xrightarrow{\text{Lose}} M_0 - fM_0 = M_0(1-f)$$

$$M_n = M_0 \prod_i (1 + \boxed{f_i} y_i)$$

$$y_i \in \{1, -1\}$$

How should a gambler choose f ?

Kelly criterion

Capital growth rate

$$g_n(f^n) \equiv \frac{1}{n} \ln \frac{M_n}{M_0} \qquad f^n \equiv \{f_1, \cdots, f_n\}$$

Choose f to maximize $\langle g_n \rangle$: "Kelly criterion" [Kelly (1956)]

Surpasses other strategies in the long run [Breiman (1961)]

Kelly criterion

When all the tosses are independent & identical

$$\max_{f^n} \langle g_n(f^n) \rangle = \max_{f_1} \langle g_1(f_1) \rangle = \max_{f_1} \langle \ln(1 + f_1 y_1) \rangle$$

The player wants to maximize

$$\begin{aligned} \langle g(f) \rangle &= \langle \ln(1 + f y) \rangle \\ &= p \ln(1 + f) + \bar{p} \ln(1 - f) \end{aligned}$$

Kelly criterion

$$\frac{\mathrm{d}}{\mathrm{d}f}\langle g(f)\rangle = 0 \quad \Rightarrow \quad \boxed{f^* = p - \bar{p}}$$

$$\boxed{\langle g(f)\rangle \leq \langle g(f^*)\rangle = \ln 2 - S(Y)}$$

$$S(Y) = S_2(p) = -p\ln p - \bar{p}\ln\bar{p}$$

Channel capacity of
a binary symmetric channel [Kelly (1956)]

Concepts in information theory

Shannon entropy $$S(X) \equiv -\sum_x P(x) \ln P(x)$$

measure of uncertainty

Mutual information $$I(X:Y) \equiv \sum_{x,y} P(x,y) \ln \frac{P(x,y)}{P(x)P(y)}$$

measure of correlation

Channel capacity $$C \equiv \max_{P(x)} I(X:Y)$$

amount of information sent reliably

sender $\longrightarrow$ receiver

x $\qquad$ y

Biased coin tossing with side information

head

Win

p

tail

Lose

$1 - p \equiv \bar{p}$

Informant

x

"side information"

Side information boosts the capital growth

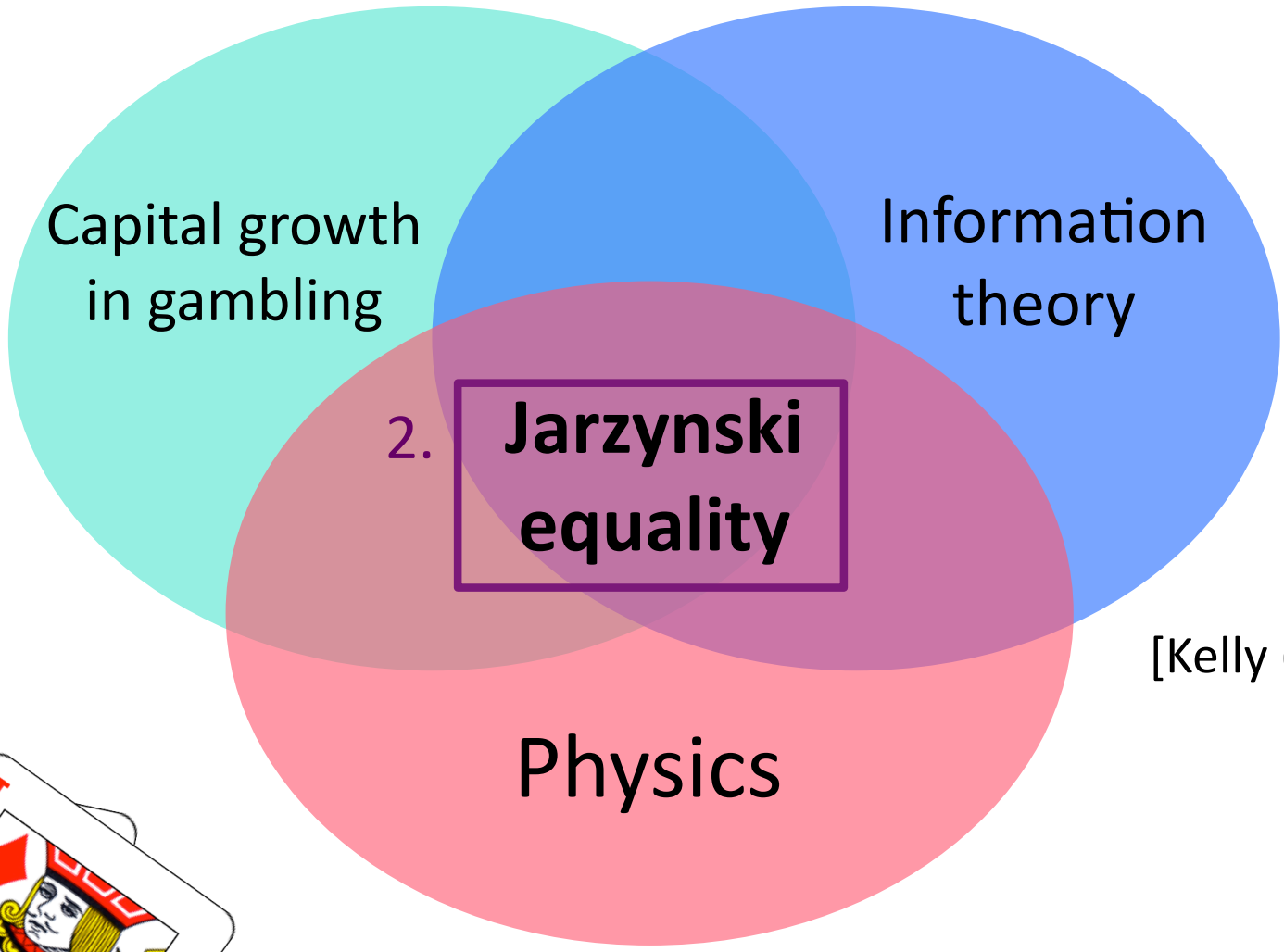
Optimal fraction $f_x^* = P(y=1|x) - P(y=-1|x)$

Maximum average capital growth rate

$$\langle g(f_x^*)\rangle_{x,y} = \sum_{x,y} P(x,y) \ln 2P(y|x)$$

$$= \ln 2 - S(Y) + I(X:Y)$$

Outline



[Kelly (1956)]

Jarzynski equality

$$\boxed{\left\langle e^{-\sigma} \right\rangle = 1}$$

[Jarzynski(1997)]

σ : entropy production

- Holds for the processes far from equilibrium
- 2nd law of thermodynamics can be derived

Jensen's inequality $\exp\left[\langle F \rangle\right] \leq \langle \exp F \rangle$

$$\Longrightarrow \langle \sigma \rangle \geq 0$$

Jarzynski equality for systems with feedback controls (demons)

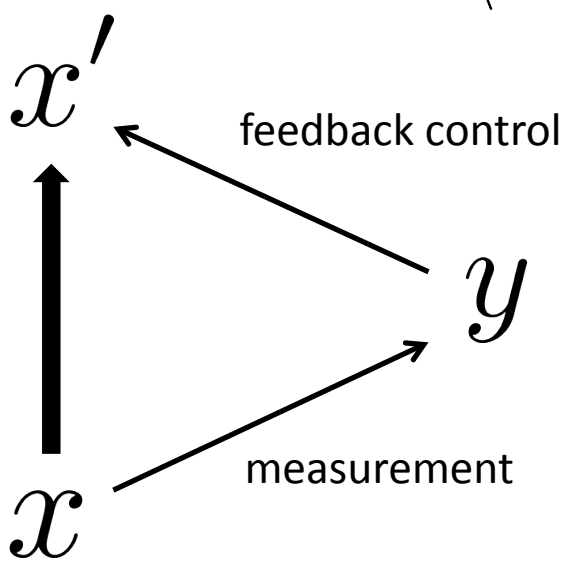
$$\boxed{\left\langle e^{-[\sigma_X + \Delta I]} \right\rangle = 1}$$

[Sagawa & Ueda(2010,2012)]

σ_X : entropy production

ΔI :change in the mutual information

$$\langle \Delta I \rangle \equiv I(X' : Y) - I(X : Y)$$



model of "Maxwell's demon"

Jarzynski equality for systems with feedback controls (demons)

$$\boxed{\left\langle e^{-[\sigma_X + \Delta I]} \right\rangle = 1}$$

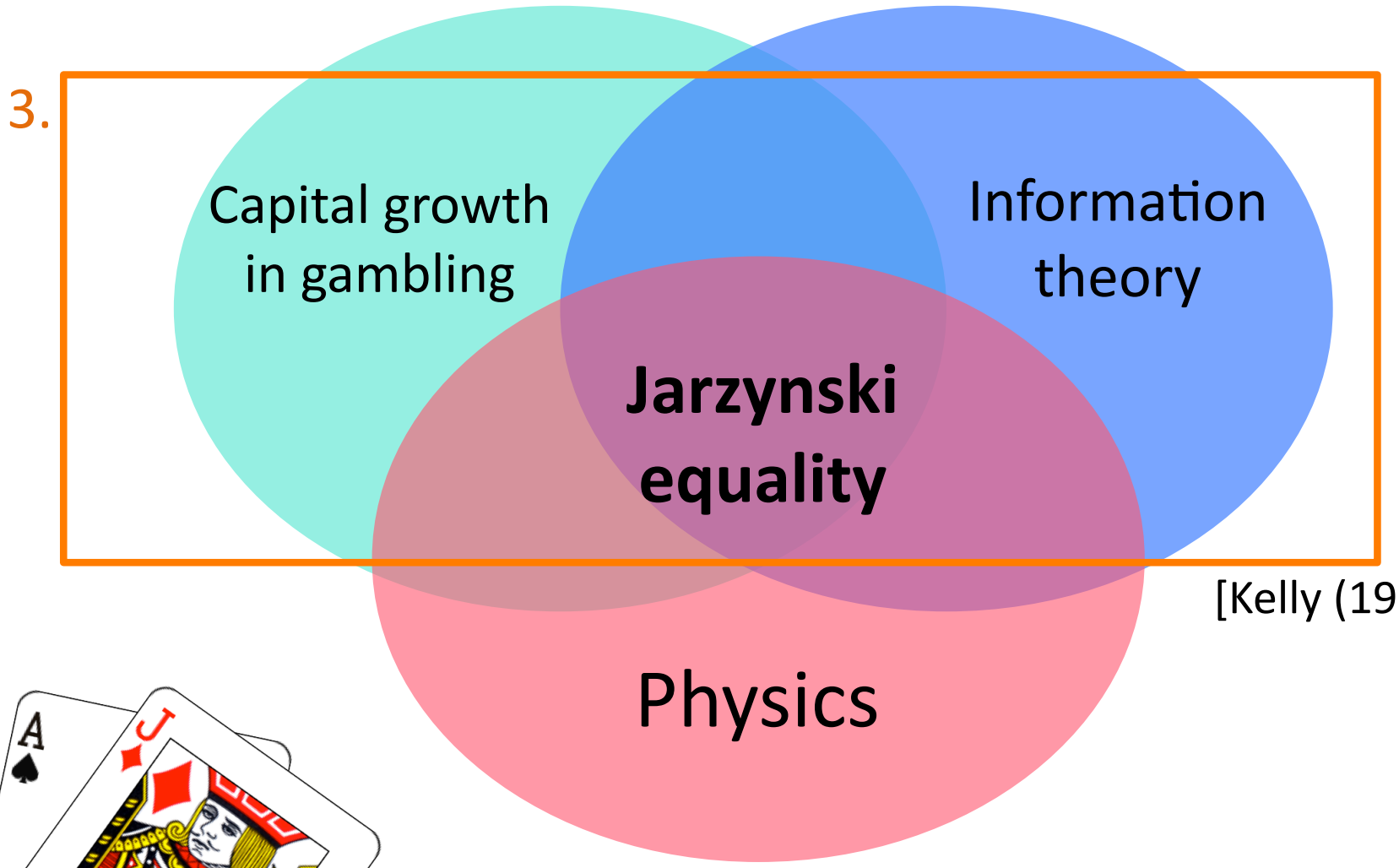
[Sagawa & Ueda(2010,2012)]

σ_X : entropy production ΔI :change in the mutual information

Jensen's inequality $\exp\left[\langle F \rangle\right] \leq \langle \exp F \rangle$

$$\Longrightarrow \langle \sigma_X \rangle \geq -\langle \Delta I \rangle$$

Outline



[Kelly (1956)]

Jarzynski-type equality for biased coin toss

$$\boxed{\langle \exp\left[g(f) + s_y\right] \rangle = 2}$$

[Hirono & Hidaka (2015)]

$$s_y \equiv -\ln P(y)$$

$\langle s_y \rangle = S(Y)$: Shannon entropy

Jensen's inequality $\exp\left[\langle F \rangle\right] \leq \langle \exp F \rangle$

$$\Longrightarrow \quad \langle g(f) \rangle_y \leq \ln 2 - S(Y)$$

Binary gambling with side information

$$\boxed{\langle \exp\left[g(f_x) + s_y - i_{xy}\right]\rangle = 2}$$

[Hirono & Hidaka (2015)]

$$i_{xy} \equiv \ln \frac{P(x,y)}{P(x)P(y)} \qquad \langle i_{xy} \rangle = I(X:Y)$$

$$\Longrightarrow \quad \langle g(f_x) \rangle \leq \ln 2 - S(Y) + I(X:Y)$$

Binary gambling with memory effects

Result on i-th game depends on the past $$P(y_i|y^{i-1})$$

Capital growth rate after n games $$g_n(f^n) \equiv \frac{1}{n} \ln \frac{M_n}{M_0}$$

Betting fraction on i-th game $$f_i = f_i(y^{i-1})$$

$$y^{i-1} = \{y_1, \cdots, y_{i-1}\}$$

Binary gambling with memory effects

$$\boxed{\langle \exp\left[n g_n(f^n) + s_{y^n}\right] \rangle = 2^n}$$

[Hirono & Hidaka (2015)]

$$s_{y^n} = -\ln P(y^n)$$

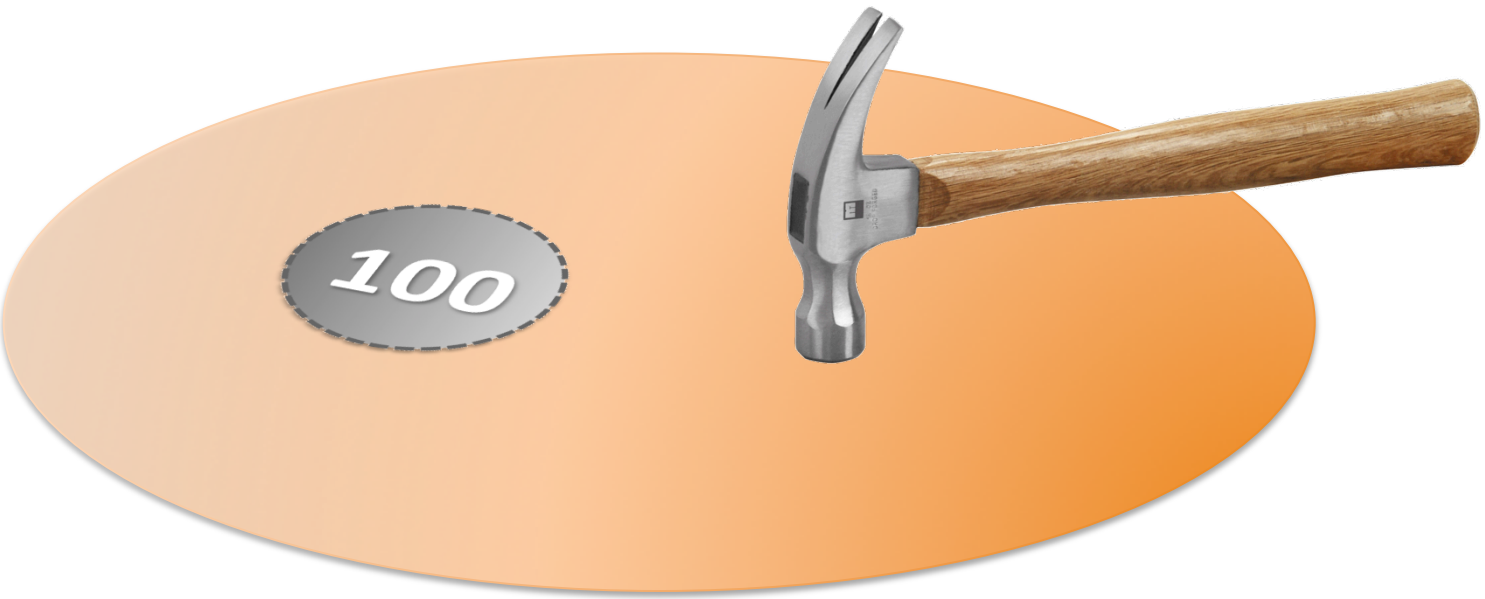
$$\langle s_{y^n} \rangle = S(Y^n)$$: Shannon entropy

$$\Longrightarrow \langle g_n(f^n) \rangle_{y^n} \leq \ln 2 - \frac{1}{n} S(Y^n)$$

Equality is attained by

$$f_i^*(y^{i-1}) = P(y_i = 1 | y^{i-1}) - P(y_i = -1 | y^{i-1})$$

Markovian coin tossing



Win

Lose

Markovian coin tossing

State of a coin after i-th hit $y_i \in \{\underset{\text{lose}}{-1}, \underset{\text{win}}{1}\}$

Conditional prob.

$$P(y_i = 1 | y_{i-1} = -1) = P(y_i = -1 | y_{i-1} = 1) = \epsilon$$

$$P(y_i = 1 | y_{i-1} = 1) = P(y_i = -1 | y_{i-1} = -1) = \bar{\epsilon}$$

ϵ : flipping probability

$$P(y_i = 1) = P(y_i = -1) = 1/2$$

Markovian coin tossing

The fraction can depend on y_{i-1} $\quad f_i = f_i(y_{i-1})$

Optimal fraction

$$f_i^*(y_{i-1}) = P(y_i = 1|y_{i-1}) - P(y_i = -1|y_{i-1})$$

Maximum capital growth rate

$$\langle g_n(f^{*n})\rangle_{y^n} = \ln 2 - S_2(\epsilon)$$

This model can be mapped to the 1D Ising model

Binary gambling with memory effects and side information

$$\boxed{\langle \exp\left[n g_n(f^n) + s_{y^n} - i_{x^n \to y^n}\right] \rangle = 2^n}$$

[Hirono & Hidaka (2015)]

$$s_{y^n} = -\ln P(y^n) \qquad i_{x^n \to y^n} = \sum_i \ln \frac{P(y_i | y^{i-1}, x^i)}{P(y_i | y^{i-1})}$$

$$\langle s_{y^n} \rangle = S(Y^n)$$: Shannon entropy

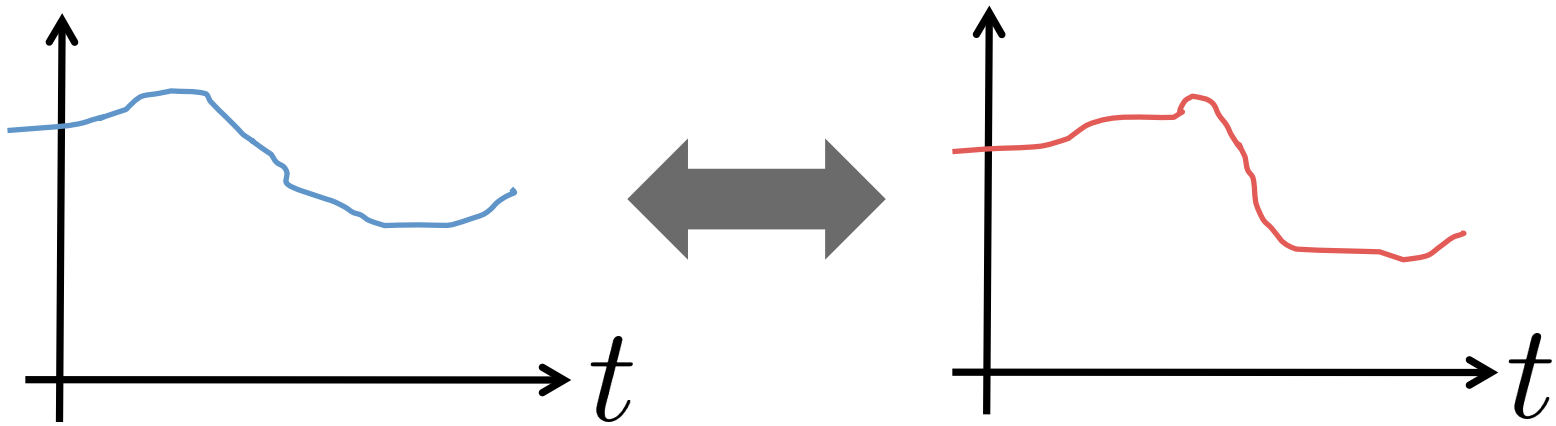
$$\langle i_{x^n \to y^n} \rangle = I_{\mathrm{dr}}(X^n \to Y^n)$$: Directed information

$$\Rightarrow \langle g_n(f^n) \rangle_{x^n, y^n} \leq \ln 2 - \frac{1}{n} S(Y^n) + \frac{1}{n} I_{\mathrm{dr}}(X^n \to Y^n)$$

Directed information [Massey (1990)]

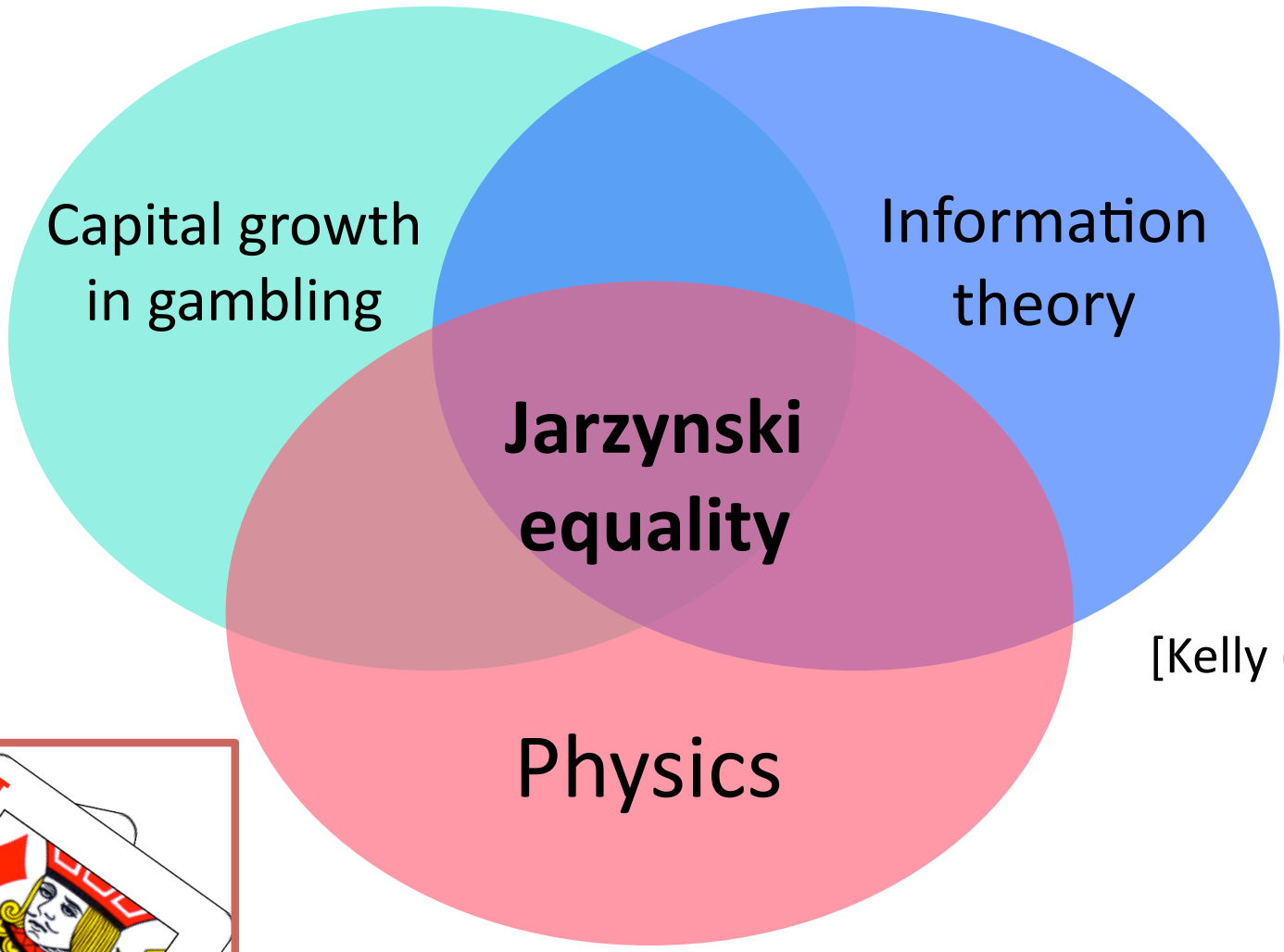
$$I_{\mathrm{dr}}(X^n \to Y^n) = \left\langle \sum_i \ln \frac{P(y_i|y^{i-1}, x^i)}{P(y_i|y^{i-1})} \right\rangle$$

quantifies *causal* correlations



used in various fields (neuroscience, currency exchange, …)

Outline



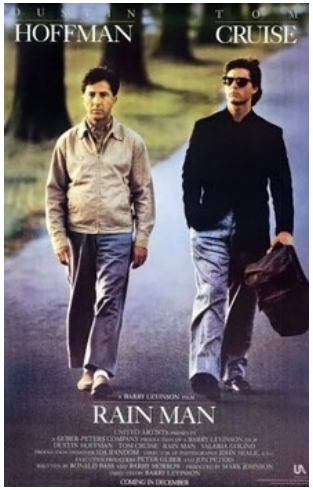
[Kelly (1956)]

4.

Blackjack

Players can beat the casino by using information of dealt cards

card counting [E. Thorp (1960)]



[“Rain man” (film) (1988)] [“21” (film) (2008)]

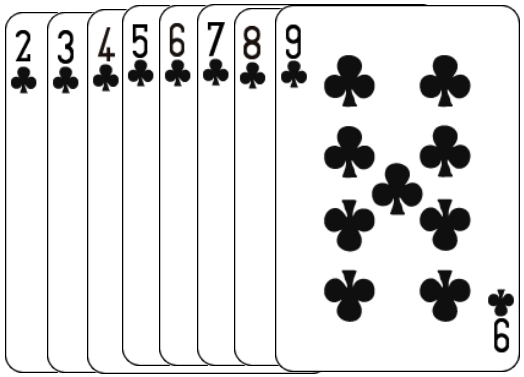
How much money can we make by counting?

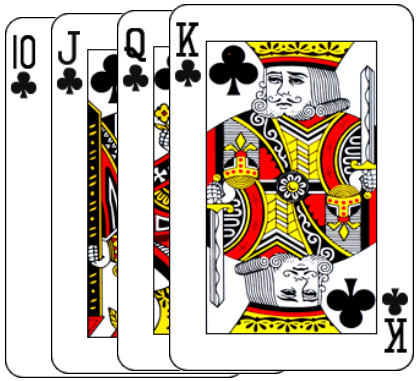
Rule of blackjack

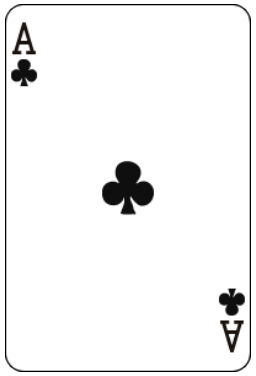
Dealer vs player. Other players are irrelevant

Purpose: reach a final score closer to 21 without <u>exceeding 21</u>

"go bust"



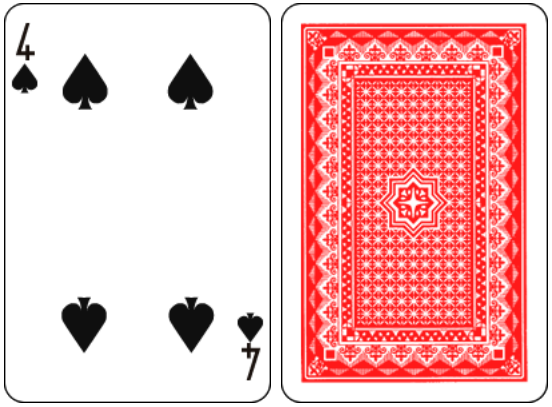


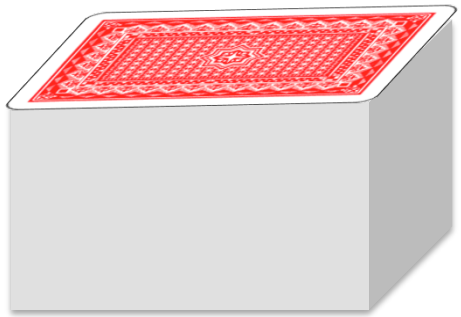


Score	number on the card	10	1 or 11

Rule of blackjack

Dealer

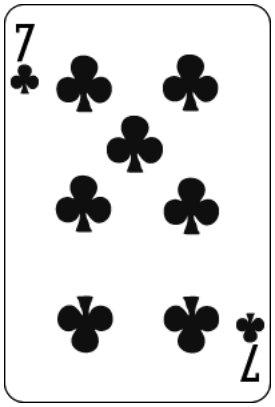




“shoe”

$\simeq$ 6 decks

Player

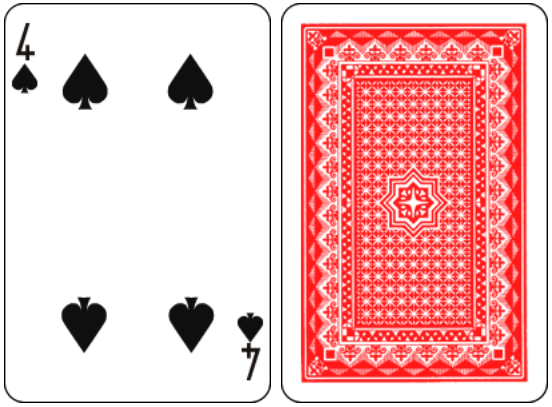


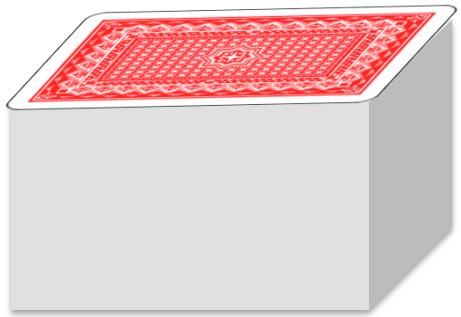
“hit” draw a card

“stand” end one’s turn

Rule of blackjack

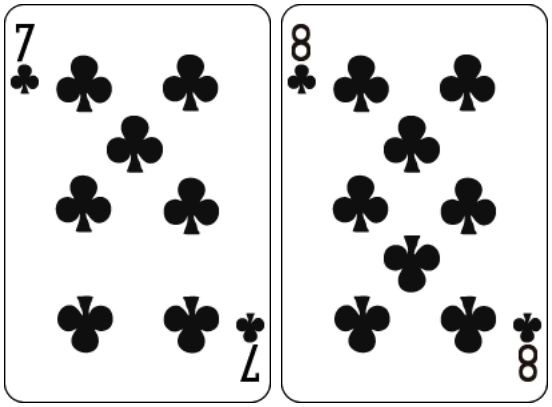
Dealer





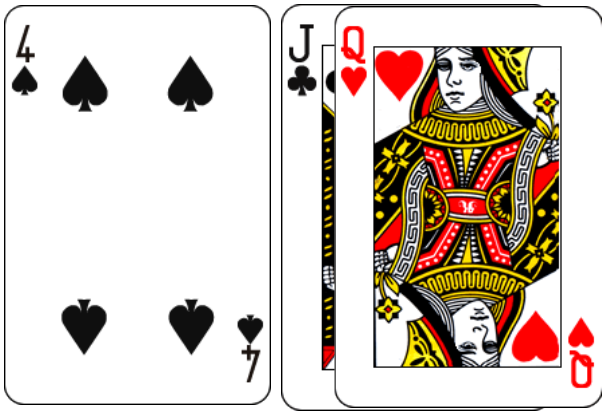
"shoe"

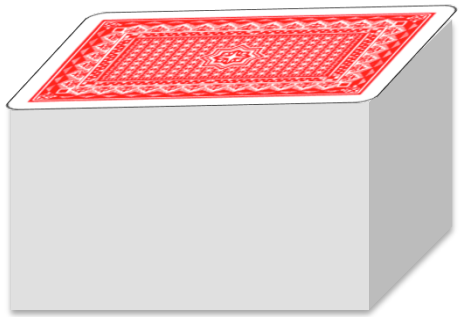
Player



Rule of blackjack

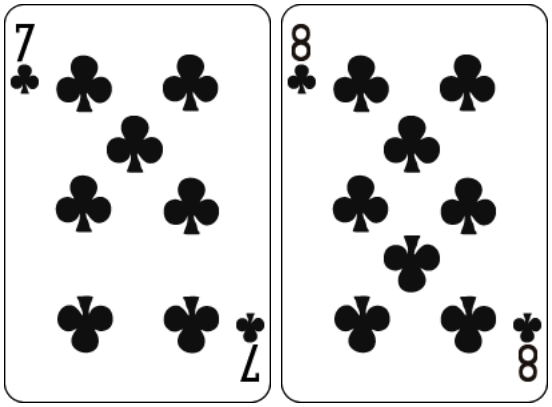
Dealer





"shoe"

Player



Dealer has to hit until (score) > 16

Dealer's action is completely specified by rules!

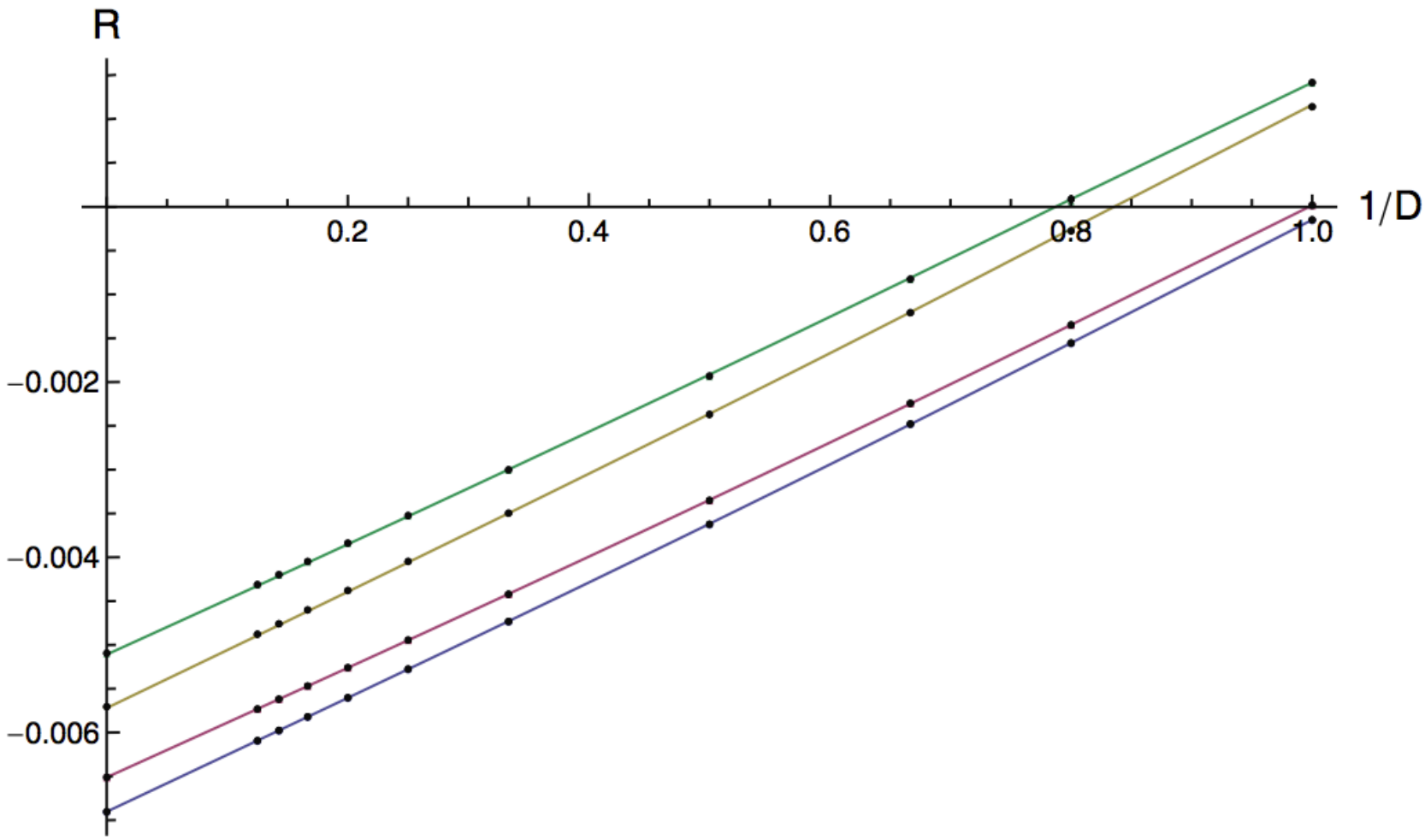
“Basic strategy”

Player hand	Dealer's face-up card									
	2	3	4	5	6	7	8	9	10	A
Hard totals (excluding pairs)										
17–20	S	S	S	S	S	S	S	S	S	S
16	S	S	S	S	S	H	H	SU	SU	SU
15	S	S	S	S	S	H	H	H	SU	H
13–14	S	S	S	S	S	H	H	H	H	H
12	H	H	S	S	S	H	H	H	H	H
11	Dh	Dh	Dh	Dh	Dh	Dh	Dh	Dh	Dh	H
10	Dh	Dh	Dh	Dh	Dh	Dh	Dh	Dh	H	H
9	H	Dh	Dh	Dh	Dh	H	H	H	H	H
5–8	H	H	H	H	H	H	H	H	H	H

S = Stand
H = Hit
Dh = Double (if not allowed, then hit)
Ds = Double (if not allowed, then stand)
SP = Split
SU = Surrender (if not allowed, then hit)

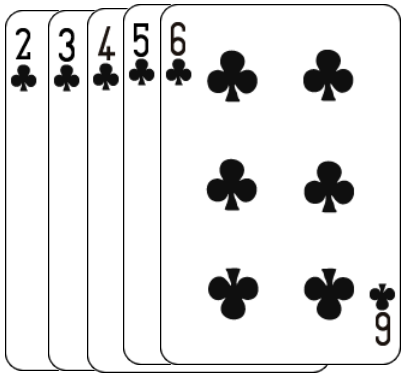
[Wikipedia]

Expected return vs inverse number of decks



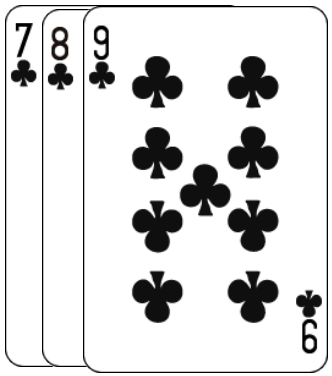
[Werthamer (2009)]

Recipe of counting – “high-low”

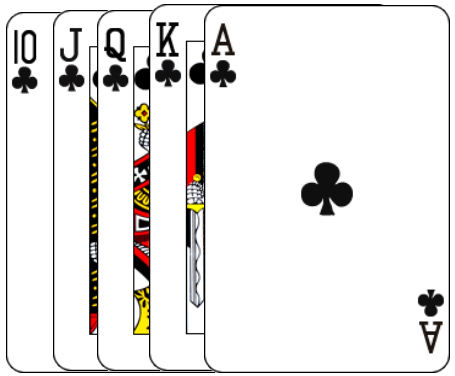


+1

“low cards”

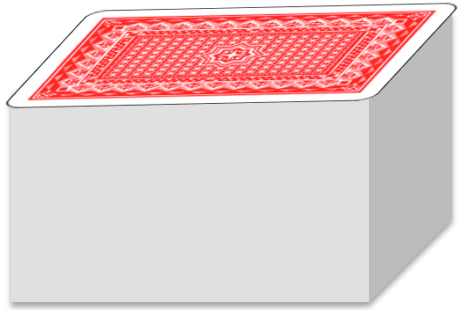


0



-1

“high cards”

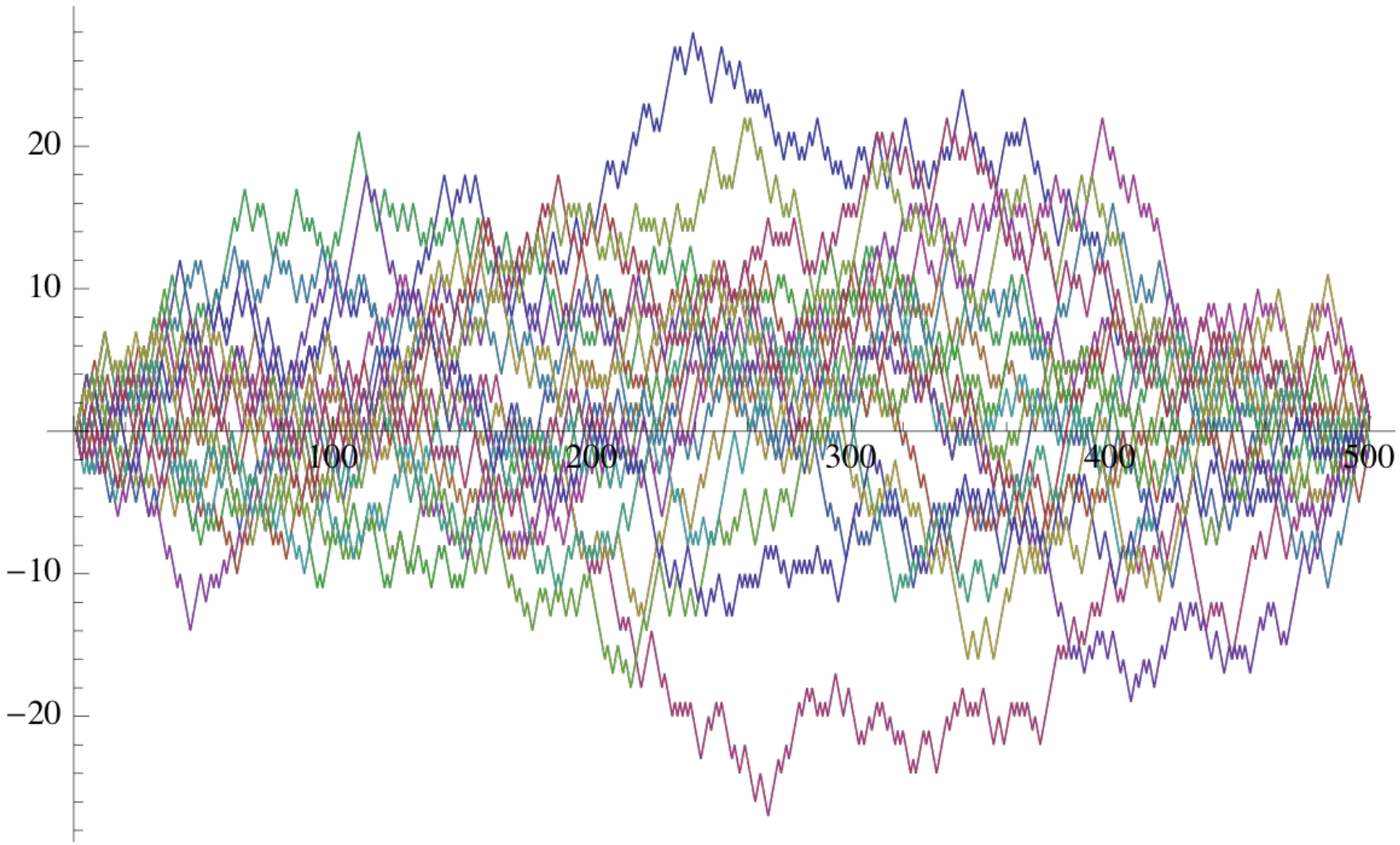


Large count

More high cards in the shoe

High return for players

Time evolution of count = random walk



Bet more money when the count is high!

Various counting systems

	A	2	3	4	5	6	7	8	9	10
Hi-Lo	-1	1	1	1	1	1	0	0	0	-1
Einstein	0	0	1	1	1	1	0	0	0	-1
Hi-Opt II	0	1	1	2	2	1	1	0	0	-2
Halves I	-1.5	1	1	1	1.5	1	0.5	0	-0.5	-1
Halves II	-1	0.5	1	1	1.5	1	0.5	0	-0.5	-1
K-O	-1	1	1	1	1	1	1	0	0	-1
Red 7	-1	1	1	1	1	1	0.5	0	0	-1
Zen	-0.5	0.5	1	1	1	1	0.5	0	0	-1

Optimal counting vector

	A	2	3	4	5	6	7	8	9	10
Optimal	-1.28	0.82	0.94	1.21	1.52	0.98	0.57	-0.06	-0.42	-1.07

[Werthamer (2009)]

- Leading order in $1/N$ expansion

N : number of cards in the deck $\sim O(100)$

True count knows expected return

True count

$$\gamma = \frac{\text{count}}{\#\text{ of remaining decks}}$$

Approximately,

$$R(\gamma) = R_0 + c\gamma$$

c indicates the performance of a counting system

$$R_0 \simeq -0.005 \qquad c \simeq 0.005$$

[Werthamer (2009)]

Distribution of the true count

f "penetration" – fraction of dealt cards

$$P(\gamma|f) = \frac{1}{\sqrt{2\pi}\sigma(f)} \exp\left[-\frac{\gamma^2}{2\sigma(f)^2}\right]$$

$$\sigma(f)^2 = C\frac{52f}{D(1-f)}$$

C : counting vector-dependent coefficient

Capital growth under Kelly betting

Assumptions:

- Players are allowed to bet no money in bad rounds
  - “minimum bet” exists in reality
- Payout variation (doubledown/natural blackjack) neglected

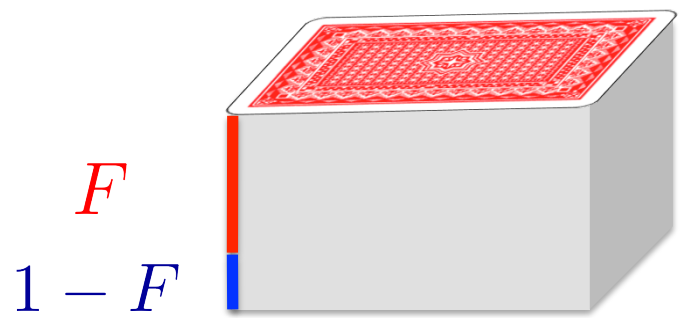
Capital growth under Kelly betting

$$\langle g \rangle_{\text{Kelly}} = \sum_{y,f,\gamma} P(y,f,\gamma)\theta(R_\gamma > 0) \ln 2 P(y|\gamma,f)$$

$$P(y,\gamma,f) = P(y|\gamma,f)P(\gamma|f)\boxed{P(f)}$$

$$P(f) = \frac{1}{F}\,\theta(0 \le f \le F)$$

F : the fraction at which cards are reshuffled



Capital growth under Kelly betting

$$\langle g \rangle_{\text{Kelly}} = \sum_{y,f,\gamma} P(y,f,\gamma)\theta(R_\gamma > 0) \ln 2P(y|\gamma,f)$$

$$P(y,\gamma,f) = P(y|\gamma,f)\boxed{P(\gamma|f)}P(f)$$

$$P(\gamma|f) = \frac{1}{\sqrt{2\pi}\sigma(f)} \exp\left[-\frac{\gamma^2}{2\sigma(f)^2}\right]$$

$$\sigma(f)^2 = C\frac{52f}{D(1-f)}$$

C : counting vector-dependent coefficient

Capital growth under Kelly betting

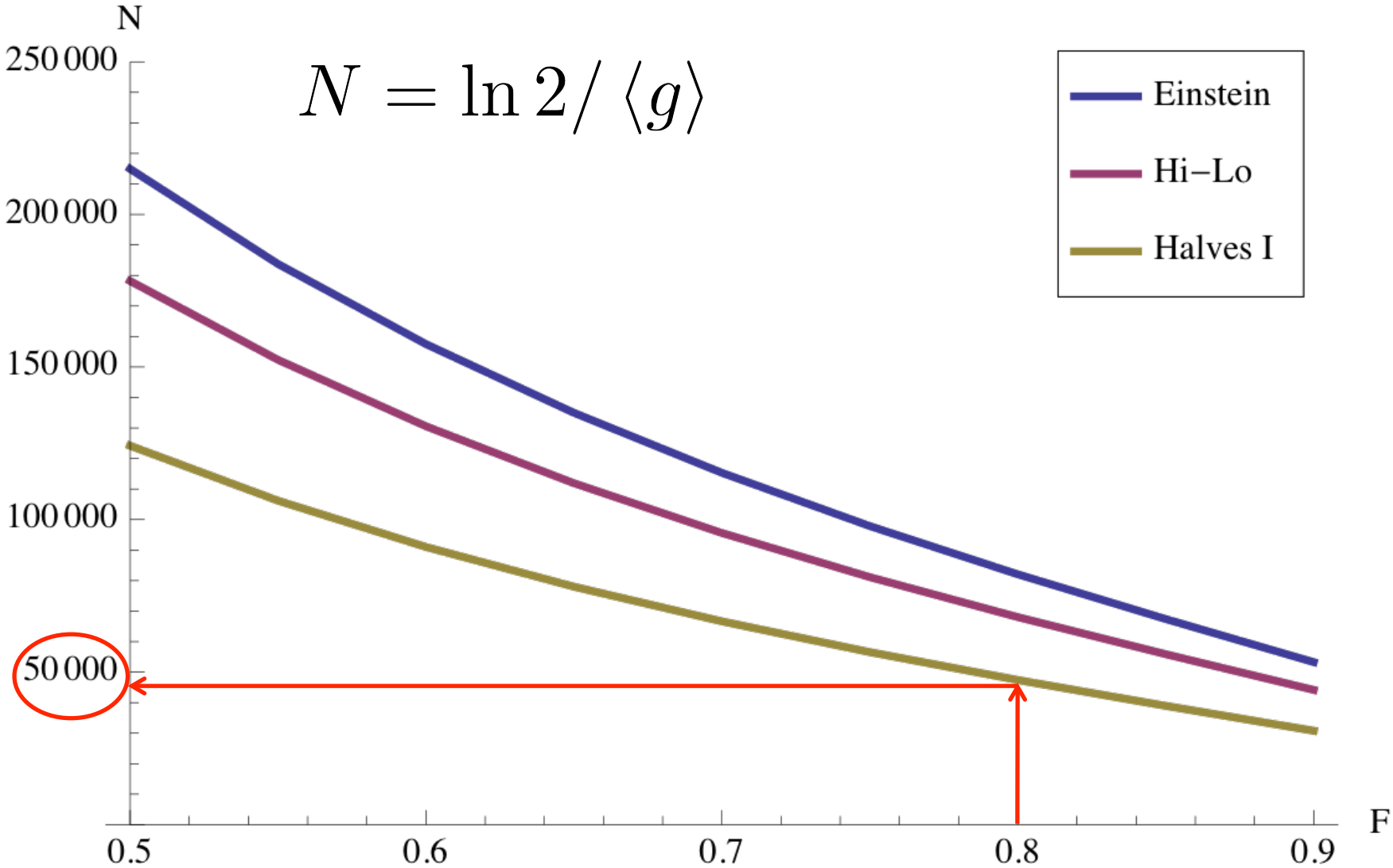
$$\langle g \rangle_{\text{Kelly}} = \sum_{y,f,\gamma} P(y,f,\gamma)\theta(R_\gamma > 0) \ln 2 \boxed{P(y|\gamma,f)}$$

$$P(y,\gamma,f) = \boxed{P(y|\gamma,f)} P(\gamma|f)P(f)$$

$$P(y=1|f,\gamma) = \frac{1+R(\gamma)}{2}$$

$$P(y=-1|f,\gamma) = \frac{1-R(\gamma)}{2}$$

Average number of rounds to double the capital



To double the capital,

$$50\text{k rounds} \simeq 800 \text{ hours}$$

60 rounds / h

$$\simeq 100 \text{ days}$$

8 hours / day

- We also have to consider:
  - Effects of minimum bet

Summary

- Derived Jarzynski-type equalities for gambling
  - Constrain the fluctuation of capital growth
  - Reproduce the Kelly bound
  - Financial value of side information is quantified by
    - Mutual information (memoryless)
    - Directed information (with memory effects)
- Analysis of blackjack
  - Capital growth rates under Kelly betting for various counting vectors