ENGG 2760A / ESTR 2018: Probability for Engineers

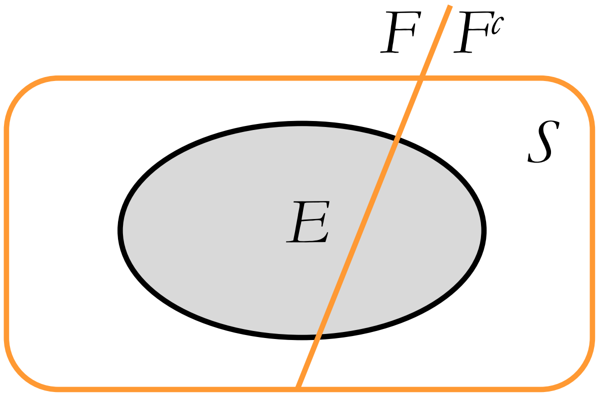
3. Conditional Probability and Independence

Prof. Hong Xu

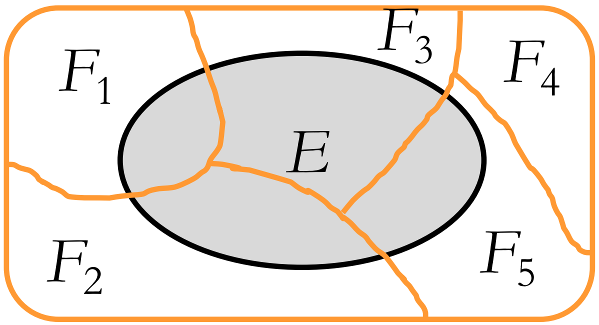
Credit to Prof. Andrej Bogdanov

Total probability theorem

$$\mathbf{P}(E) = \mathbf{P}(E \cap F) + \mathbf{P}(E \cap F^c)$$
$$= \mathbf{P}(E \mid F)\mathbf{P}(F) + \mathbf{P}(E \mid F^c)\mathbf{P}(F^c)$$



More generally, if $F_1, \ldots, F_n$ **partition** Ω **then**



$$\mathbf{P}(E) = \mathbf{P}(E \mid F_1)\mathbf{P}(F_1) + \ldots + \mathbf{P}(E \mid F_n)\mathbf{P}(F_n)$$

An urn has 10 white balls and 20 black balls. You draw two at random. What is the probability that their colors are different?

W_1 = 1st ball is white

D = 1st & 2nd balls have diff colors.

$$P(D) = P(D|W_1) \cdot P(W_1) + P(D|W_1^c) \cdot P(W_1^c)$$

$$= \frac{20}{29} \cdot \frac{10}{30} + \frac{10}{29} \cdot \frac{20}{30}$$

Geography quiz…

What is the capital of Romania?

A: Brasov 1%

B: Budapest 2%

C: Bucharest 5%

D: Bratislava 92%

Did you know or were you lucky?

Geography quiz

Probability model

There are two types of students:

Type K: Knows the answer

Type K^c: Picks a random answer

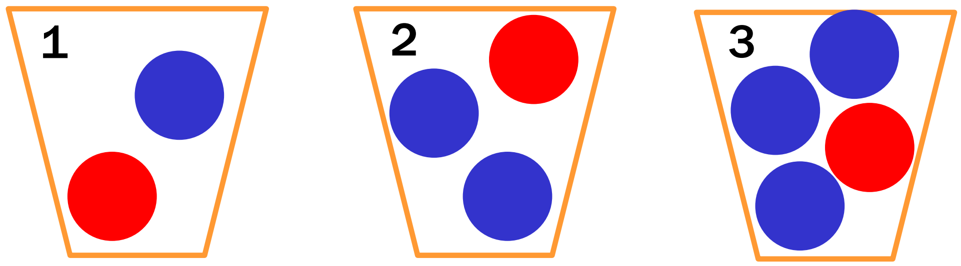
$P(K)$?

C = Give the correct answer.

$$P(C) = P(C|K)\cdot P(K) + P(C|K^c)\cdot P(K^c)$$

$$= 1 \cdot P(K) + 1/4 \quad (1-P(K))$$

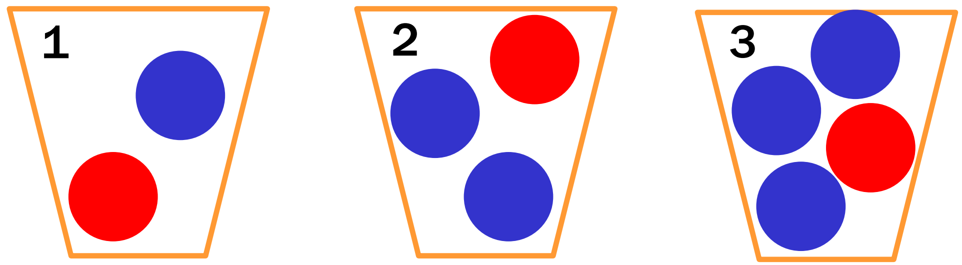
$$\Rightarrow P(K) = \frac{4}{3}\left(P(C) - \frac{1}{4}\right)$$



I choose a cup at random and then a random ball from that cup. The ball is red. You need to guess where the ball came from.

Which cup would you guess? CUP 1

Cause and effect



cause: C_1 C_2 C_3

effect: R

$P(C_1|R)$

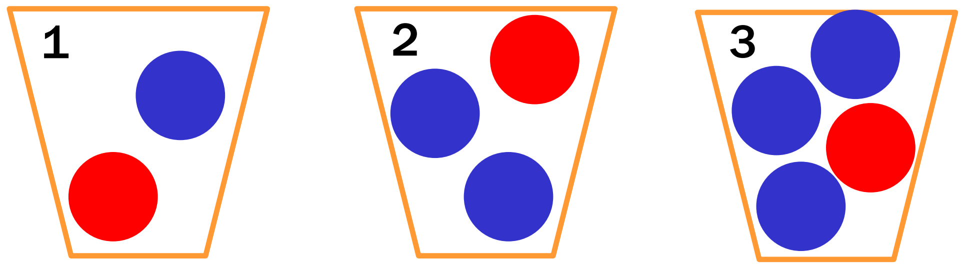
Bayes' rule

$$\mathbf{P}(C \mid E) = \frac{\mathbf{P}(E \mid C)\ \mathbf{P}(C)}{\mathbf{P}(E)} = \frac{\mathbf{P}(E \mid C)\ \mathbf{P}(C)}{\mathbf{P}(E \mid C)\ \mathbf{P}(C) + \mathbf{P}(E \mid C^c)\ \mathbf{P}(C^c)}$$

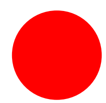
More generally, if $C_1, \ldots, C_n$ partition S then

$$\mathbf{P}(C_i \mid E) = \frac{\mathbf{P}(E \mid C_i)\ \mathbf{P}(C_i)}{\mathbf{P}(E \mid C_1)\ \mathbf{P}(C_1) + \ldots + \mathbf{P}(E \mid C_n)\ \mathbf{P}(C_n)}$$

Cause and effect



cause: C_1 C_2 C_3

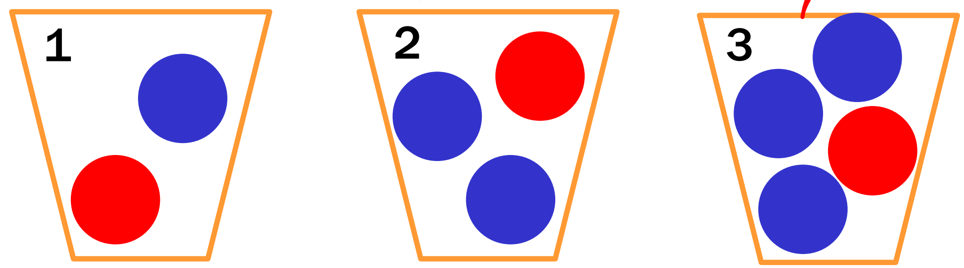


effect: R

$$\mathbf{P}(C_i \mid R) = \frac{\mathbf{P}(R \mid C_i)\,\mathbf{P}(C_i)}{\mathbf{P}(R \mid C_1)\,\mathbf{P}(C_1) + \mathbf{P}(R \mid C_2)\,\mathbf{P}(C_2) + \mathbf{P}(R \mid C_3)\,\mathbf{P}(C_3)}$$

Cause and effect

1/3 1/3 1/3



$$\Omega = \{R_1, B_1, R_2, B_{21}, B_{22}, R_3, B_{31}, B_{32}, B_{33}\}$$

1/6 1/6 1/9 1/12.

$$\mathbf{P}(C_i) = \quad 1/3 \qquad 1/3 \qquad 1/3$$

$$\mathbf{P}(R \mid C_i) = \quad \frac{1}{2} \qquad \frac{1}{3} \qquad \frac{1}{4}$$

$$\mathbf{P}(R) = \text{total prob.} \quad \frac{1}{3}\cdot\frac{1}{2} + \frac{1}{3}\cdot\frac{1}{3} + \frac{1}{3}\cdot\frac{1}{4} = \frac{13}{36}$$

$$\mathbf{P}(C_i \mid R) = \quad \frac{1/6}{13/36} = \frac{6}{13} \; ; \; \frac{1/9}{13/36} = \frac{4}{13} \; ; \; \frac{3}{13}$$

Two classes take place in Lady Shaw Building.

ENGG2430 has 100 students, 20% are girls.

NURS2400 has 10 students, 80% are girls.

A girl walks out. What are the chances that she is from the engineering class?

Causes : E (ENGG) , E^c (NURS).

Effect : G

$$P(E|G) = \frac{P(G|E) \cdot P(E)}{P(G|E) \cdot P(E) + P(G|E^c) \cdot P(E^c)}$$

$$= \frac{0.2 \cdot \frac{100}{110}}{0.2 \cdot \frac{100}{110} + 0.8 \cdot \frac{10}{110}} = \frac{20}{20+8} \approx 0.71$$

Summary of conditional probability

Conditional probabilities are used:

① When there are causes and effects

to estimate the probability of a cause when we observe an effect

② To calculate ordinary probabilities

Conditioning on the right event can simplify the description of the sample space

Independence of two events

Let E_1 **be "first coin comes up** $\mathtt{H}$**"**

E_2 **be "second coin comes up** $\mathtt{H}$**"**

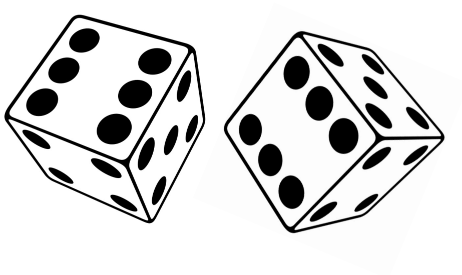
Then $\mathbf{P}(E_2 \mid E_1) = \mathbf{P}(E_2)$

$$\mathbf{P}(E_2 \cap E_1) = \mathbf{P}(E_2)\mathbf{P}(E_1)$$

Events A **and** B **are independent if**

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\ \mathbf{P}(B)$$

Examples of (in)dependence



Let E_1 be "first die is a 4"

S_6 be "sum of dice is a 6"

S_7 be "sum of dice is a 7"

NO

E_1 **and** S_6**?** $\quad \frac{1}{36} = P(E_1 \cap S_6) \stackrel{?}{=} P(E_1) \cdot P(S_6) = \frac{1}{6} \cdot \frac{5}{36}$

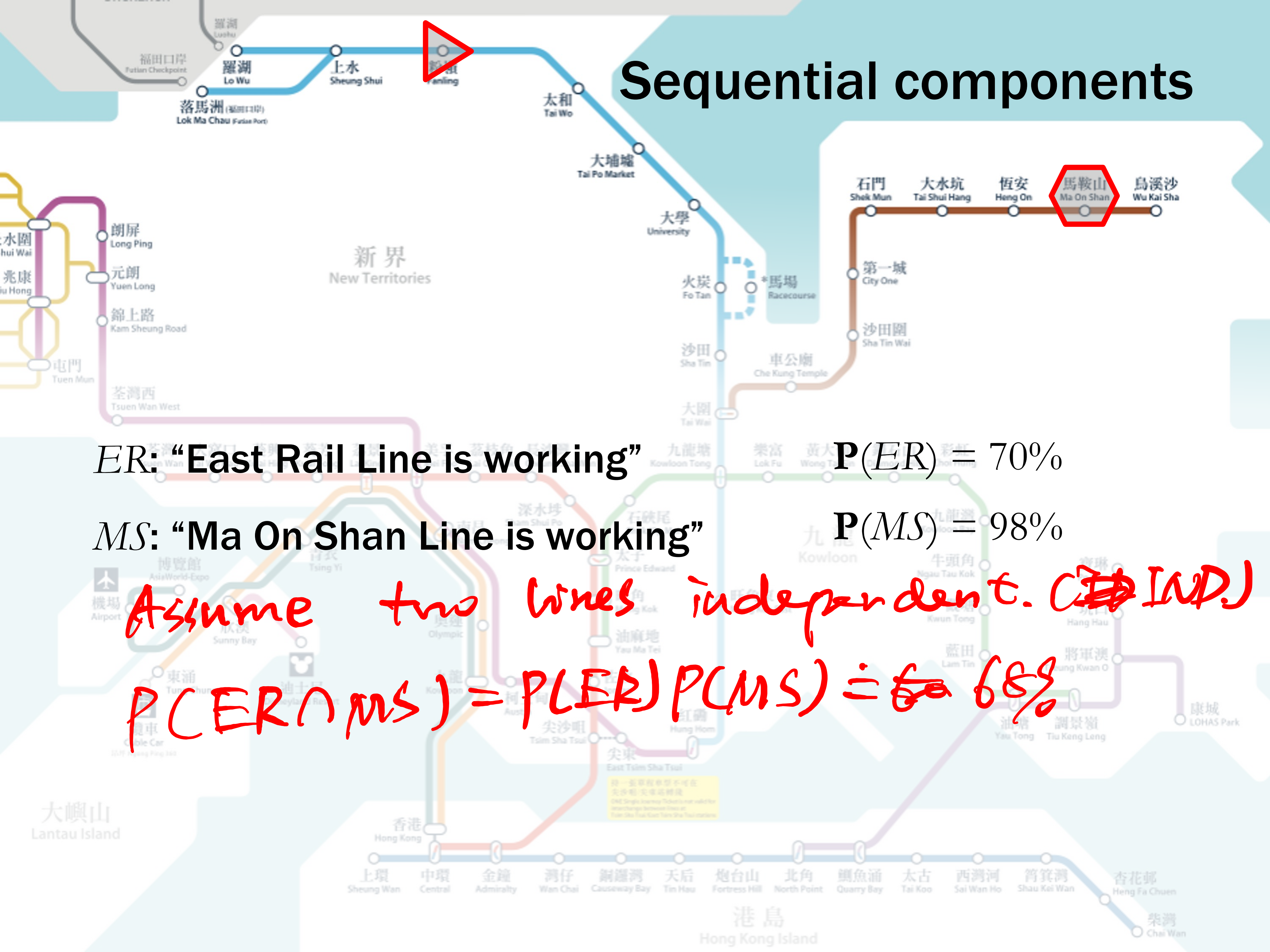
YES

E_1 **and** S_7**?** $\quad \frac{1}{36} = P(E_1 \cap S_7) \stackrel{?}{=} P(E_1) \cdot P(S_7) = \frac{1}{6} \cdot \frac{6}{36}$

NO

S_6 **and** S_7**?** $\quad 0 = P(S_6 \cap S_7) \stackrel{?}{=} P(S_6) P(S_7)$

Sequential components



ER: "East Rail Line is working" $\mathbf{P}(ER) = 70\%$

MS: "Ma On Shan Line is working" $\mathbf{P}(MS) = 98\%$

Assume two lines independent. (~~ER~~ IND.)

$P(ER \cap MS) = P(ER)\,P(MS) \doteq \cancel{60}\ 68\%$

Algebra of independent events

If A and B are independent, then A and B^c are also independent.

Proof: By axioms

$$P(A\cap B^c) = P(A) - P(A\cap B)$$
$$= P(A) - P(A)P(B)$$
$$= P(A)(1-P(B))$$
$$= P(A)\cdot P(B^c)$$

Parallel components

TW: **"Tsuen Wan Line is operational"** $P(TW) = 80\%$

TC: **"Tung Chung Line is operational"** $P(TC) = 85\%$

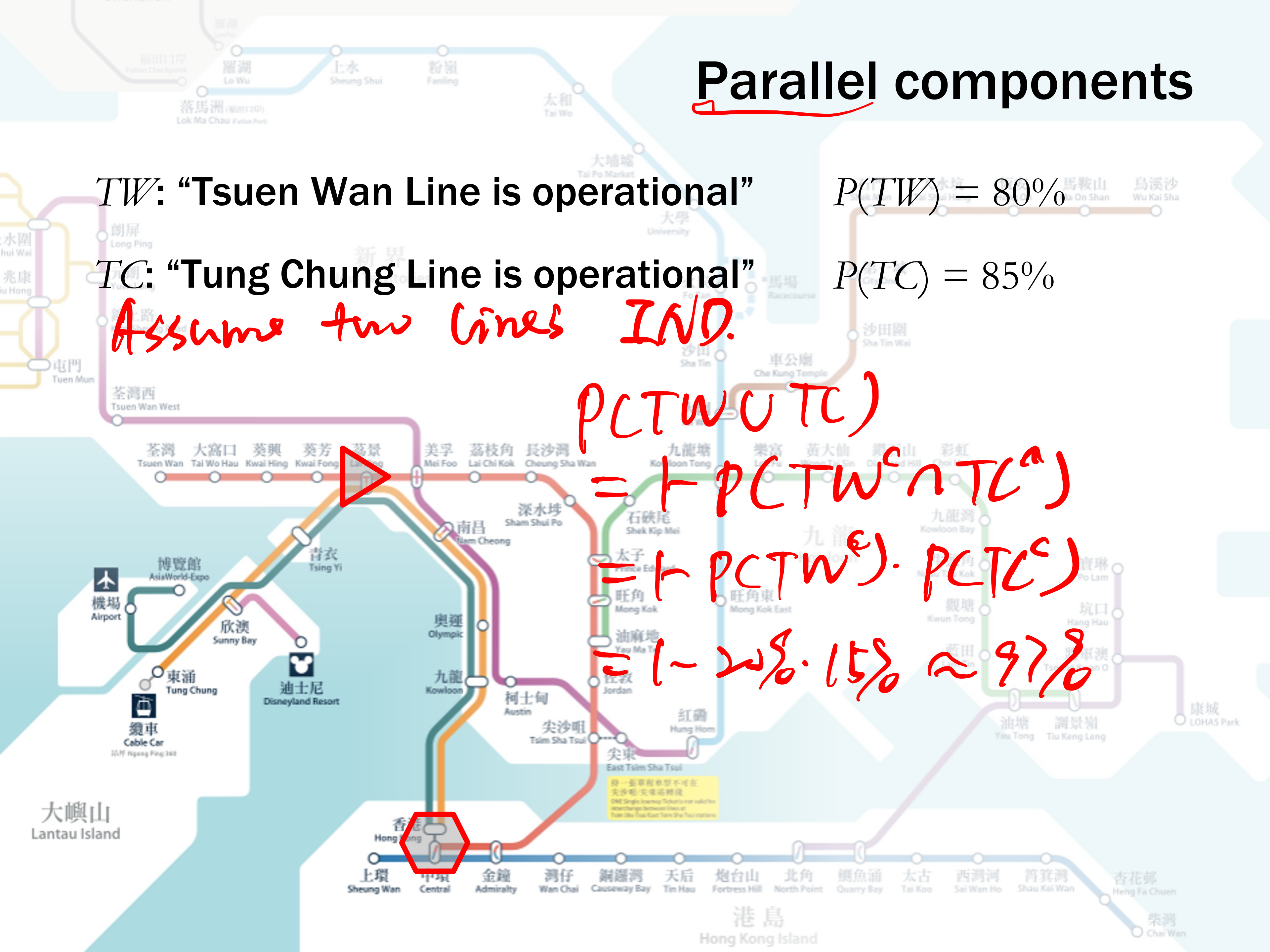
Assume two lines IND.

$P(TW \cup TC)$

$= 1 - P(TW^c \cap TC^c)$

$= 1 - P(TW^c) \cdot P(TC^c)$

$= 1 - 20\% \cdot 15\% \approx 97\%$



Independence of three events

Events A, B, and C are **independent** if

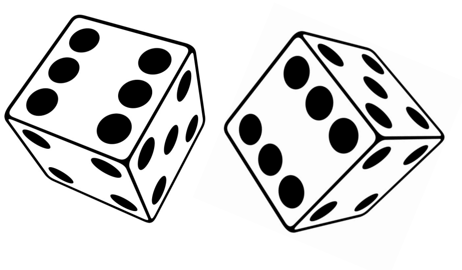
$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\,\mathbf{P}(B)$$

$$\mathbf{P}(B \cap C) = \mathbf{P}(B)\,\mathbf{P}(C)$$

$$\mathbf{P}(A \cap C) = \mathbf{P}(A)\,\mathbf{P}(C)$$

and $\mathbf{P}(A \cap B \cap C) = \mathbf{P}(A)\,\mathbf{P}(B)\,\mathbf{P}(C)$.

(In)dependence of three events



Let E_1 be "first die is a 4"

E_2 be "second die is a 3"

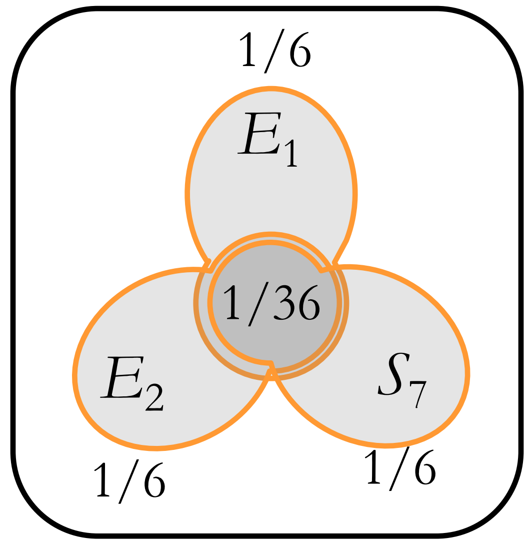
S_7 be "sum of dice is a 7"

E_1, E_2? $\frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6}$ ✓

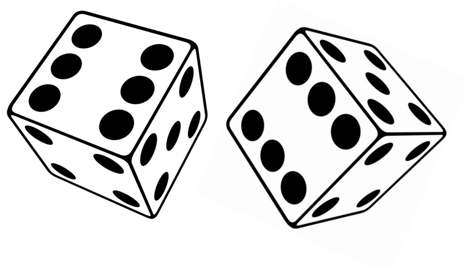
E_1, S_7? $\frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6}$ ✓

E_2, S_7? $\frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6}$ ✓

E_1, E_2, S_7? $\frac{1}{36} \neq \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6}$ NO



(In)dependence of three events



Let A be "first roll is 1, 2, or 3"

B be "first roll is 3, 4, or 5"

C be "sum of rolls is 9"

A, B? $\frac{1}{6} = P(A \cap B) \stackrel{?}{=} P(A)P(B) = \frac{1}{2} \cdot \frac{1}{2}$ NO

A, C? $\frac{1}{36} = P(A \cap C) \stackrel{?}{=} P(A)P(C) = \frac{1}{2} \cdot \frac{4}{36}$ NO

B, C? $\frac{3}{36} = P(B \cap C) \stackrel{?}{=} P(B)P(C) = \frac{1}{2} \cdot \frac{4}{36}$ NO

A, B, C? $\frac{1}{36} = P(A \cap B \cap C) \stackrel{?}{=} P(A)P(B)P(C) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{4}{36}$

YES!

Independence of many events

Events $A_1, A_2, \ldots$ are independent if **for every subset** of the events, the probability of the intersection is the product of their probabilities.

$n, \quad \binom{n}{2} \cdot \binom{n}{3} \cdot \binom{n}{4} \cdots \binom{n}{n}$

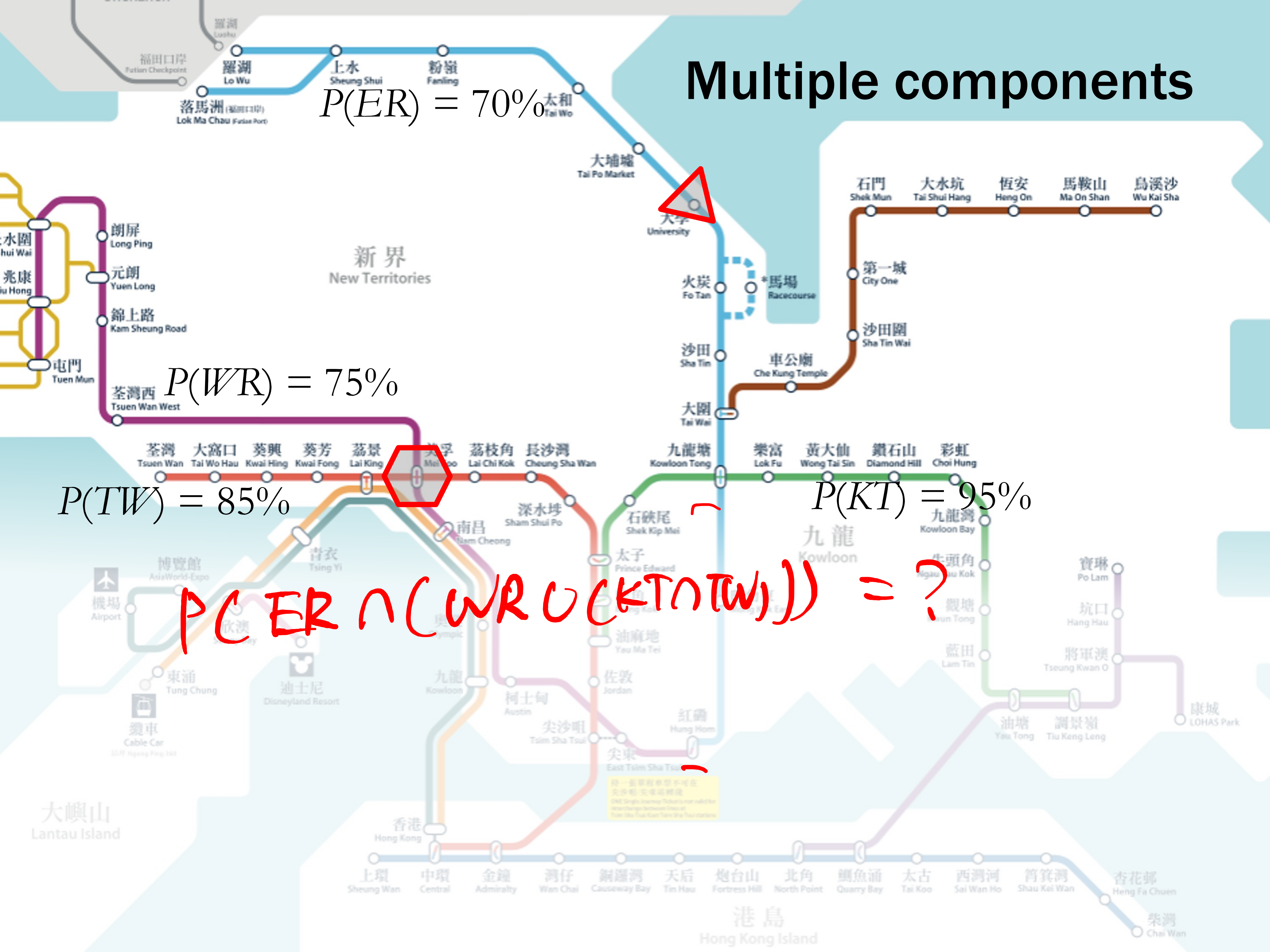
Algebra of independent events

Independence is preserved if we replace some event(s) by their complements, intersections, unions

Eg., A, B, C IND., $\Rightarrow P((A \cup B^c) \cap C)$

$= P(A \cup B^c) \cdot P(C)$

Multiple components



$P(ER) = 70\%$

$P(WR) = 75\%$

$P(TW) = 85\%$

$P(KT) = 95\%$

$P(ER \cap (WR \cup (KT \cap TW))) = ?$

Multiple components

$$P(ER) = 70\%$$
$$P(WR) = 75\%$$
$$P(KT) = 95\%$$
$$P(TW) = 85\%$$

$P(ER \cap (WR \cup (KT \cap TW)))$

$= P(ER) \cdot P(WR \cup (KT \cap TW))$

$= P(ER) \cdot (1-(1-P(WR)) \cdot (1-P(KT \cap TW)))$

$= P(ER) \cdot (1- - -) (1-P(KT) \cdot P(TW))$

$\approx 67\%$

Conditional independence

A **and** B **are independent conditioned on** F **if**

$$\mathbf{P}(A \cap B \mid F) = \mathbf{P}(A \mid F)\, \mathbf{P}(B \mid F)$$

Alternative definition:

$$\mathbf{P}(A \mid B \cap F) = \mathbf{P}(A \mid F)$$

today	tomorrow
☀️	80% ☀️, 20% 🌧️
🌧️	40% ☀️, 60% 🌧️

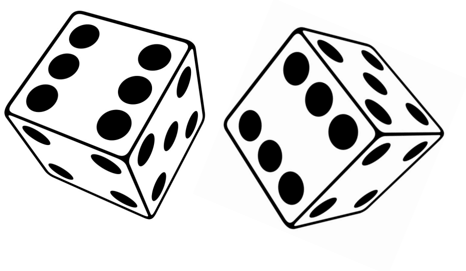
It is ☀️ on Monday. Will it 🌧️ on Wednesday?

$$P(W_R \mid M_S) = P(W_R \mid T_S \cancel{\cap M_S}) \cdot P(T_S \mid M_S) + P(W_R \mid T_R \cancel{\cap M_S}) \cdot P(T_R \mid M_S)$$

$$= 0.2 \cdot 0.8 + 0.6 \cdot 0.2$$

$$= 28\%$$

Conditioning does not preserve independence



Let E_1 be "first die is a 4"

E_2 be "second die is a 3"

S_7 be "sum of dice is a 7"

E_1 E_2 IND.

But, $P(E_1 \cap E_2 \mid S_7) \stackrel{?}{=} P(E_1 \mid S_7) \cdot P(E_2 \mid S_7)$

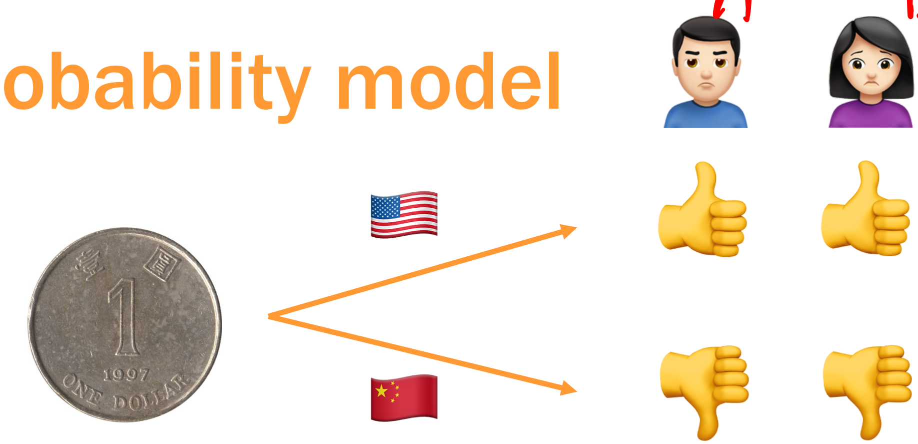
$1/6 \neq 1/6 \cdot 1/6$

The New York Times

‘Crazy Rich Asians’ Has Soared, but It May Not Fly in China

Conditioning may destroy dependence

Probability model



99%

1%

C: in the U.S.

Given C: A, B IND. $P(A\cap B)\neq P(A)P(B)$

C^c: A, B IND.

$P(A\cap B) = P(A\cap B|C)\cdot P(C) + P(A\cap B|C^c)\cdot P(C^c)$

$= \frac{1}{2}\cdot 0.99^2 + \frac{1}{2}\cdot 0.01^2$

≈ 0.49

$P(A) = P(A|C)P(C)\cdot P(A|C^c)\cdot P(C^c)$

$= 0.99\cdot\frac{1}{2} + 0.01\cdot\frac{1}{2} = 0.5$

$P(B) = 0.5$