Distributions

class – 7 (28.8.24)

LS2103 (Autumn 2024)

Dr. Neelanjana Sengupta

Associate Professor, DBS

https://www.iiserkol.ac.in/~n.sengupta/

Continuous distributions

Consider a quantity f that depends on the variable x, *ie.*

$$f \longrightarrow f(x)$$

The mean value of the f,

$$\bar{f} = \int_{x_{low}}^{x_{high}} f(x)\, p(x)\, dx$$

The mean value of the f^2,

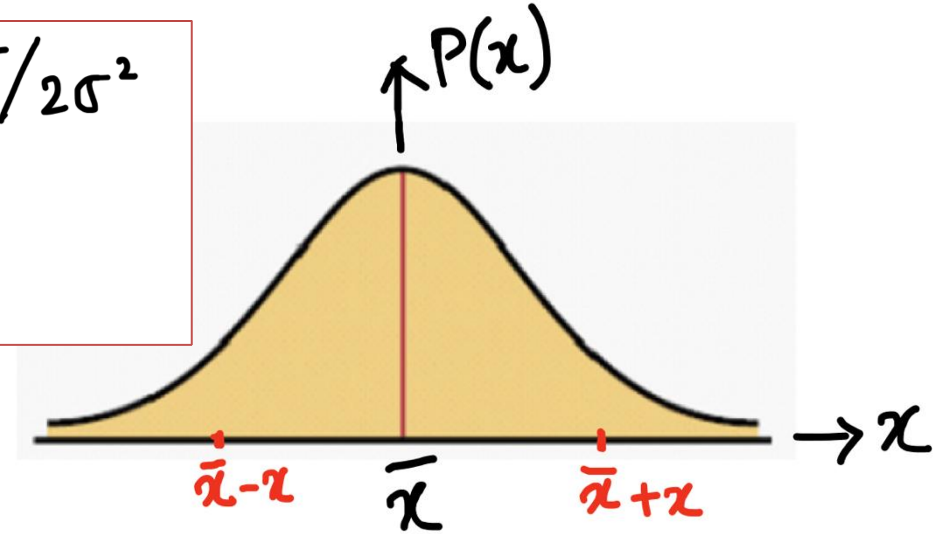
$$\overline{f^2} = \int_{x_{low}}^{x_{high}} [f(x)]^2\, p(x)\, dx$$

The variance,

$$\sigma_f^2 = \overline{f^2} - (\bar{f})^2$$

Gaussian or Normal Distribution

$$P(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\bar{x})^2/2\sigma^2}$$



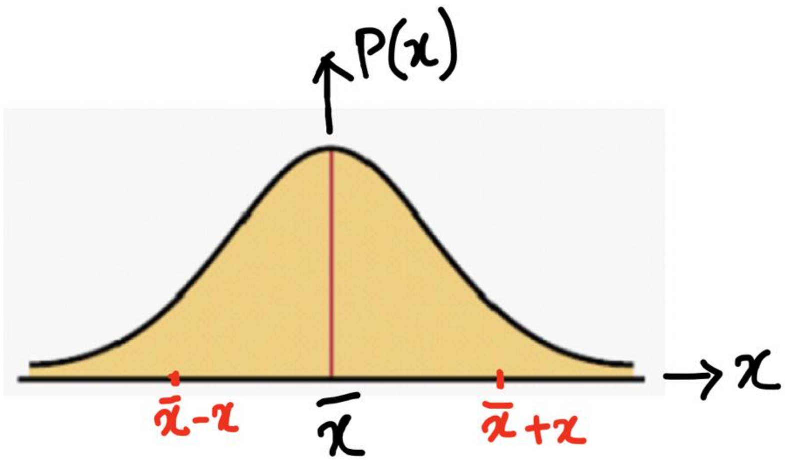
1) Symmetric function: $P(\bar{x}+x) = P(\bar{x}-x)$

2) If centered at the origin, this becomes an *even function*

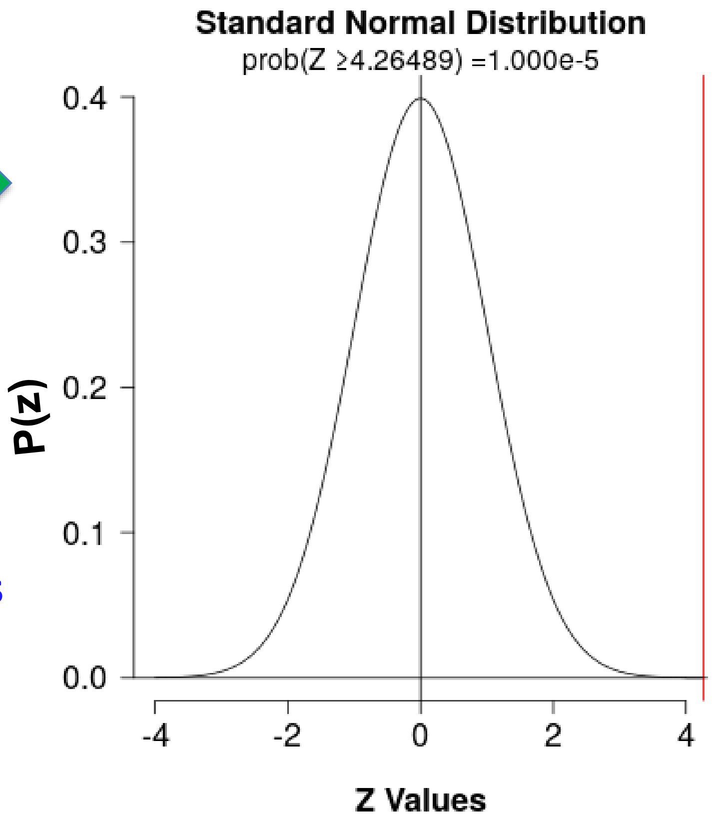
3) Full width and half maximum (FWHM) depends on σ

$$\text{FWHM} = 2\sqrt{2\ln 2}\,\sigma \approx 2.355\,\sigma.$$

Gaussian or Normal Distribution



μ = 0; σ = 1.0



4) For any data point x_i, the no. of standard deviations away from the mean is,

$$Z_i = \frac{x_i - \mu}{\sigma}$$

Prob. In treatment of colorectal cancer, the cellular response to irradiation varies depending on the expression of tumor suppressor p53, which can be estimated by the corresponding "DNA content".

3 stages of cell growth are labelled growth-1 (G1), suppression (S) and growth-2 (G2).

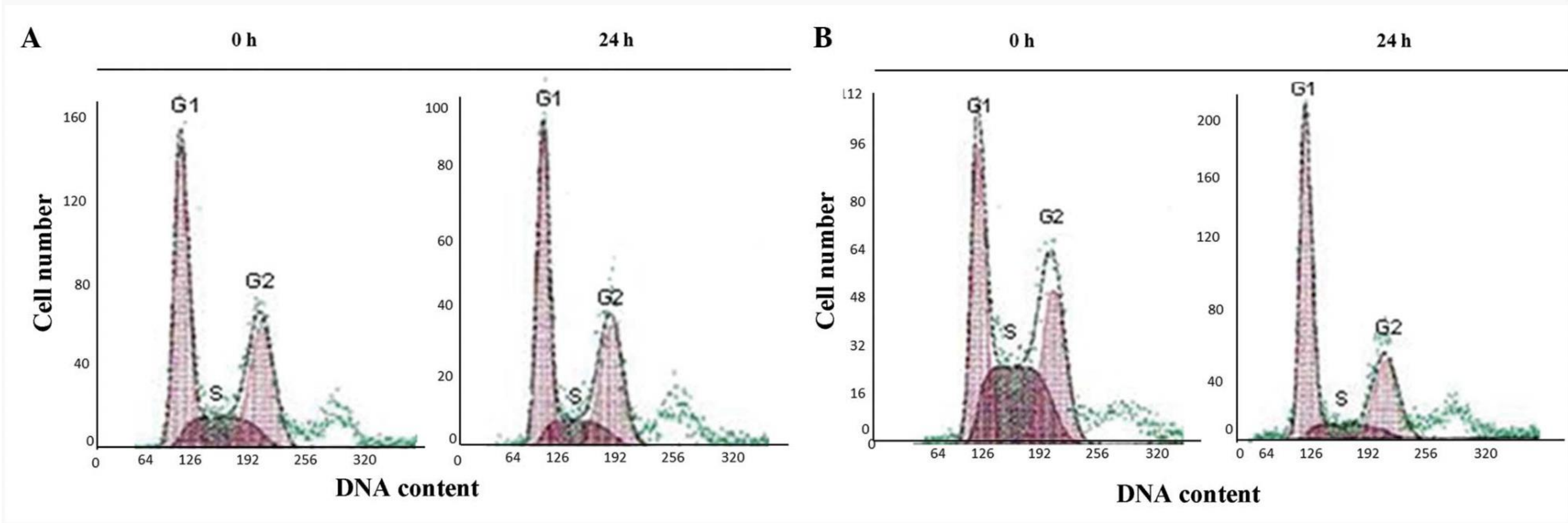
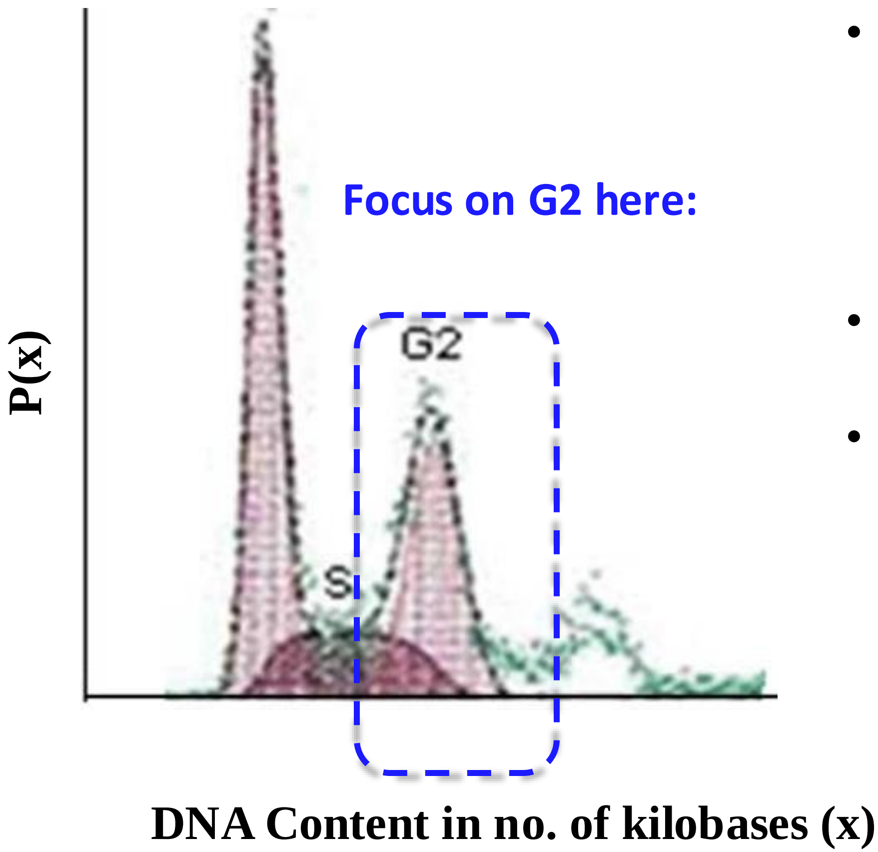


Figure 2 - Normal distribution of G1, S and G2 stages of the cell cycle in non-irradiated (A) p53 wild-type (+/+) and (B) p53 deficient (–/–) HCT116 cells at 0 and 24 h.

Halacli et al., Oncololgy Lett 2013, https://doi.org/10.3892/ol.2013.1441

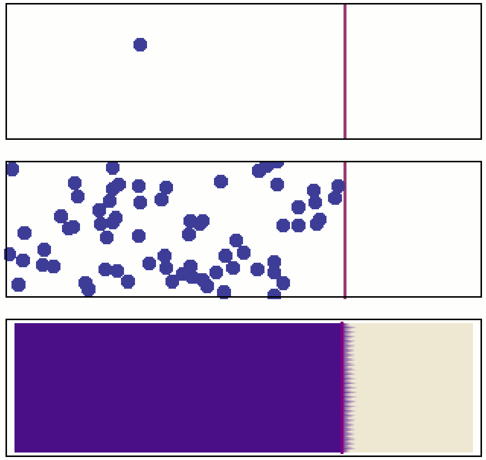
Prob. In treatment of colorectal cancer, the cellular response to irradiation varies depending on the expression of tumor suppressor p53, which can be estimated by the corresponding “DNA content”.



- Let’s hypothesize that patient survival (in years) scales with square of the DNA content (x) in the G2-phase, with a suitable prefactor, A = 25
- **Find the mean patient survival time.**
- What are the units of A?

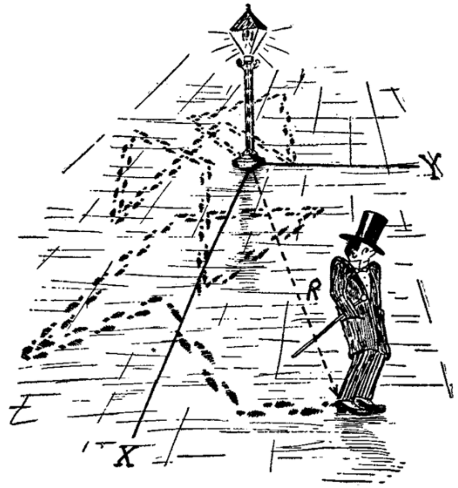
The raw data was converted to a zero-centered normal distribution with $\sigma = 1$

Random Processes: Biology at microscopic scales



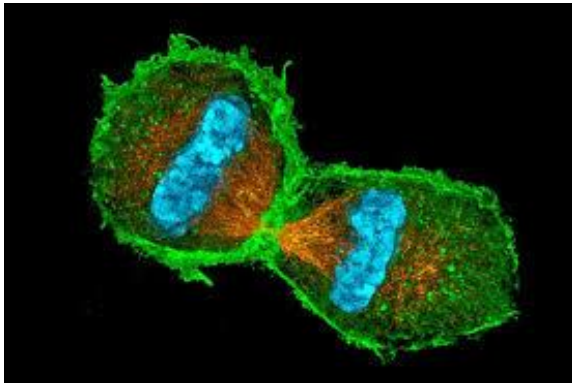
Diffusion

https://commons.wikimedia.org/w/index.php?curid=8995324



Random Walk

George Gamow, 1961

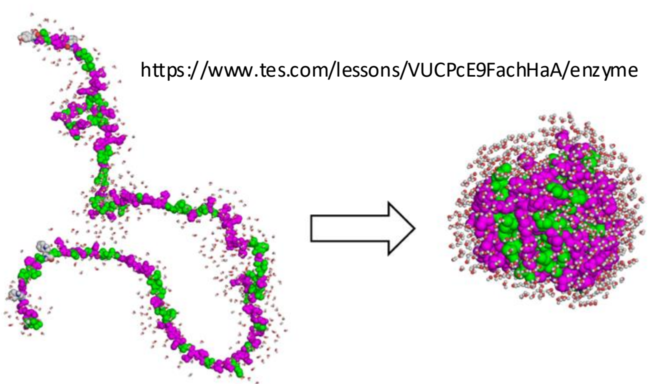


Cell division

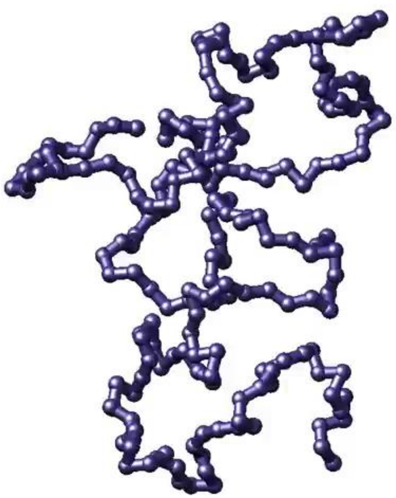
https://www.thoughtco.com/mitosis-and-cell-division-quiz-4078417

Polypeptide collapse

https://www.tes.com/lessons/VUCPcE9FachHaA/enzyme



Unfolded **Folded**



Polymer collapse

https://www.youtube.com/watch?v=ek9vd7_PDjk

Binomial Distribution

Only **2 possible outcomes** of an event with ***N*** attempts:

- **Success**, probability ***s***
- **Failure**, probability ***f = (1 – s)***

Mean success, $\langle n \rangle = s \cdot N$

n = No. of success

Probabity of ['n' success and (N−n) failure] is,

$$P(n, s, N) = \frac{N!}{n!\,(N-n)!}\, s^{n} (1-s)^{N-n}$$

The sum of probabilities for the (n, s, N) is 1·0, ie.

$$\sum_{n=0}^{N} P(n, s, N) = 1 \cdot 0$$

Binomial Distribution

Only **2 possible outcomes** of an event with *N* attempts:
N → trials .

Success, probability ***s***

Failure, probability ***f = (1 – s)***

Mean success, $\langle n \rangle = s \cdot N$

n = No. of success

Variance in success, $\sigma^2 = s(1-s)N$

Standard deviation, $\sigma = \sqrt{s(1-s)N}$

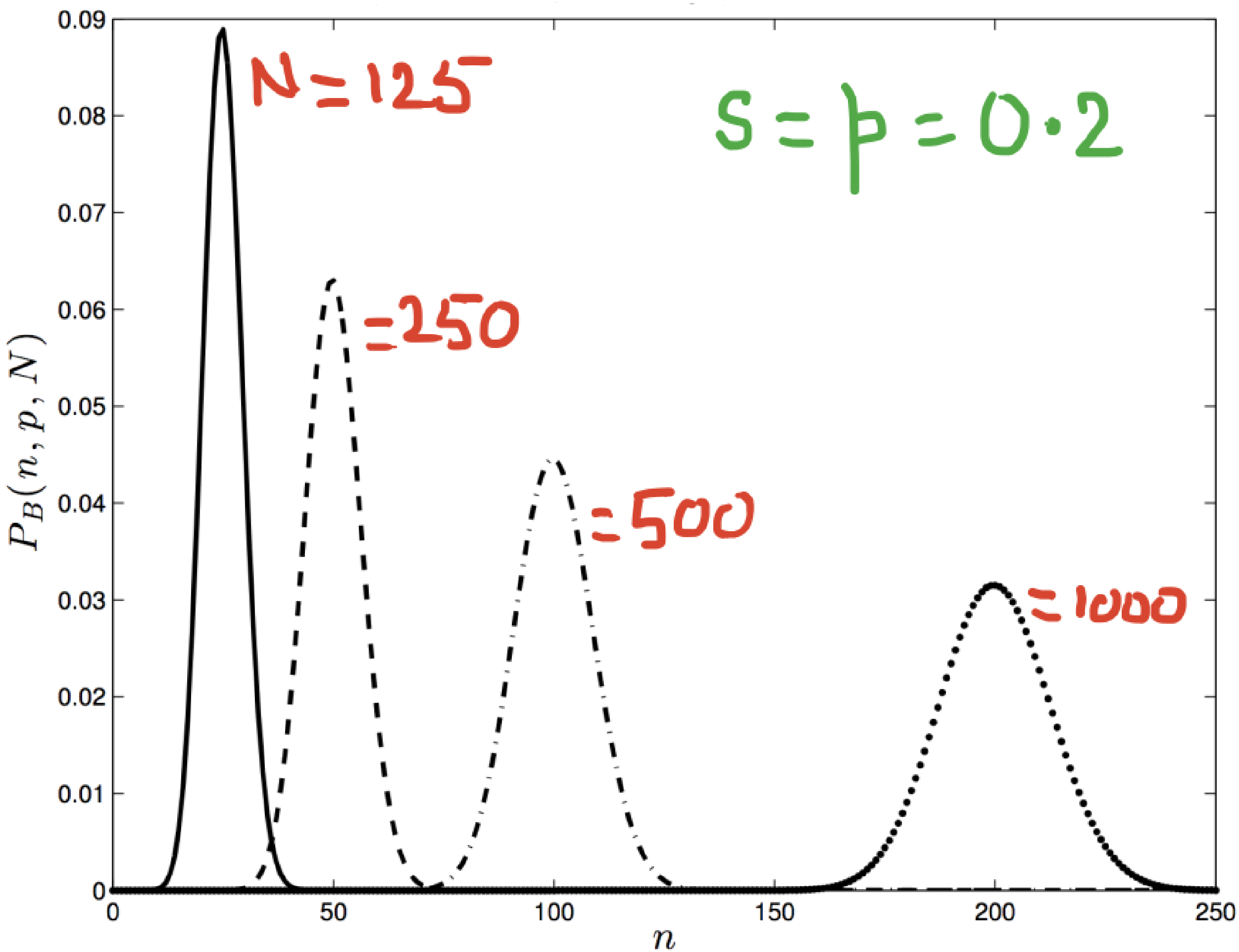
The ratio of std. dev. to mean $= \sqrt{\dfrac{1-s}{sN}} \propto \dfrac{1}{\sqrt{N}}$

$r = \dfrac{\sigma}{\mu}$

Binomial disbn. approaches Normal disbn. at high N

Only **2 possible outcomes** of an event with ***N*** attempts:

- **Success**, probability ***s (= p)***
- **Failure**, probability ***f = (1 – s)***



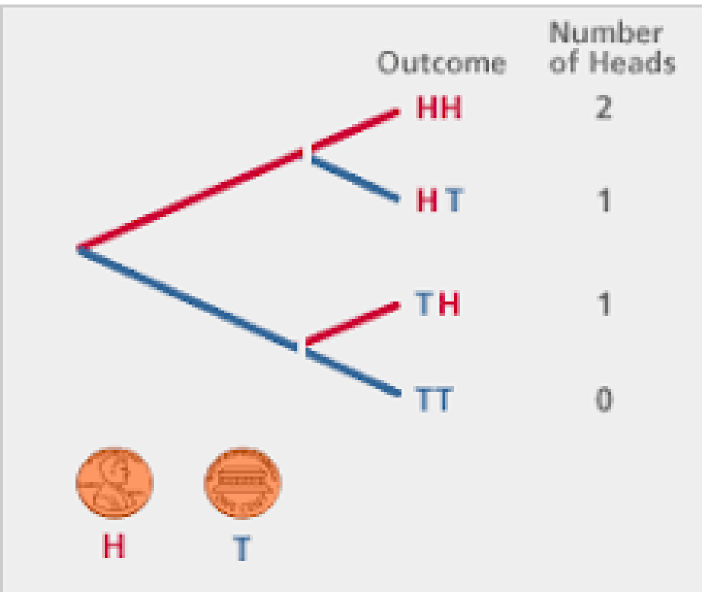
n = No. of success

Prob. disbn. of one outcome, say "success"

Only **2 possible outcomes**
of an event with ***N*** attempts,
with ***n*** no. of success

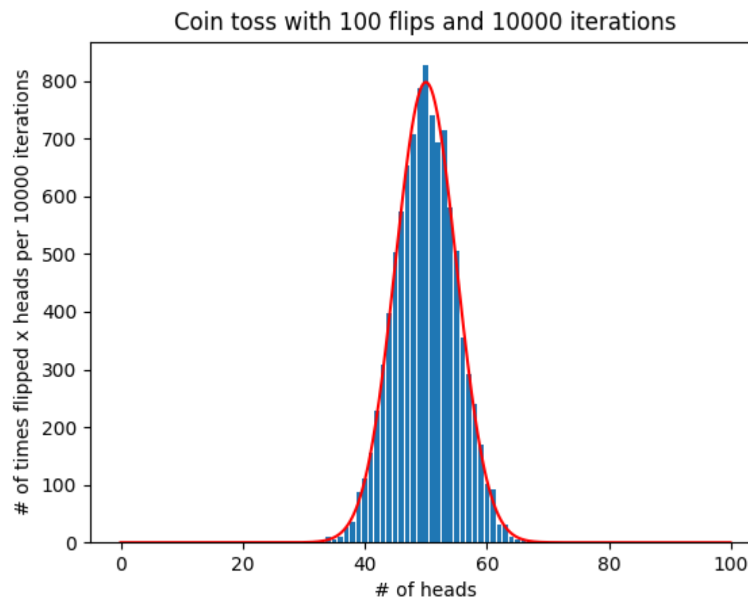
Binomial Distribution

Gaussian Distribution



Consider toss of *unbiased* coin:

$$P_{head} = P_{tail} = 0.5$$



nearly exact 50% head only for N_toss >> 1

Binomial Distribution

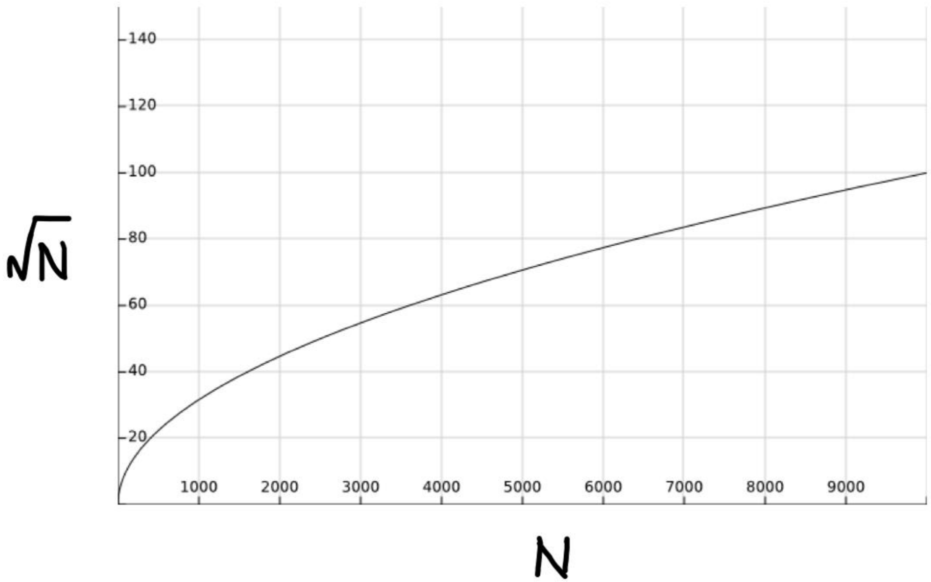
Only **2 possible outcomes** of an event with ***N*** attempts:

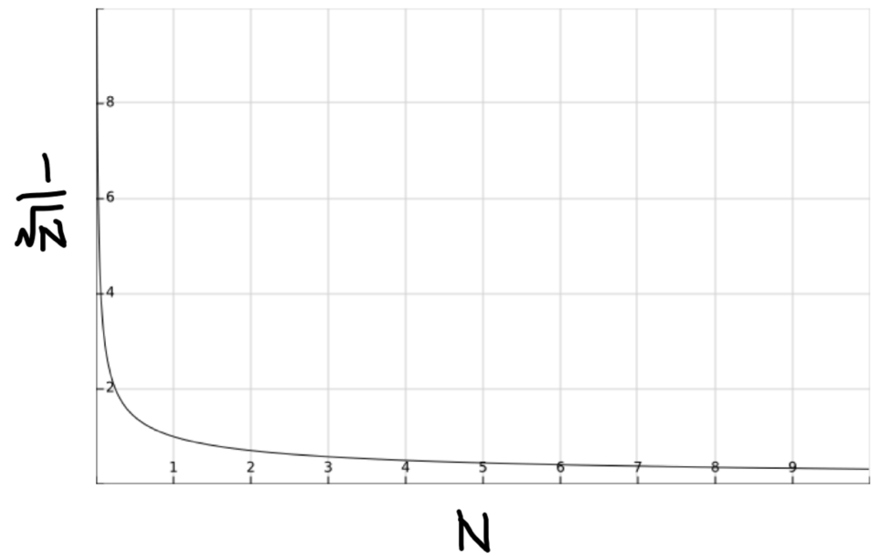
Success, probability ***s***

Failure, probability ***f = (1 – s)***

n = No. of success

Note:





Binomial Distribution

N no. of gas molecules under normal conditions is left in a cubical box marked in two halves.

- Find the *ratio* of the *uncertainty to the mean probability* of finding any one molecule

in one of the halves.

- How does the ratio vary when:

a) $N = 1000$

b) $N = 6 \times 10^{23}$