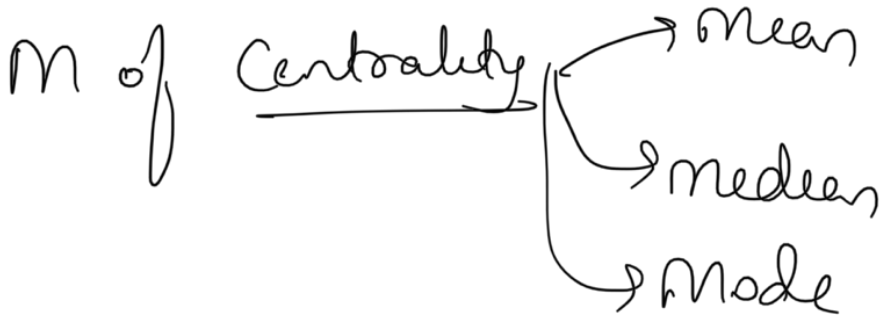
Measure of Centrality and Spread

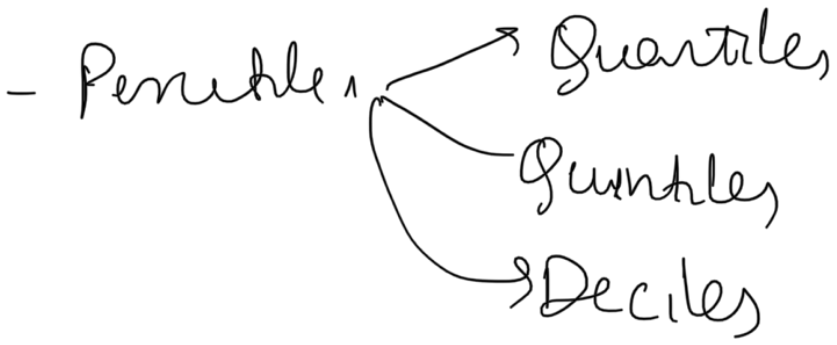
Parameter =

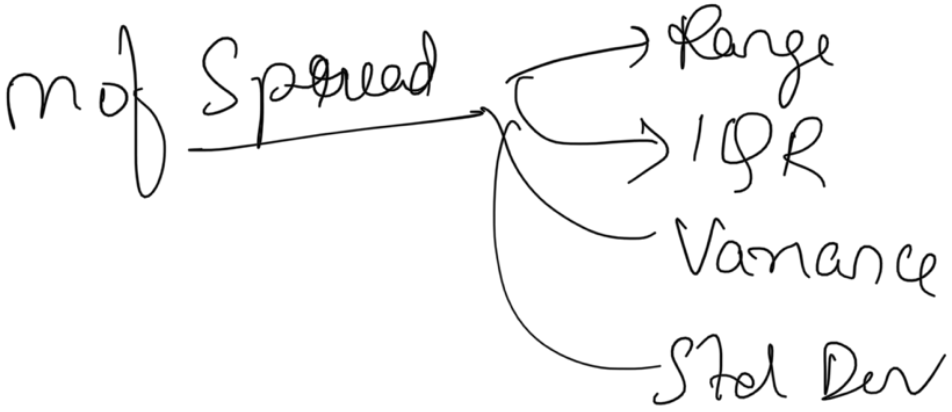
- numeric property of entire population to study

Statistic

[any numerical property of a sample.
for a parameter







Data, $x_1, x_2, x_3 \ldots\ldots x_n$

→ sum of all elements / data devided by elements

Mean

- sample $\bar{x}$
- populat μ

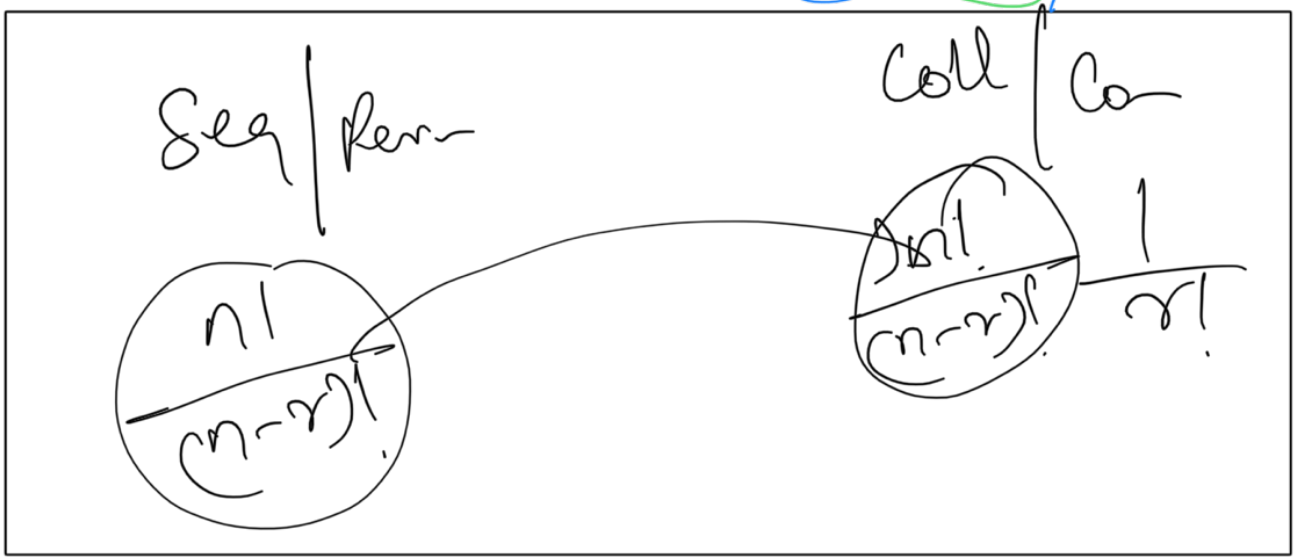
total no. $= \frac{1}{n} \sum_{i=1}^{n} x_i$

Median

- value appears at centre of data when data is sorted

If N = odd $= Med = x_{\frac{n+1}{2}}$ i.e. 1,2,3

Even $= \left(x_{\frac{n}{2}} + x_{\frac{n+1}{2}} \right) / 2$ 1,2,3,4,5,6,7 ; 1,2,3,4,5,6,7,8



Mode

- Most freq. occurg element in dataset

Mean

- centre of gravity of the data.

point / score

Deviation

- 1 [illegible] from mean is defined as

↳ of a point

difference between the score & mean

$$Dev = x_i - \bar{x}$$

Sum of deviations of all points from mean is zero

$$\sum_{i=1}^{n} (x_i - \bar{x})$$

$$x_1 - \bar{x} + x_2 - \bar{x} + \dots + x_n - \bar{x}$$

$$(x_1 + x_2 + \dots + x_n) - n \times \bar{x}$$

$$\sum_{1}^{n} x_i - n \times \frac{\sum_{1}^{n} x_i}{n}$$

$$= 0$$

Score x_i	Mean $\bar{x}$	Devian $x_i - \bar{x}$
8	9.67	-1.67
⋮	⋮	⋮

Outlier

↳ point which is far from other value in dataset

Sensitive to Outlier

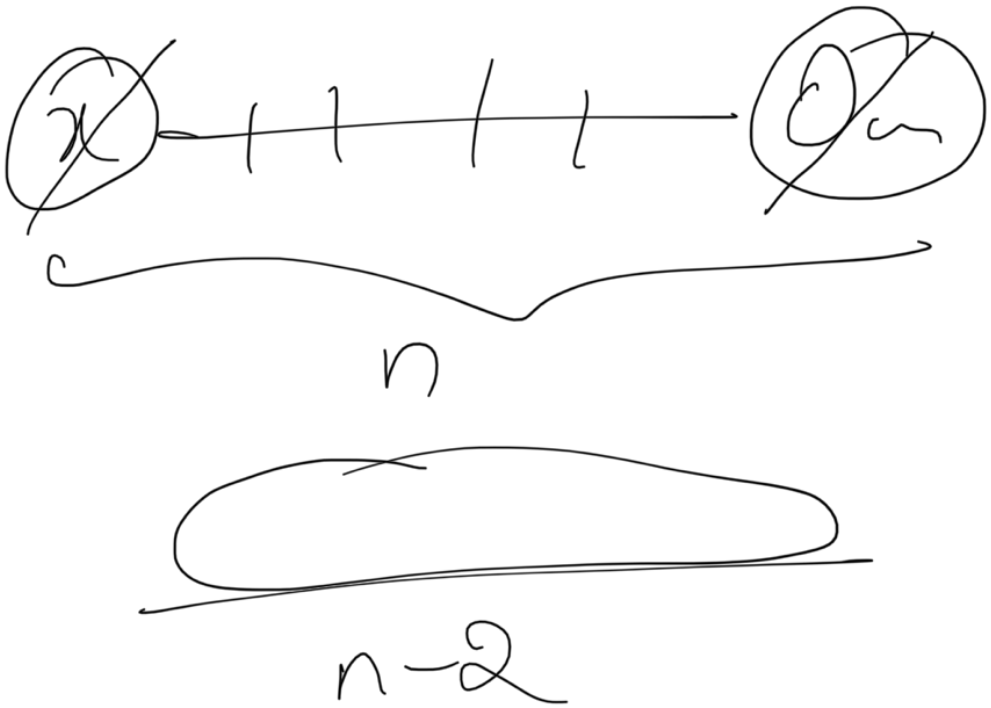
Mean — very sensitive

Median, Mode — not

To check consistency

- gap b/w Mean & median should be less ≈ 0

Trimmed mean



Trimmed Mean is calculated by dropping extreme elements from either side

- drop same no. of elements from both sides

Mode

→ is very NOT sensitive to outliers unless mode itself an outlier (very rare case)

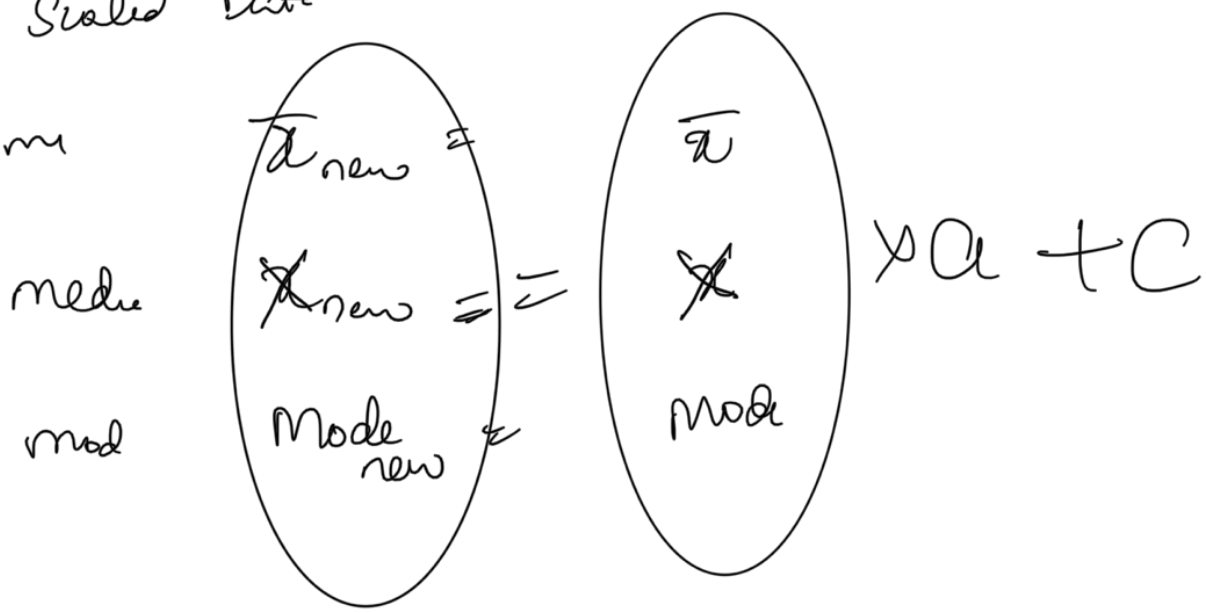
S

Effect of Transformation on Measure of Centrality

→ Scale & Shift

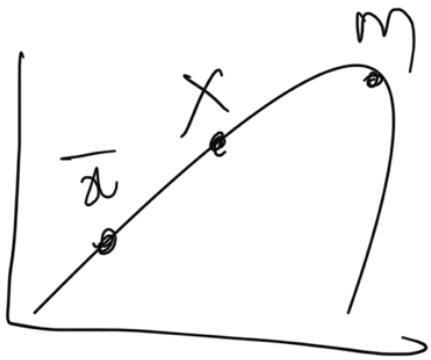
Ex Cm to mtr
F to C

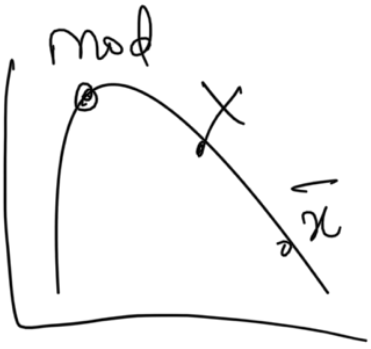
- Org Date
Scaled Date



Skewness

Left skew | Symmetric | Right skew





$\bar{x} < X < Mod$

$\bar{x} = X = Mod$

$\bar{x} > X > mod$

Percentile

① sort

② compute loc of p^{th} prod

$$l_p = \frac{P}{100}(n+1)$$

$l_p = i.f$

③ compute integ & frac part

④ comp

$$y_p = x_i + f_p(x_{i_{p+1}} - x_i)$$

pth percentile

Alternate

$\frac{p}{100} \times n$

if fraction next digit

if integer

$x_i + x_{i+1}$

Percentiles Spread & IQR

Quartiles

25^{th} pcile Q_1

50^{th} pcile Q_2

75^{th} pcile Q_3

→ divide data into 4 equal parts

note median = Q_2 50^{th} pcile

same

Quintiles

20^{th} 40 60 80

pcile

Q_1 Q_2 Q_3 Q_4

Deciles

10^{th} 20^{th} ——— 90^{th}

Q_1 Q_2 Q_9

Percentile rank of given score

$$PR_S = \frac{C_S + (0.5\ f_S)}{n} \times 100$$

C_S : no. of values < S

f_S : no. of values = S

n : percentile rank of 'S' score

round to nearest whole no. if fraction

Effect of Transformation

SD of Scale & Shift

$$Lp_{new} = Lp$$

no change

$$y_p^{new} = a \times y_p + C$$

Measure of Centrality not tells about spread & variance of data

$$\text{Range} = \text{Max value} - \text{min value}$$

Range also sensitive with outlier

$$IQR = Q_3 - Q_1$$

75% 25%

Dev

$$\sum_{i=1}^{n} (x_i - \bar{x}) = 0$$

$$\frac{1}{n} \left(\right) = 0$$

$$\frac{1}{n} \sum_{i=1}^{n} |x_i - \bar{x}|^2 \quad \text{Comp}$$

mee of spread

Variance

$$\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

S^2

σ^2

sample $(n-1)$

pop (N)

Std dev

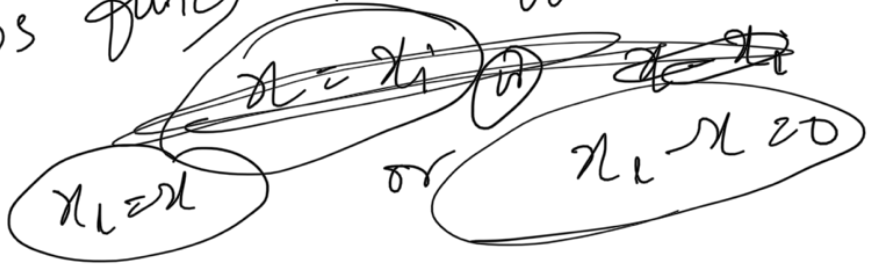
$S/\sigma = \sqrt{\text{variance}}$

Comp devat

why sq

sq. fn better prop

sq. fn (abs func) not differentiab @



sq func magnifies contribn of outliers

Ex: topic context

variance by sq > V by abs

COUNTING & Collections

Probability Theory

Goal: to study large collection of objects

- challenges: infeasible, expensive, time consuming

Sol: Survey only few ele

draw inference about all element from this group

Pop — total collect of objects to study

Sample : subgroup of population to draw inference abt pop

---

if no. of outcon 'n'

if equal chances:

probab 1/n

52 cards

Face cards

Ace (A) 2 3 4 5 6 7 8 9 10 Jack Q. K

Red ♡

Red ◇ Dia

Black ♤ spade

Black ♧ clubs

prop of getting 4 aces

$\frac{4}{52} \cdot \frac{3}{51} \cdot \frac{2}{50} \times \frac{1}{49}$

P1 Num of nums b/w 1 & n = n

P2 Num ——— in k & n = $n-(k-1)$

$n-k+1$

if frac given, in sum
convert to prop frac & multiply

P3 The Multiplication Principle

no. of ways of making seq. of indep choices is just product of no. of choices @ each step
(food idli dose combs)

P4 ——— Mul prin: Special case 1

no of seq. of k objects made from

given 'n' objects, when any object in seq. can be repeated any no of times is n^k

(Gym 10 days + 7 days) w rep

P5 : mul prnc : spec case 2

The no. of seq. of k objects made from given 'n' objects. such that no object in seq can be repeated is

$$n \times (n-1) \cdot (n-2) \times \cdots (n-(k-1))$$

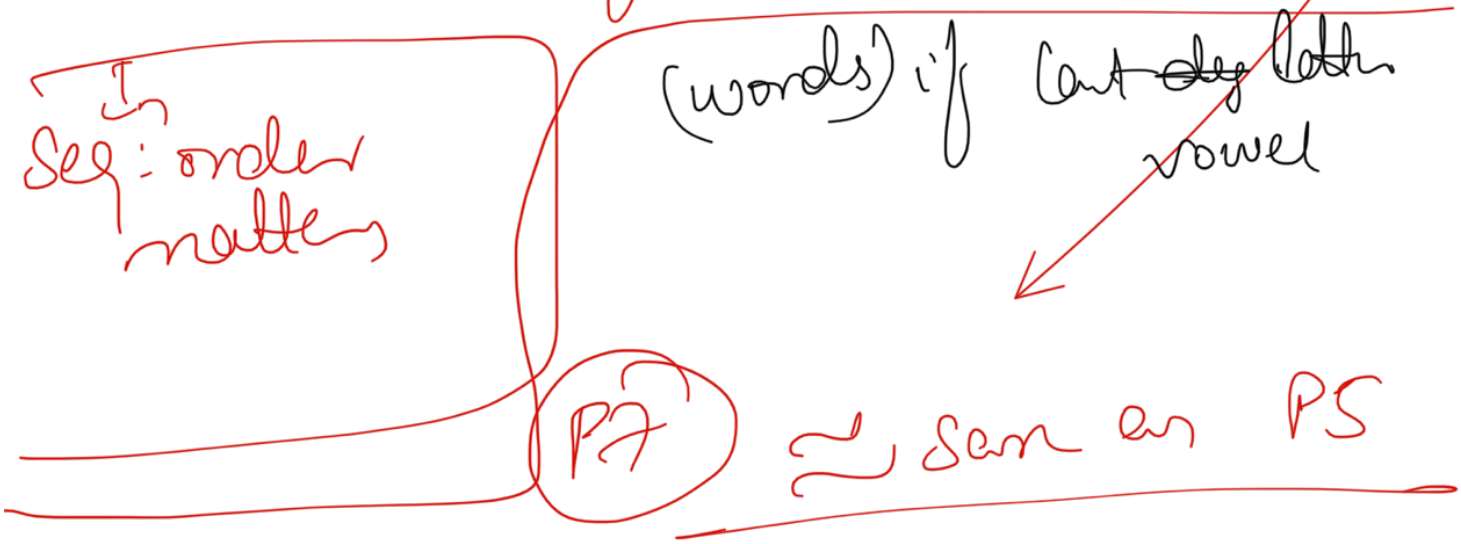
$$= {}^nP_k = \frac{n!}{(n-k)!}$$

(Gym 10 C w/o rep)

with
${}^nC_r = {}^nC_{n-r}$

(P6): Mult prnc spec ~~~

If problem specifies a constraint or restriction then always start by addressing the restriction first



The no. of ways of filling k named slots using collection of 'n' objects is same as no. of ways of creating a seq. of 'k' elements such no seq. can be

repeated in

$n(n-1)(n-2)\ldots\ldots(n-(k-1))$

$\Rightarrow {}^nP_r = \frac{n!}{(n-r)!}$

Ex: 111 PS committee of 4 from 15 $\Rightarrow P_1 VP_3 S_7 AS$

PS: The no. of ~~ways~~ n seq. of length n that can be formed using n objects, such that no objects in seq. is repeate in n! (factor of n)

- [no. of ways n obj can be arrange am themselves in n!]
- no. of perm of n obj in n!

Ex: 9 flower plot rearrange

SUBTRACTION PRINCIPLE

The no. of objects that satisfy some condⁿ = (is eq. to) total no. of object in collection − (minus) ones which not satisfy condⁿ

Ex: 3 letter word with atleast 1 vowel = All collection 3 letter word − collection with no vowel

Question with sol = ( ) − no 2 consecutive same

Sequence
Perm

order matters

cat ≠ act

n!

Collection
Comb

not

cat = act

$\frac{n!}{r!(n-r)!}$

$(n-k)!$

$k!$

rearrange k objs in $k!$ ways

$\binom{n}{k}$

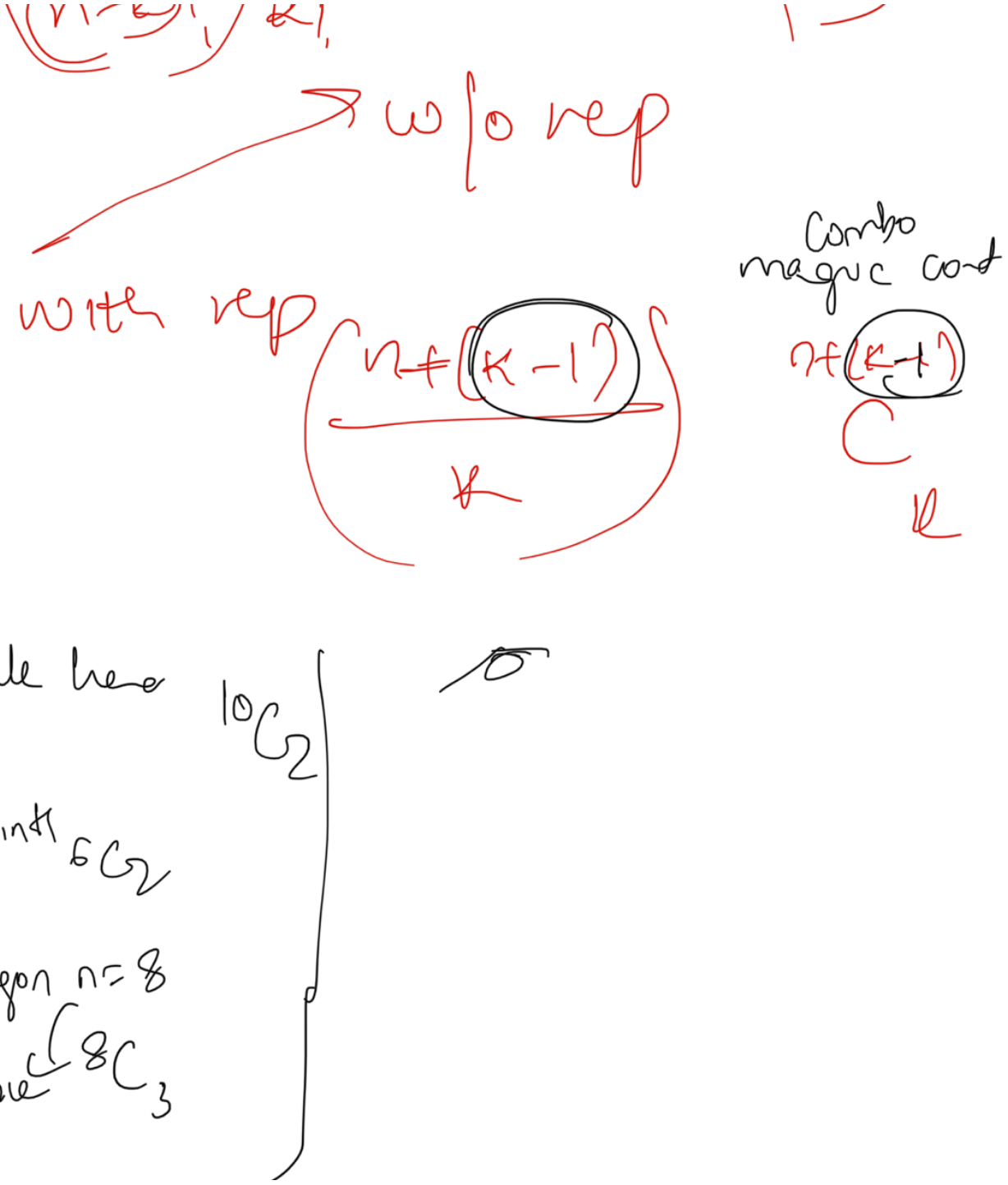
$^8C_2 = {}^8C_6$

Ex: of 15 stud, com of 4

Collection Principle

no. of ways of selecting k objects from n objects is

$\left(\frac{n!}{(n-k)!}\right)\frac{1}{k!}$ don = $\binom{n}{k}$


w/o rep
with rep
Combo
magic cord
shade here
10C2
6points 6C2
octagon n=8
true 8C3