Day 10 - October 2:

Continuation of previous ⑥

o Now suppose you are told that the city has 10 million people i.e 2 million have genetic defect. Find probability that 3 in a sample of ten will have the defect.

lution (1) Do exactly as the previous solution — City is much large now.

This would infact be an approximate solution since 10 outcomes in sample aren't Really independant cause sampling without replacement (we consider it is though cause big no's)

ution (2) We can calculate the exact probability as follows
See fish in Lake problem

o have a box with 10 million objects of which 2 million are Type 1 and 8 million of type 2.
Draw a sample of 10 without replacement
Then P(3 in 10 of Type 1).

$$= \frac{\binom{2\text{ mill}}{3}\binom{8\text{ mill}}{10}}{\binom{10\text{ mill}}{10}}$$

Compared to Solution 1 this is exact answer

Note: Too small of a probability eg (2/million), and first outcome having those 2, renders the situation dependant

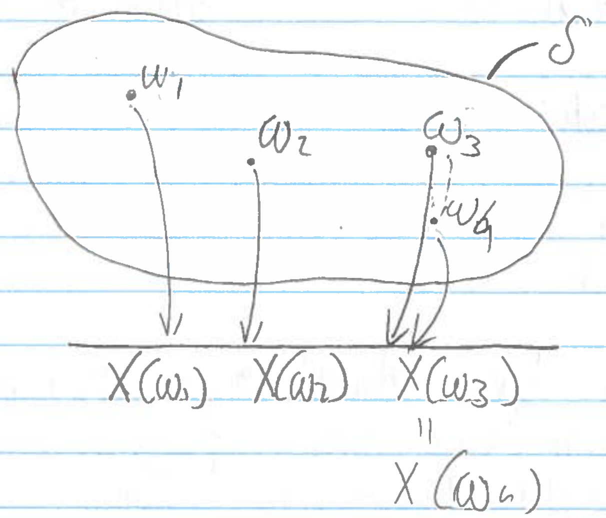
New Section, Random Variables : RV

Idea :- Sometimes not care about outcomes but no's associated with the outcomes example no of Red ball's in a sack of different colours

~~Function~~

Hence Random variable is a function that assigns a number each of outcome

Def : Let S be a sample space with outcomes ω. A Random variable is defined to be a real valued function, $X : S \longrightarrow \mathbb{R}$ (Real line) such that $X(\omega)$ is a real no for each $\omega \in S$.



$X(\omega_3) = X(\omega_4)$ means ω_3 can be mapped to same point ω_4

o RANDOM Variable (RV) is Big course part

Note o Usually RV's are denoted by capitals as in $X, Y, Z, \ldots$

Their realized values are denoted by lower cases x, y, z ---- ie once you have substituted a specific ω.

$P[X = x]$

- Random Variable is a function: Called Random cause before experiment the outcomes in S are uncertain. (They occur only with certain probability)
  - This uncertainty is transferred to a pre-experiment uncertainty in a value of the R.V which we'll see after experiment is complete
- RV's come in two main types 1) discrete, 2) Continous

Discrete : A RV X is said to be discrete if X can take on at most a contable set of real no's. or Range of X is countable

eg of Discrete:

- Toss a coin 10 times. let X be no of times you observe head.
- Here sample space consists of all vectors of length 10 that contain outcomes {H, T} in each position
  eg $\omega_1 = \{H, H, H \cdots H\}$ then X can assume values 0, 1 ---- 10 { ie the } range of X is {0, 1 —— 10}

Eg 2

- Toss a coin till you observe 1st head. let X be toss at which this occurs then the range of X = {1, 2 —— ∞}
  Then the range of X is {1, 2 ---} is infinite but countable

Eg3:

o Let X be the strength of an earthquake in a certain region measured to one decimal point of richter scale. Since X can take on countably many values, X is discrete.

In contrast a continous R.V is one that can assume ANY VALUE in some interval

Example

1) The exact value of an earthquake on the scale
2) The exact error in reading from altimeter on a plane

3) The time that someone survives after the onset of some disease.

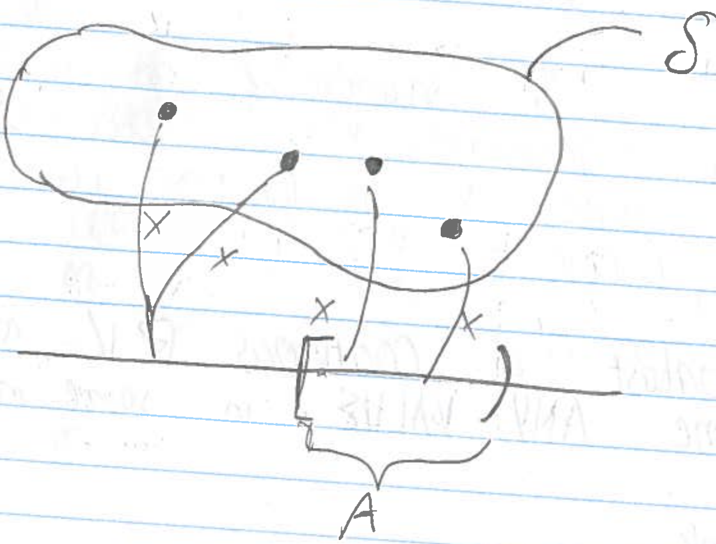
Big Question: How we specify RV.

o In regular math we specify function by giving its domain and values that function can assume. ie either through a formula or List.

When it comes to a RV, we need to specify the probabilities of all events in to events on real line into which the RV can fall

Picture →

Picture Representation:-



- We call the set $\{P(A) : A \text{ is a subset of } \mathbb{R}\}$ the probability distribution of the random variable

i.e The probability distribution is a specification of probability $P[X \in A]$ for all sets A in $\mathbb{R}$