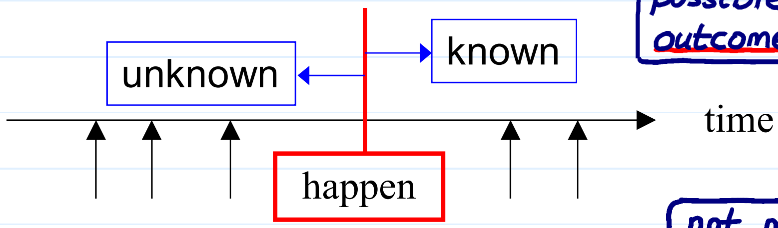
Introduction to Probability

- Uncertainty/Randomness (不確定性/隨機性) in our life
  - ➢ Many events are random in that their result is unknowable *before* the event happens.

at least 2 possible outcomes



not necessarily always the case

- ▪ Will it rain tomorrow?
- ▪ How many wins will a player/team achieve this season?
- ▪ What numbers will I roll on two dice?
- ▪ Q: Is your height/weight measure random?

c.f.

outcome: random
probability: fixed

➢ We often want to assess *how likely* it is the outcomes of interest occur. Probability is that measurement.

Random vs. Deterministic Patterns

random	混亂 隨機	noise (雜訊)	uncertain result
決定性論的 deterministic	規律 秩序	signal (信號)	predictable result 100%

- Consider the two cases:
  - ➢ Case I (← random pattern?)

1	2	3	4	5	6	7	8	9	10
🔴	🔴	🟢	🔴	🟢	🔴	🔴	🔴	🟢	?
R	R	G	R	G	R	R	R	G	

  - ➢ Case II (← deterministic pattern?)

1	2	3	4	5	6	7	8	9	10
🔴	🔴	🟢	🔴	🔴	🟢	🔴	🔴	🟢	?
R	R	G	R	R	G	R	R	G	

  - ➢ Note. #R : #G = 2 : 1
    number

- (Possible) modeling: the color in the nth trial.
  - Case I. $X_1, X_2, \ldots, X_n, \ldots$ are independent, for i=1, 2, …,

$$X_i = \begin{cases} R, & \text{with prob. } 2/3, \\ G, & \text{with prob. } 1/3. \end{cases}$$

random variable (future lecture)

cf.

  - Case II. For i=3, 4, …,

$$X_i = \begin{cases} R, & \text{if } (X_{i-2}, X_{i-1}) \in \{(R,G),(G,R)\}, \\ G, & \text{if } (X_{i-2}, X_{i-1}) = (R,R). \end{cases} \qquad (*)$$

- Prediction strategy:
  - Case I: always guess X_i=R (why? next slide)
  - Case II: decide X_i by X_{i-1}, X_{i-2} using (∗)

cf.

- **Q**: why always guess X_i=R for Case 1? To possibly reach 100% accuracy, we need to guess 100p% of R and 100(1-p)% of G. Is this a good strategy?
  - Let $X_i = \begin{cases} R, & \text{with prob. } p, \in [0,1] \\ G, & \text{with prob. } 1-p. \end{cases}$

guessing strategy $\rightarrow$ $Y_i = \begin{cases} R, & \text{with prob. } q, \in [0,1] \\ G, & \text{with prob. } 1-q. \end{cases}$

  - Then,

$$\begin{aligned} P(X_i = Y_i) &= P((X_i, Y_i) \in \{(G,G),(R,R)\}) \\ &= pq + (1-p)(1-q) \\ &= 1 - p + (2p-1)\, q \end{aligned}$$

∵ independent assumption

		Y_i: R	Y_i: G
X_i	R	✓	×
	G	×	✓

  - The $P(X_i$=$Y_i)$ is maximized at

$$q = \begin{cases} 1, & \text{if } p > 0.5, \Leftrightarrow 2p-1>0 \leftarrow \text{always guess } R \\ 0, & \text{if } p < 0.5. \Leftrightarrow 2p-1<0 \leftarrow \text{always guess } G \end{cases}$$

What if $p=1/2$?

and

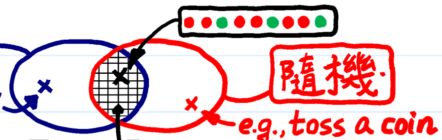
$$\max_q P(X_i = Y_i) = \begin{cases} p, & \text{if } p > 0.5, \\ 1-p, & \text{if } p < 0.5. \end{cases}$$

Giving up 100% accuracy and accepting errors in predition when G (or R) occurs increase overall predictive accuracy.

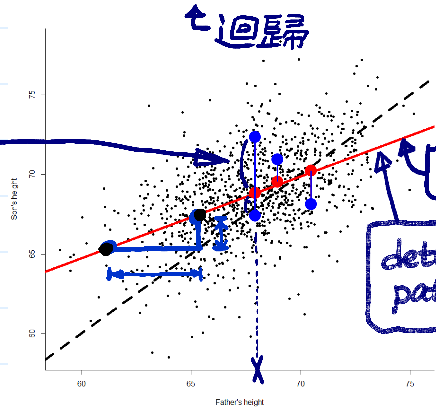
- **Q**: Is Case II really a deterministic pattern?
  - ➢ Under the model for Case I, ← LNp.1-3

$$P(RRGRRGRRG) = \left[\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)\right]^3 = 0.325\%$$

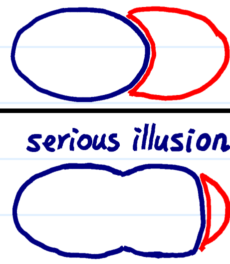
  - ➢ Random pattern $\xrightarrow{\text{誤判}}$ Deterministic pattern (remedy: replicate)
    - ← 隨機 ← 規律
  - ➢ Deterministic pattern $\xrightarrow{\text{誤判}}$ Random pattern (remedy: controlled experiment)
- System containing both random and deterministic patterns



  - ➢ Galton (1875): regression effect
    - ← 迴歸



illusion

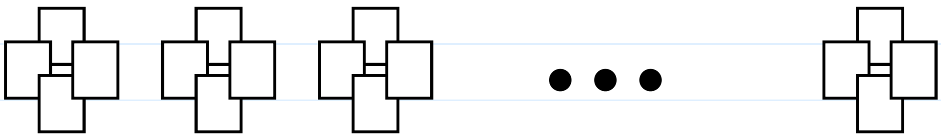


serious illusion

Should everyone have the same probability for an event?

more information
less random

- Example: 52 cards



	1	2	3		13
Player 1	32	24	10	• • •	deterministic
Player 2	X	X	X	• • •	random ← c.f.

- Conditional probability ← probability evolves → check the graph in LNp.1-1.
- Subjective (Bayesian) probability:

紅樓夢的作者是曹雪芹嗎?

信者恆信，不信者恆不信

It's the Chance (Probability, Proportion, Frequency), Stupid

- Bill Clinton, 1992, Campaign slogan

It's the Economy, Stupid.

- Examples
  - ➢ 該買某保險嗎?
  - ➢ 發生飛機失事事件後，該改成開車嗎?
  - ➢ 規畫謬誤：蚊子館、該創業嗎?
  - ➢ 敍述謬誤：偉人（成功）的故事 $\frac{\text{特質}}{\text{成功}}$ large $\overset{?}{\nRightarrow}$ $\frac{\text{成功}}{\text{特質}}$ large
  - ➢ 賭徒謬誤：擲筊多次未成，則擲出聖筊機會變大?
  - ➢ 馬路三寶? 汽車保險金額, 男>女
  - ➢ 車禍先問酒駕? 酒駕易肇事, yes, 但肇事者多酒駕?
  - ➢ 屏東人：你怎麼不黑? $\frac{\text{肇事}}{\text{酒駕}}$ large $\overset{?}{\nRightarrow}$ $\frac{\text{酒駕}}{\text{肇事}}$ large

Distinction between Discovery (發現) and Invention (發明)

- Examples
  - ➢ 哥倫布 "發現" 新大陸 ← 原本就有
  - ➢ 愛迪生 "發明" 電燈泡 ← 無中生有
  - ➢ **Q**: 相對論是發明還是發現?
- 機率論是人類 "發明" 來處理生活中的不確定性之理論
- 愛因斯坦: "上帝永遠不會擲骰子"

Indeterminableness could change with time.

other approaches ?
- 算命、占卜、…
- 風水、改運、…
- 祈禱、拜拜、…

something is due to chance
- not necessarily mean that it's indeterminate
- only mean that it is currently indeterminable

e.g.

❖ **Further Readings**: ← optional

✓ Kahneman (2011), *Thinking, Fast and Slow*. (中譯：快思慢想)

✓ Silver (2012), *The Signal and the Noise*. (中譯：精準預測)