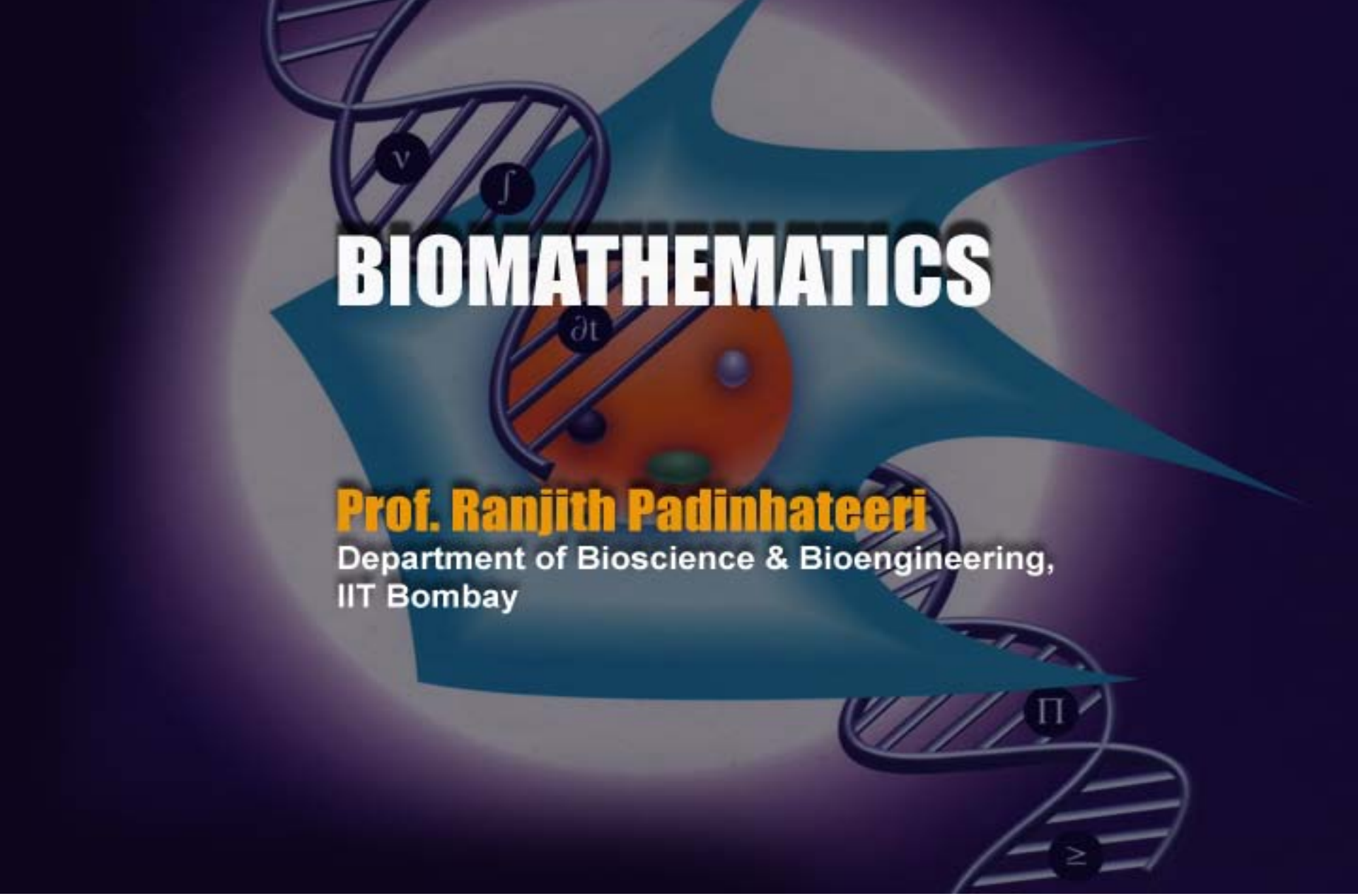
BIOMATHEMATICS
Prof. Ranjith Padinhateeri
Department of Bioscience & Bioengineering,
IIT Bombay

Lecture 21

Statistics

Experiments/simulations produce large amount of data.

How do we make sense out of these numbers ?

We often say:

"It takes 30 minutes to travel to my college, from my home."

What do we mean here ? (and what we do not mean ?!)

We do not mean that, each and every day, we take precisely 30 minutes.

What we mean is that, on an average,
it will take 30 minutes to reach the college.

Here we are using the idea from
statistics called “average” or “mean”

Travelling to college as an experiment

Each day note the time it takes to reach college

Day	Time
Day 1	27 minutes
Day 2	33 minutes
Day 3	31 minutes
Day 4	29 minutes

Average/mean time to reach college

Day	Time (min)
Day 1	27
Day 2	33
Day 3	31
Day 4	29

$$\langle t \rangle = \frac{27+33+31+29}{4} = 30$$

Average over many experiments!

Average/mean

$$\langle t \rangle = \frac{t_1 + t_2 + t_3 + \ldots + t_N}{N}$$

$$\langle t \rangle = \frac{1}{N} \sum_{i=1}^{N} t_i$$

Average mark

What do we mean when we say :

I have got 70% mark for my class 12 exam

We mean, average of all subjects
is 70 out of 100

Average mark

Case 1

Biology: 75
Physics: 65
Chemistry: 70

Average=(75+65+70)/3=70

Average mark

Case 2

Biology: 95
Physics: 40
Chemistry: 75

Average=(95+40+75)/3=70

What is the difference ?

Biology: 75
Physics: 65
Chemistry: 70

Biology: 95
Physics: 40
Chemistry: 75

Since average is the same for both the cases, how do we differentiate the two cases ?

Answer: “standard deviation”

Deviation from the average

Average marks = 70

Mark	Deviation= Mark-average
75	5
65	-5
70	0

Mark	Deviation= Mark-average
95	25
40	-30
75	5

Sum of the deviations is zero!

Deviation from the average

Average marks = 70

Mark	(Deviation)^2
75	$(5)^2=25$
65	$(-5)^2=25$
70	$(0)^2=0$

Mean of deviation square $\langle D^2 \rangle = \frac{25+25+0}{3} = 16.667$

Standard deviation $\sigma = \sqrt{\langle D^2 \rangle} = 4.08$

Deviation from the average

Average marks = 70

Mark	(Deviation)^2
95	$(25)^2=625$
40	$(-30)^2=900$
75	$(5)^2=25$

Mean of deviation square

$$\langle D^2 \rangle = \frac{625+900+25}{3} = 516.667$$

Standard deviation $\sigma = \sqrt{\langle D^2 \rangle} = 22.73$

Deviation from the average

Average marks = 70

Case1
75
65
70

Case2
95
40
75

70 ± 4.08 70 ± 22.73

Variance= $\frac{1}{N}\sum_i (x_i - \langle x \rangle)^2$

$$= \frac{1}{N}\sum_i \left[x_i^{\,2} + \langle x \rangle^2 - 2x_i \langle x \rangle \right]$$

$$= \langle x^2 \rangle + \langle x \rangle^2 - 2\langle x \rangle \frac{1}{N}\sum_i x_i$$

$$where, \langle x \rangle = \frac{1}{N}\sum_i x_i$$

Variance $= \langle x^2 \rangle + \langle x \rangle^2 - 2\langle x \rangle^2$

$= \langle x^2 \rangle - \langle x \rangle^2$

Standard Deviation $\sigma = \sqrt{Variance}$

Graphical representation

If you were to mark these numbers on a graph, how would you think about average and standard deviation ?

Summary

Average: $\langle x \rangle = \frac{1}{N} \sum_i x_i$

Standard deviation: $\sqrt{\langle x^2 \rangle - \langle x \rangle^2}$