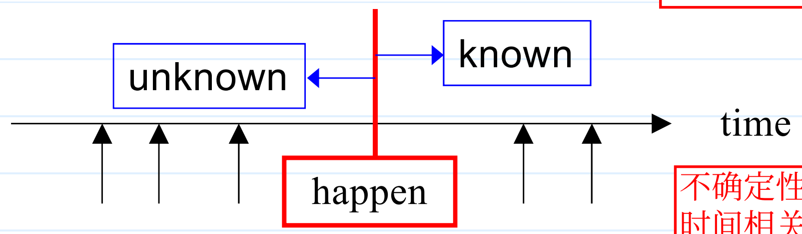
Introduction to Probability

- 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.

随机事件：参数特别多的确定性事件



不确定性也不一定与时间相关，可能是地点的不确定性

  - ▪ 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?

现实中很多事情可能看起来是确定性的，但是事实不是

c.f.

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

Random v.s. Deterministic Patterns

random	隨機 (混亂)	noise (雜訊)	uncertain result
deterministic	規律 (秩序)	signal (信號)	predictable result

決定性論的

之前学习的很多知识，如物理、化学的很多规律都是确定性的

- 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?)

最有可能的是哪个，建模需要考虑的因素是什么？

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 $(*)$

可以建模现实中的很多情景

这就是为什么佩服陈少平老师的原因了！

- **Q**: why always guess $X_i=R$ for Case 1?
  - Let

$$X_i = \begin{cases} R, & \text{with prob. } p, \\ G, & \text{with prob. } 1-p. \end{cases} \qquad p \in [0,1]$$

guessing strategy

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

  - 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	G
X_i	R	✓	×
	G	×	✓

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

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

What if $p=1/2$?

这种情况是最糟糕的情况，也是最好的情况

and

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

- **Q**: Is Case II really a deterministic pattern?

几率很小*时间很久--
一定发生！
黑天鹅现象！

  - Under the model for Case I,

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

  - Radom pattern $\xrightarrow{\text{誤判}}$ Deterministic pattern

损失准确率

股票分析师切
一半的方法！

  - Deterministic pattern $\xrightarrow{\text{誤判}}$ Random pattern

损失准确率

- System containing both random and deterministic patterns
  - Galton (1875): regression effect

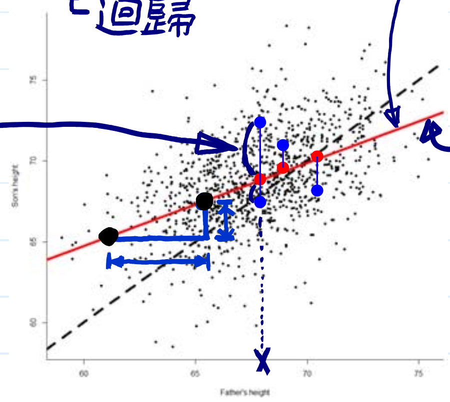
迴歸

slope < 1

随机事件与确定性事件
随着参数的变化（认
知）会不断改变

random
pattern
(隨机)

son's
height



father's height

deterministic
pattern (規律)

数据挖掘：抽取
规律

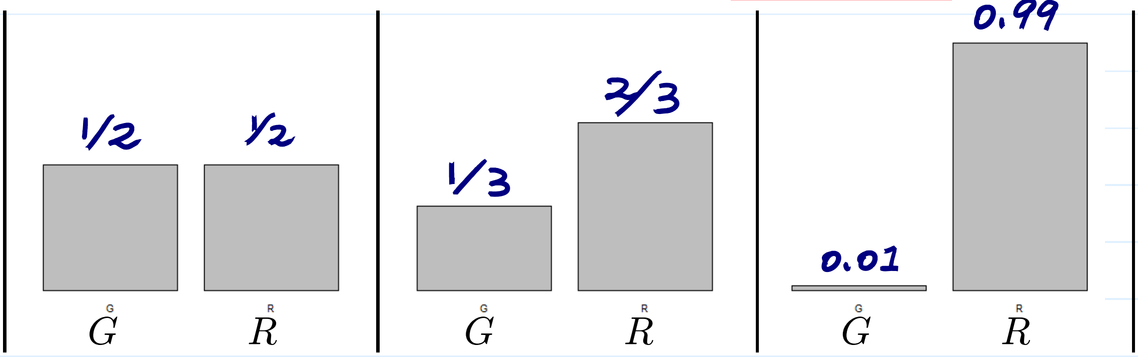
Random = totally unpredictable?

- Entropy: $-[p\log(p) + (1-p)\log(1-p)]$

亂度

图像是什么样
的？

知道规律后猜对的
可能性有多大
随机事件的可预测
性



$\max_q P(X_i = Y_i)$ ½ ⅔ 0.99

more random more unpredictable	more deterministic more predictable
larger entropy	smaller entropy

cf.

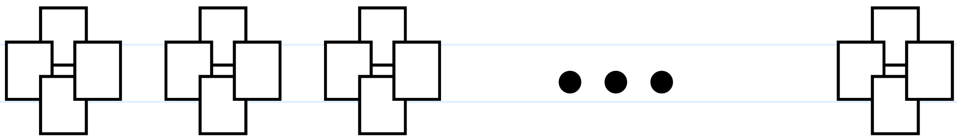
- Black swan effect

概率很小但是影响很大

easy to be misjudged as
deterministic system

Should everyone have the same probability for an event?

- Example: 52 cards



	1	2	3	• • •	13
Player 1	32	24	10	• • •	deterministic
Player 2	X	X	X	• • •	random

c.f.

信息战

- Conditional probability ← probability evolves

- Subjective (Bayesian) probability:

紅樓夢的作者是曹雪芹嗎?

信者恆信，不信者恆不信

咨询不对等！咨询越公平，竞争才公平，才有人原因参与游戏

概率值的判断不通才引起了市场的交易

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

- Bill Clinton, 1992, Campaign slogan

处理随机事件最好用的工具：概率论

It's the Economy, Stupid.

笨蛋，考虑问题的重点在经济上面

- 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
  - ➢ 賭徒謬誤：擲筊多次未成，則擲出聖筊機會變大? 否极泰来？
  - ➢ 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?

注意他人的描述有没有把随机事件描述成确定性事件

$\frac{\text{肇事}}{\text{酒駕}}$ large $\not\Rightarrow$ $\frac{\text{酒駕}}{\text{肇事}}$ 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? 理论是发明，现象是原始就有的
- 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"

参数太多，认为是随机，用概率论更方便地去处理（但是这不是真理）

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

❖ **Further Readings**:

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

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