Part IIa Inferential Statistics ①

- probability
- Distributions

((Part II b will be the actual Inference Process...))

Chapt 5 Probability

5.1 Introduction to Probability

Def: **Emperical probability** (vs. mathematical) is the proportion of times that an actual experiment produces a certain Result.

Ex Coin toss : After tossing a silver dollar 100 times we see 52 Heads and 48 tails.
so we can say that this coin has an emperical probability of 52% Heads

The Law of Large Numbers states that over many times the "true" probability will be revealed

Ex After 1000 tosses we see P(Head) = 0.502
much closer to the mathematical prob. of
50 - 50 or $P(H) = \frac{1}{2}$

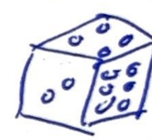
Mathematical Probability is the ratio of the number of ways to obtain a certain outcome with the total possible outcomes

Outcomes of an experiment are all possible results that could occur.

$$P(A) = \frac{\text{number of outcomes with } A}{\text{number of total outcomes possible}}$$

EX P(tossing a 5 when a single die is rolled)

$$= \frac{1 \text{ side with a } 5}{6 \text{ sides on the die}} = \frac{1}{6}$$



The **Sample Space** of an experiment is a statement of all possible outcomes

EX
- Die : $S = \{1, 2, 3, 4, 5, 6\}$
- Coin : $S = \{H, T\}$
- Deck of Cards :

S = { Ace of Diamonds, 2 of diamonds, ..., Jack, Queen, King
Ace of Spades 2 of spades, ... J, Q & K of spades
Ace of Heart, 2 of hearts ..., J, Q, K of Hearts
Ace of Clubs, 2 of Clubs, ... 9, 10, J, Q, King of Clubs }

let $N(S)$ = the number of outcomes in a sample space S

Ex Die : $N(S) = 6$

Coin : $N(S) = 2$

Deck : $N(S) = 4 \times 13 = 52$

The probability of an outcome is a number between $\boxed{0 \text{ and } 1.}$

Ex $P(\text{rolling a } 7) = 0.0$

$P(\text{rolling a } 1, 2, 3, 4, 5, \text{ or } 6) = 1.0$

- Statistical Outcomes account for the fact that not always do we get the mathematical outcome.

Ex Toss a coin 10 times

we might get 6H and 4T vs. the mathematical outcome of 5H and 5T

"Statistical Variability"

Ex Lets now toss 2 coins at a time
{or one after the other}

- Sample Space : {HH, HT, TH, TT} , N(S) = 4

Q_1 : $P(HH) = \frac{1}{4}$

Q_2 : $P(H \text{ and } T) = \frac{2}{4}$

Q_3 : $P(\text{at least one tail}) = \frac{3}{4}$

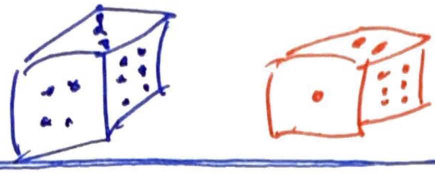
Q_4 : $P(\text{no tails}) = \frac{1}{4}$ {only HH}

Note : $P(\text{at least } 1T) + P(\text{no "T"}) = 1$

$\frac{3}{4} + \frac{1}{4} = \frac{4}{4}$

$$\boxed{\sum_{i=1}^{N} P(A_i) = 1.0}$$

EX Toss two dice, one red & one blue



we use color to assure we are properly counting the outcomes all of which are Equally Likely to occur.

- Sample Space = {(1, 1), (1, 2), (1, 3), ...
  (2, 1), (2, 2), (2, 3), ...
  ⋮
  (6, 1), (6, 2), (6, 3) ... (6, 6)}

- Table:

	1	2	3	4	5	6
1	1 1	1 2	1 3	1 4	1 5	1 6
2	2 1	2 2	2 3	2 4	2 5	2 6
3	3 1	3 2	3 3	3 4	3 5	3 6
4	4 1	4 2	4 3	4 4	4 5	4 6
5	5 1	5 2	5 3	5 4	5 5	5 6
6	6 1	6 2	6 3	6 4	5 5	6 6

all outcomes are equally likely

$N = 36$

blue is ↘ red is ↘

(a) $P(1,4) = \boxed{1/36}$, $P(\underbrace{1 \,\&\, 4}_{\text{We Don't care about color}}) = \dfrac{(1,4) \text{ or } (1,4)}{36} = \boxed{\dfrac{2}{36}}$ w/o regard to color

(b) $P(\text{doubles}) = \boxed{6/36}$

(c) $P(\text{sum} = 5) = \boxed{4/36}$

(d) $P(\text{sum} \geq 10) = \boxed{6/36}$, $P(\text{sum} < 10) = \boxed{30/36}$

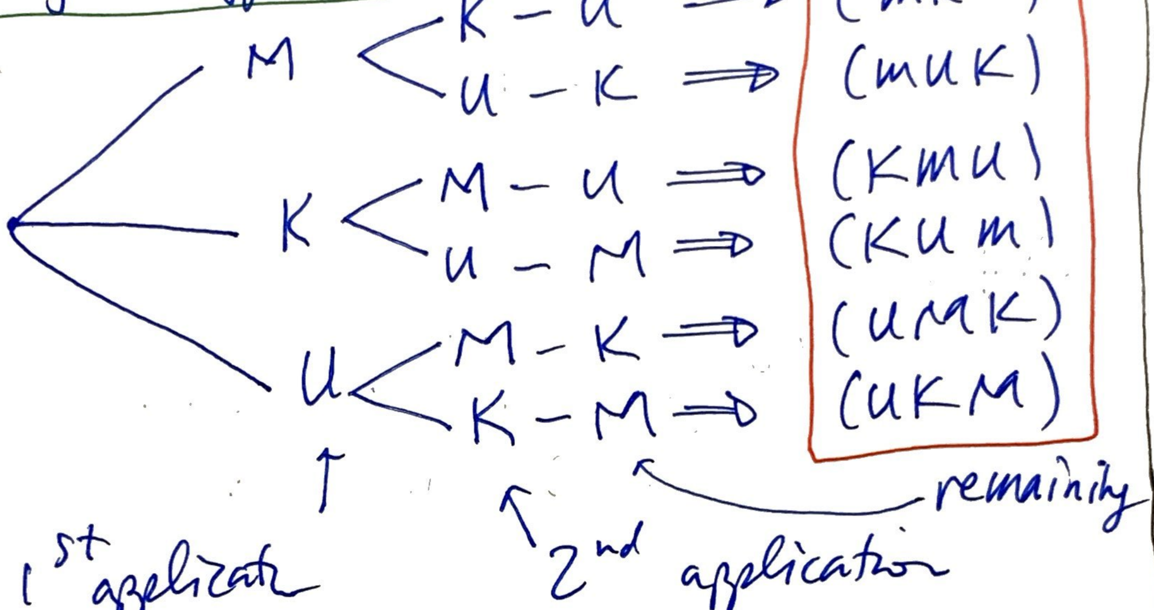
compliment : $\boxed{P(\geq 10) + P(< 10) = 1}$

⊛ Tree diagrams can be used to count all possible outcomes

Ex A concessions stand offers 3 condiment sauces:

Mayo M, Ketchup K, Mustard U

Use a tree diagram to write all possible orderings of application



S =
(MKU)
(MUK)
(KMU)
(KUM)
(UMK)
(UKM)

- $N(S) = 6$ different orderings in which we can apply the condiments.

⊛ Probability and Sampling

Drawing a sample from a population is a probability experiment.

Ex 10,000 families live in Saugus. Their dwelling can be displayed in a table

Results of a Census

Own a house	4753	own
Own a Condo	1478	
Rent a House	919	Rent
Rent a Condo	2857	
	10,000	

Emperical Probability (vs. mathematical)

(a) P(a randomly selected family own a condo)

$$= \frac{\text{Specific}}{\text{Generic}}$$

$$= \frac{\text{Own a condo}}{\text{All families}} = \frac{1478}{10{,}000} = \boxed{15\%}$$

(b) P(rent either a house or condo)

$$= \frac{919 + 2857}{10{,}000}$$

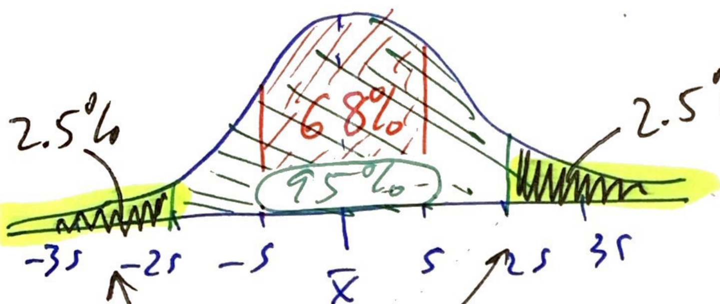
$$= \boxed{38\%}$$

⊛ Unusual Events

- Statement by Statisticians:

> If the probability of an event occuring is less than 5%, that event is **unusual**

- Recall in Distributions (unimodal & Symmetric)



"wings of the distribution"

unusual for a data point to be out here.

EX For 150 math majors at Berkeley out of 35,000 students, is it unusual to randomly bump into a math major at the boba house?

P(being next to a math major in the Student Union Bobatime line)

$$= \frac{\text{Specific}}{\text{Generic}}$$

$$= \frac{150}{35{,}000} = \underset{\text{proportion}}{\boxed{0.004}} \text{ or } \underset{\%}{\boxed{0.4\%}}$$

very unusual