Extension: Central Limit Theorem

- **What is it?**

It says[+]: Many independent random variables x_n each has finite expected value (mean) and variance, form a variable

$$X = \sum_{n=1}^{N} x_n$$

for which $P(X)$ is a Gaussian distribution function in the limit $N \to \infty$.

$$P(X) = \frac{1}{\sqrt{2\pi V_N}} e^{-\frac{(X-E_N)^2}{2V_N}}$$

$$E_N = \text{Expected value of } X \equiv \overbrace{\sum_{n=1}^{N}}^{\text{sum of means}} e_n$$

expected value of x_n (might call it μ_n)

$$V_N = \text{Variance (of } X) \equiv \underbrace{\sum_{n=1}^{N}}_{\text{sum of variances}} v_n$$

variance of x_n (might call it σ_n^2)

Valid when $V_N \to \infty$, $\frac{v_N}{V_N} \to 0$ as $N \to \infty$.

[+] CLT is not stated rigorously here (about the conditions). The point is to convey what CLT says. The conditions are usually valid in physics contexts.

Illustration: Back to spatial distribution of gas molecules

- Consider one molecule: Prob. p in fictitious box
  Prob. $1-p=q$ not in box

e = mean [what one molecule contributes to the number of molecules in box on average]

$$= \underset{\text{in}}{p \cdot 1} + \underset{\text{not in}}{(1-p)\cdot 0} = p$$

$$v = \text{Variance} = \langle (x-e)^2 \rangle = \langle x^2 \rangle - e^2 = \langle x^2 \rangle - p^2$$

$$\langle x^2 \rangle = p \cdot 1^2 + (1-p)\cdot 0^2 = p$$

$$\therefore v = p - p^2 = p(1-p) = pq$$

- Now we have $N\ (\gg 1)$ molecules

$$X = \sum_{n=1}^{N} x_n \quad \text{(variable giving \# molecules in box)}$$

$$E_N = \langle X \rangle = \sum_{n=1}^{N} e_n = Np \quad \text{(same result as before)}$$

$$V_N = \sum_{n=1}^{N} v_n = Npq \quad \text{(same result as before)}$$

Note that $\frac{v_n}{V_N} = \frac{1}{N} \to 0$ as $N \to \infty$ and $V_N = Npq \to \infty$ as $N \to \infty$

By CLT, Variable X follows the distribution function:

$$P(X) = \frac{1}{\sqrt{2\pi V_N}} e^{-\frac{(X-\langle X \rangle)^2}{2V_N}} \quad \text{as } N \to \infty$$

(same result as before)

Other cases:

- Random Walk: $\underbrace{\text{Each } x_n \text{ is a variable}}$

prob. p (or ½) walk to the right

prob. $(1-p)$ (or ½) walk to the left

$X = \sum_{n=1}^{N} x_n$ variable giving location after N steps

Try it many times (each N steps), end points fall on different places

E_N = mean position of end point $= N\cdot(p-q)$ ($=0$ for $p=½$)

V_N = Variance of end point positions $= N\cdot 4pq$ ($=1$ for $p=½$)

$SD = \sqrt{V_N} = \sqrt{4pq}\cdot\underbrace{\sqrt{N}}_{\sqrt{\text{time}}}$

On average how far from origin after time (N steps)

Diffusion