Why Probability?

We want to combine uncertain statements/observations to end up with logical inferences.

- $\boxed{\text{frequentist}}$
- Bayesian approach (subjective)

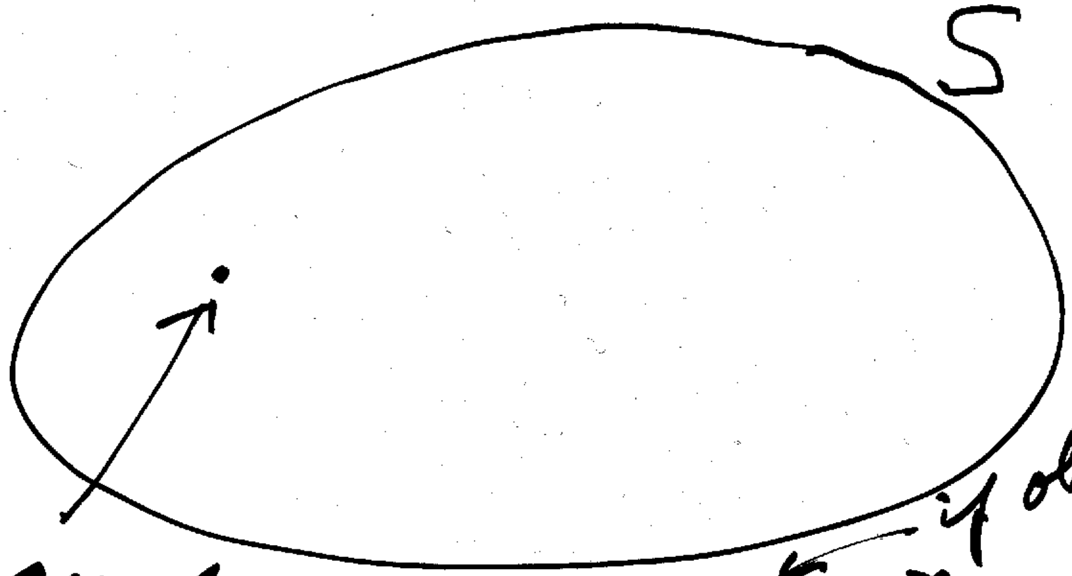
We talk about probabilities of events & denote the probability of an event A by P(A).

<u>Impossible Event</u> – $\emptyset \rightarrow P(\emptyset) = 0$

<u>Sure event</u> $- S \rightarrow P(S) = 1$
(Ω)

Events either <u>occur</u> or they <u>do not</u>.

Suppose we are interested in the analysis of some situation. We decide what we want to observe. The set of (ideal) possible observations is called the <u>sample space</u>. It is usually denoted by S.
Note that S is a <u>set</u> of outcomes or <u>sample points</u>.



if observing #'s   "   vectors

sample point $s, \omega, x, \underset{\sim}{x}, g$

eg depth of a lake at a certain point over a certain time

You might be interested in

$$X = \max_{a \le t \le b} g(t)$$

Notice $X : S \to \mathbb{R}$. This is a <u>random variable</u> (rv)

Another rv is

$$Y = \min_{a \le t \le b} g(t)$$

The vector

$$\begin{pmatrix} X \\ Y \end{pmatrix} : S \longrightarrow \mathbb{R}^2$$

is a <u>random vector</u> (r vec).

<u>eg</u> Toss a coin & observe H or T.

$$S = \{H, T\}$$

sample points

Set $X(H) = 1$ & $X(T) = 0$.
This is the rv "# of H's on one toss".

eg Toss a coin 3 times & observe the sequence of faces.

$$S = \{HHH, TTT, THT, \ldots\}$$

8 sample points

Let X = # of H's tossed.
Then X is a rv. The possible values of X are 0, 1, 2, 3

Now measure X. This yields a number x.

eg Take a course & your grade is observed.

$$S = \{x \mid 0 \le x \le 100\} = [0, 100]$$
$$= \{\omega \mid 0 \le \omega \le 100\}$$

Let X = your grade. Notice
$X(\omega) = \omega \quad X(x) = x$

eg Roll a die & observe the # of dots.

$$S = \{1, 2, 3, 4, 5, 6\}$$

a sample point

Consider the subset

$$A = \{2, 4, 6\}$$

= the event of rolling an even # of dots

= {rolling an even # of dots}

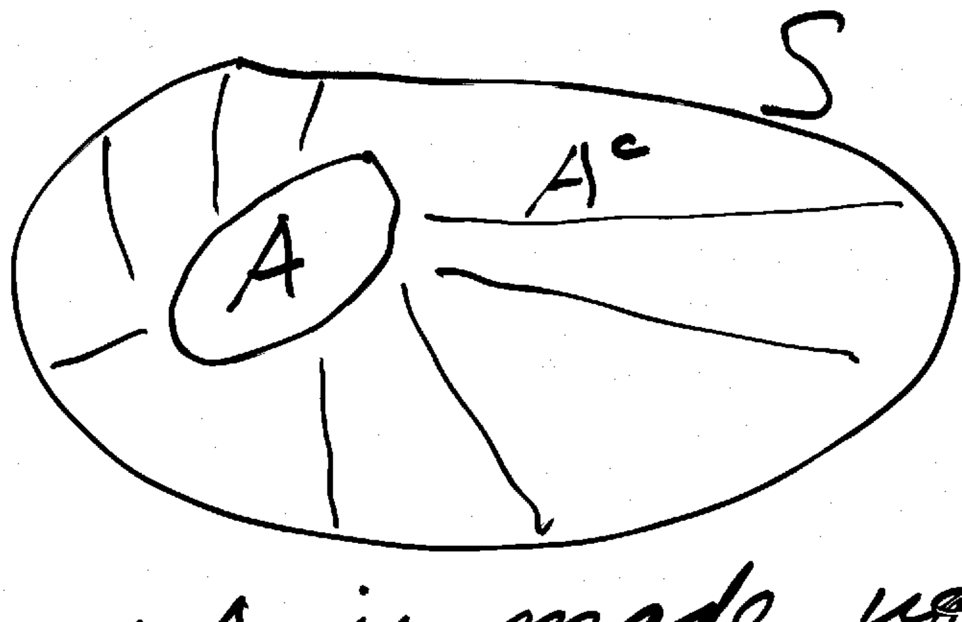
Def'n An event is a subset of S.

Note ① S is an event (the sure event)

② An event occurs if and only if the observation is an element of the event.

③ The empty set never occurs & is the <u>impossible event</u> $\emptyset$.

We need to know more about sets / events.



The event A is made up of "possible outcomes", ie sample points.

A^c is made up of points outside A.

Let A & B be events. If they are useful in predicting each other, they are called <u>dependent</u> events. Otherwise, they are <u>independent</u>.

<u>Notation</u>

and/intersection

AB (or $A \cap B$)

= the event that both A & B occur.

$A_1 A_2 \cdots$ = the event that all the A_i's occur.

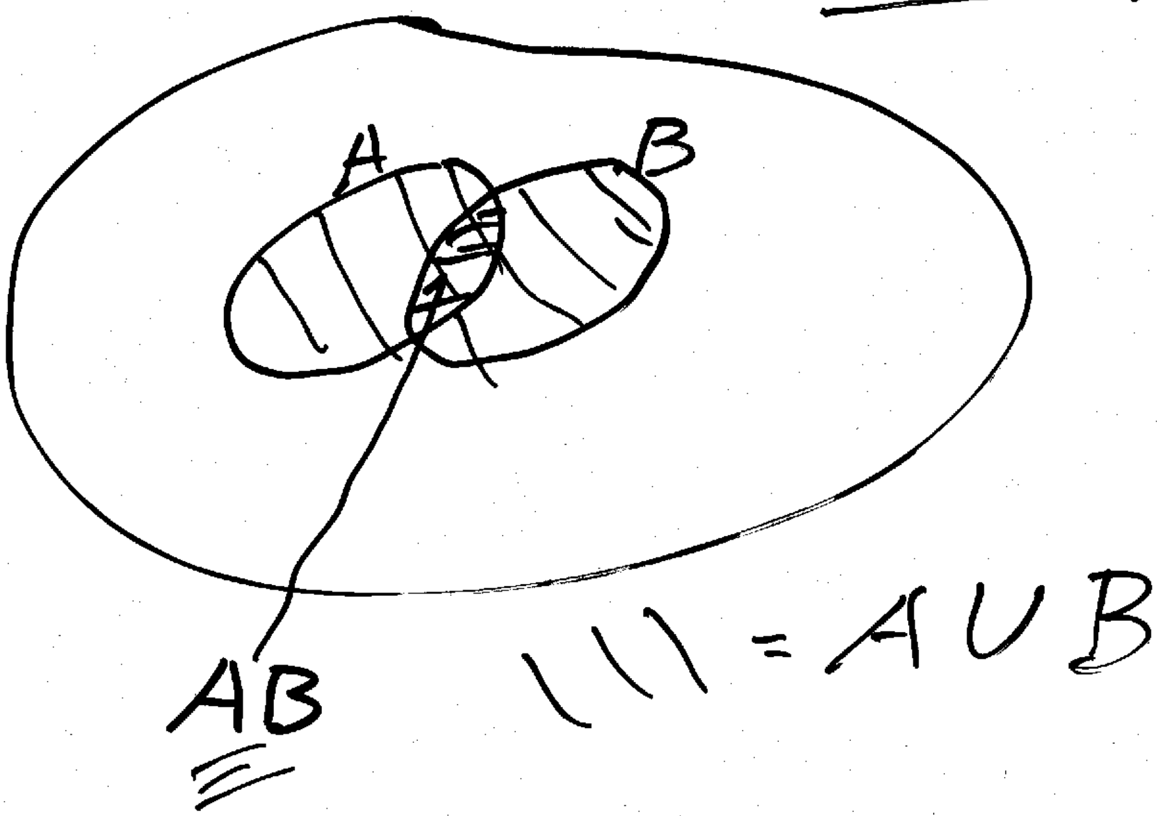
$A \cup B$ = the event that at least one of the two occurs.

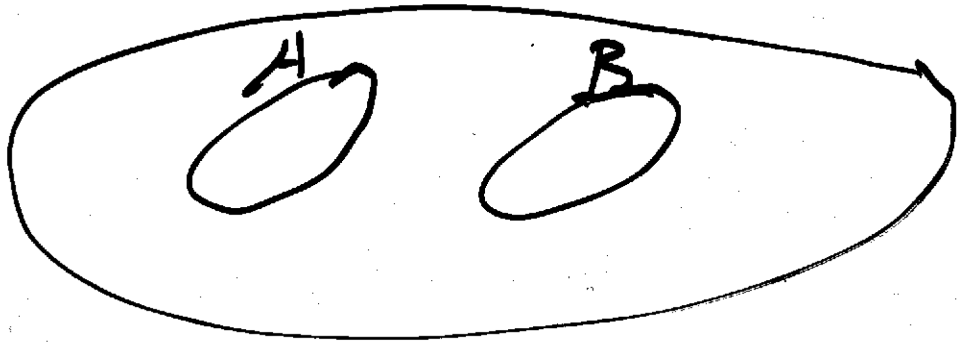
or/union

$A_1 \cup A_2 \cup \cdots$ = the event that at least one of the A_i's occurs.

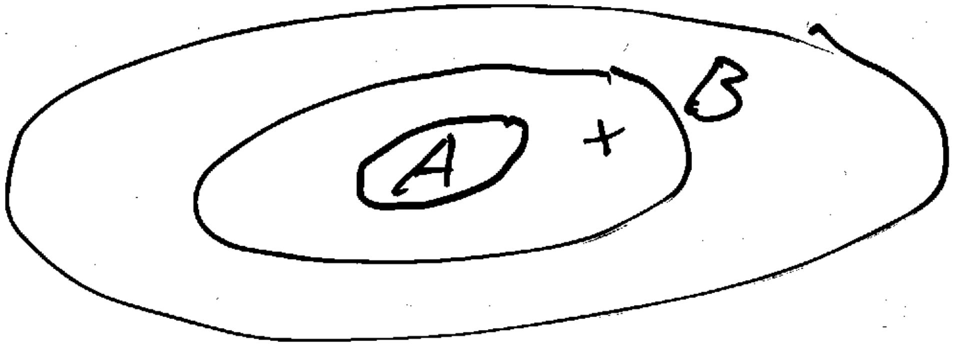
<u>Notation</u> $\bigcup_{i=1}^{\infty} A_i$, $\bigcap_{i=1}^{\infty} A_i$

<u>Note</u> <u>Venn diagram</u>





no overlap = <u>disjoint</u>



A occurs implies B occurs

$$A \Longrightarrow B$$

$$(A \subset B)$$

<u>Note</u> Suppose $A \Longrightarrow B$ and $B \Longrightarrow A$. Then

$$A = B$$

<u>DeMorgan's Laws</u>

(1) $\left[\bigcap_i A_i\right]^c = \bigcup_i A_i^c$ ← try to show

(2) $\left(\bigcup_i A_i\right)^c = \bigcap_i A_i^c$

Def'n A & B are independent if $P(AB) = P(A)\,P(B)$

Def'n $A_1, A_2, A_3, \cdots$ are ind if for any $A_{i_1}, A_{i_2}, \cdots, A_{i_m}$

$$P(A_{i_1} A_{i_2} \cdots A_{i_m}) = P(A_{i_1})\,P(A_{i_2}) \cdots P(A_{i_m})$$

Kolmogorov Axioms for P

1. $P(S) = 1$
2. $P(A) \geq 0$
3. $P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$

disjoint

Proposition $P(\emptyset) = 0$

empty event

Proof Let A be an event. Then

$$A = A \cup \emptyset \cup \emptyset \cup \cdots$$

disjoint

$\Rightarrow$ implies $P(A) = P(A \cup \emptyset \cup \cdots)$

$= P(A) + P(\emptyset) + \cdots$, by Law 3 of P

$\Rightarrow$ $P(\emptyset) = 0$

"The End" "~~—~~" "$\square$" "qed"

Proposition If $A_1, \ldots, A_n$ are disjoint then

$$P(A_1 \cup \cdots \cup A_n) = P(A_1) + \cdots + P(A_n)$$

Proof $P(A_1 \cup \cdots \cup A_n) = P(A_1 \cup \cdots \cup A_n \cup \emptyset \cup \cdots)$

$$= P(A_1) + \cdots + P(A_n) + \underbrace{P(\emptyset)}_{0} + \underbrace{P(\emptyset)}_{0} + \cdots$$

$$= P(A_1) + \cdots + P(A_n)$$

qed

Application

Roll a fair die & observe the # of dots. Then

$$P(\{3\}) = \frac{1}{6}$$

Proof $S = \{1\} \cup \{2\} \cup \{3\} \cup \cdots \cup \{6\}$

disjoint

$$\Rightarrow \underbrace{P(S)}_{=1} = P(\{1\}) + \cdots + P(\{6\})$$

equal

$\therefore$ die is fair

$$= 6\,P(\{3\})$$

$$\Rightarrow P(\{3\}) = 1/6$$

qed

Important The previous argument shows that a finite sample space with equiprobable sample points yields

$$P(A) = \frac{|A|}{|S|}$$ ← # of outcomes in A

A bit of counting

$3 \times 2 \times 1 = 3!$ (3 factorial)

$n \times (n-1) \cdots \times 2 \times 1 = n!$ (n ")

$0! = 1$

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

"n choose k"

$$(x_1 + x_2)^n = \sum_{k=0}^{n} \binom{n}{k} x_1^k x_2^{n-k}$$

$\boxed{\text{binomial theorem}}$

$\binom{n}{k}$ = # of subsets of size k of set of n elements

So $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n}$

= total # of subset of a set of size n

$= (1+1)^n = 2^n$

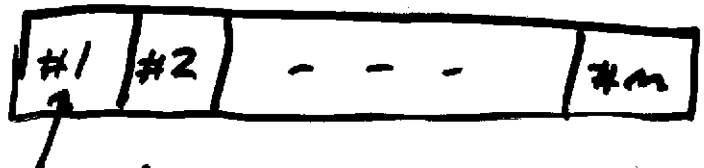
Set

$$\binom{n}{n_1, n_2, \ldots, n_k} = \frac{n!}{n_1! \cdots n_k!},$$

where $n_1 + n_2 + \cdots + n_k = n$ is the multinomial coefficient. ($k = 2$ yields the binomial coefficient).

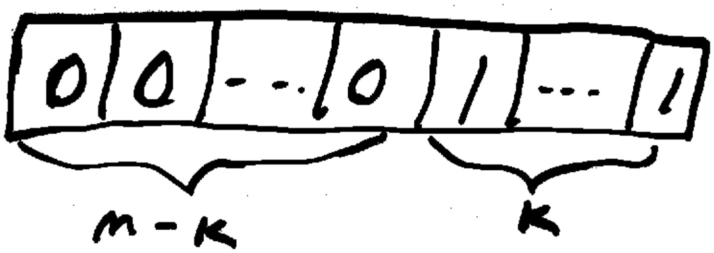
Then

$$(x_1 + \cdots + x_k)^n = \sum_{n_1 + \cdots + n_k = n} \binom{n}{n_1, \ldots, n_k} x_1^{n_1} \cdots x_k^{n_k}$$



people in boxes/cells

# of arrangements = $n!$



# of arrangements $\times (n-k)! \times (k!) = n!$

$\Rightarrow$ # of arrangements $= \dfrac{n!}{k!\,(n-k)!} = \dbinom{n}{k}$