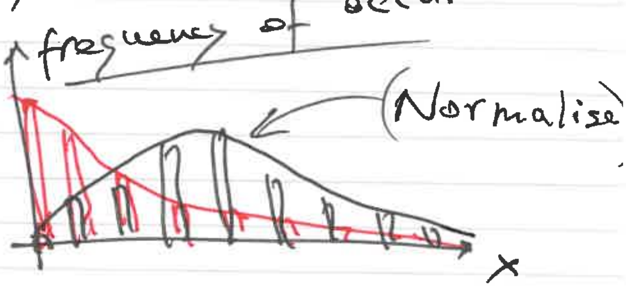
Lecture 2 Basic terms and definitions of probability theory.

- "Random", or "Stochastic" variable: X

Probability $P(x)$:



Such that

$$\langle x \rangle = \int x\, P(x)\, dx \qquad \int P(x)\, dx = 1$$

Also $$\langle \Delta x^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2 \equiv \sigma^2$$

Variance

- Characteristic function (of X)

$$\phi_X(k) \overset{\text{def.}}{=} \langle e^{ikx} \rangle_{P(x)}$$

This is just a FT of $P(x)$:

$$\phi_X(k) = \int e^{ikx} P(x)\, dx$$

$$= \int \sum_n \frac{1}{n!} (ikx)^n P(x)\, dx$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} (ik)^n \langle x^n \rangle$$

Moments of $P(x)$

Separately, define:

$$\phi_x(k) = \exp\left[\sum_{m=0}^{\infty} \frac{(ik)^m}{m!} C_m\right]$$

cumulants of $P(x)$

So that

$$\boxed{C_n = \frac{d^n}{d(ik)^n} \ln \phi_x(k) \Big|_{k=0}}$$

Now compare ...

$n=1$

$$C_1 = \frac{d}{d(ik)} \Big/ \sum_n \frac{(ik)^n}{n!} C_n \Big|_{n=1} \quad (1)$$

$$= \frac{d}{d(ik)} \ln \sum_n \frac{1}{n!} (ik)^n \langle x^n \rangle \Big|_{k=0} \quad (2)$$

cancel {n} and renumber the sum in numerator

$$= \frac{\sum_n n \cdot \frac{(ik)^{n-1}}{n!} \langle x^n \rangle}{\sum_n \frac{(ik)^n}{n!} \langle x^n \rangle} = \langle x^1 \rangle = \langle x \rangle$$

Check yourself:

$$C_2 = \langle x^2 \rangle - \langle x \rangle^2 \ldots$$

If $P(x)$ is Gaussian, then $\phi_x(k)$ is also: $P(x) \sim e^{-\frac{(x-x_0)^2}{2\sigma^2}}$

$$\phi_x(k) \approx \int e^{ikx} e^{-\frac{(x-x_0)^2}{2\sigma^2}} dx \Rightarrow \boxed{e^{-\frac{\sigma^2}{2}k^2 + ikx_0}}$$

⊙ If there is a linked sequence of random variables:

$$Y_N = X_1 + X_2 + \dots X_N$$

Each of these has same $P(x)$

is called "Random walk"

or

"Stochastic process"

Then: $\langle Y_N \rangle = N \langle x \rangle$

Variance: $\langle \Delta Y_N^2 \rangle = \langle Y_N^2 \rangle - \langle Y_N \rangle^2$

$$= \sum_{i,j} \langle x_i x_j \rangle - \sum_i \langle x_i \rangle \sum_j \langle x_j \rangle$$

$$= \sum_i \langle x_i^2 \rangle + \sum_{i \neq j} \langle x_i x_j \rangle - N^2 \langle x \rangle^2$$

$$= N \langle x^2 \rangle + N(N-1) \langle x \rangle^2 - \cancel{N^2 \langle x \rangle^2}$$

$$= N \left(\langle x^2 \rangle - \langle x^2 \rangle \right) = N \sigma^2$$

Central Limit Theorem

Take a "normalised" random walk

$S_N = \frac{y_N}{N}$ such that $\langle S_N \rangle = \langle x \rangle$

and Variance $= \frac{\sigma^2}{N}$

for large N: $P(S_N)$ converges to a Gaussian

Proof via characteristic function:

$$\phi_S(k) = \langle e^{iks} \rangle_{P(s)} = \underbrace{\langle e^{\frac{ik}{N}\sum_m x_m} \rangle}_{\text{product of } N \text{ exponentials}}$$

$$= \langle e^{\frac{ik}{N}x} \rangle^N_{P(x)}$$

But this is the characteristic function $\phi_x(k/N)$ raised to N ...

Separately:

$$\phi_S(k) = \exp\left[\sum_{m=0}^{\infty} \frac{(ik)^m}{m!} c_m(s)\right]$$

$$= [\phi_x(k/N)]^N$$

Now, write $c_n(x)$ for $\phi_x(k/N)$:

$$\ldots = \left[\exp\left(\overbrace{\sum_{n=0}^{\infty} \frac{(ik)^n}{n!\,N^n} c_n(x)} \right) \right]^N$$

$$= \exp\left[\sum_{n=0}^{\infty} \frac{(ik)^n}{n!\,N^{n-1}} c_n(x) \right]$$

product of exponentials ....

$\rightarrow$ $n=1$ $\qquad \exp\left(ik\, c_1(x) \right)$

See the added "note" at the end

$\rightarrow$ $n=2$ $\qquad \exp\left(-\frac{k^2}{2N} c_2(x) \right)$

No need to go further if $N \gg 1$...

this cuts out the range of possible "k" when $|k| > \sqrt{N}$, this $\exp(-k^2/N) \longrightarrow 0$

Then:

$$\phi_s(k) = \exp\left(N + ik\, c_1(x) - \frac{k^2}{2N} c_2(x) + \ldots \right)$$

(irrelevant if $|k| < \sqrt{N}$)

And we recover $P(s)$:

$$P(s) \overset{\text{def.}}{=} \int e^{-iks} \phi_s(k)\, dk$$

over the range $|k| < \sqrt{N}$

$$P(S) \approx \int e^{-ikS + ik\langle x\rangle} \cdot e^{-\frac{k^2}{2N}\sigma^2}\, dk$$

$$= \underline{\underline{\text{const}}} \cdot e^{-\frac{(S-\langle x\rangle)^2}{2\sigma^2/N}}$$

Extend range (k) to ∞, ignoring higher terms

So if is Gaussian with

$$\langle S\rangle = \langle x\rangle$$

$$\mathrm{var}(S) = \sigma^2/N$$

⊙ Replace N with time t

$\longrightarrow$ Stochastic process $y(t)$

(random walk $y_N = \sum_{m=0}^{N} x_m$)

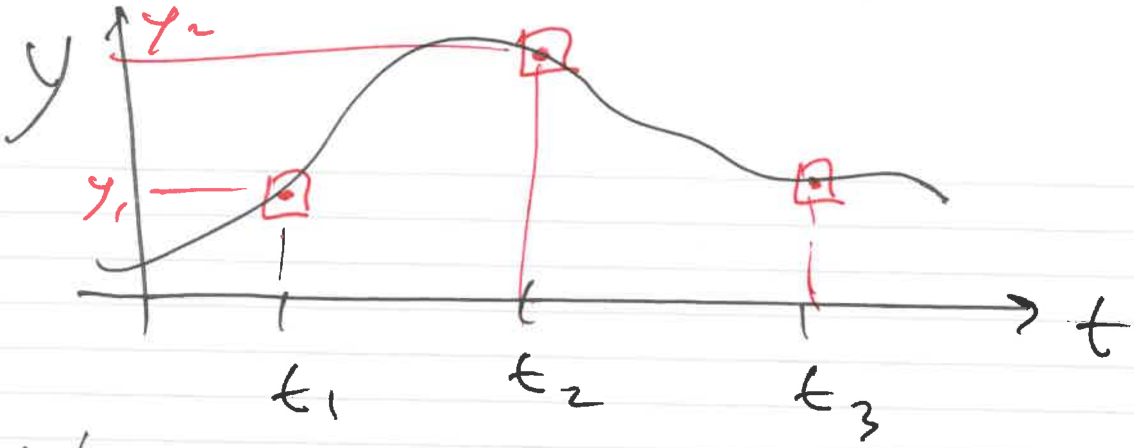
and stay with continuous t...

Define set of probalities:

$P_1(y,t)$: $\boxed{\text{reach value } y \text{ in } t}$

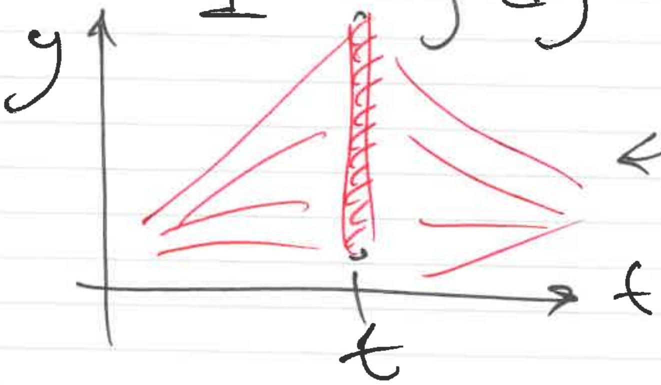
$P_2(y_1 y_2 \mid t_1 t_2)$: (simultaneous) y_1 at t_1 and y_2 at t_2

etc.



Normalisation:

$$1 = \int dy \, P(y,t)$$



Also: $$1 = \int dy_1 \, dy_2 \ldots P(y_1 y_2 \ldots | t_1 t_2 \ldots)$$

⊙ Reduction:

$$\int dy_N \, P_N(y_1 y_2 \ldots y_N | \ldots) = P_{N-1}(y_1 y_2 \ldots y_{N-1})$$

i.e. $$\boxed{\int dy_2 \, P_2(y_1 y_2 | t_1 t_2) = P_1(y_1 t)}$$

⊙ Correlation function:

$$\langle y(t) \rangle = \int y \cdot P(y,t) \, dy$$

$$\langle y_1(t_1) \, y_2(t_2) \rangle = \int dy_1 \, dy_2 \; y_1 y_2 \; P_2(1,2)$$

etc...

⊙ Stationary process $y(t)$

is when $P(y,t)$ is

invariant with time shift

$t \to t+\Delta t$

$\langle y(t) \rangle$ is not t-dependent

$\langle y(t_1)\, y(t_2) \rangle \sim f(t_1 - t_2)$

(so that $t_1 + \Delta t$ and $t_2 + \Delta t$)

etc.

⊙ Conditional probability

= Propagator

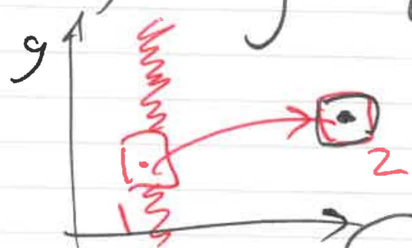
$$P_2(y_1 y_2 | t_1 t_2) = \boxed{G[2|1]}\, P_1(y_1 | t_1)$$

Simultaneous (pair) probability $P(1,2)$

- Markov process (without memory)
  fully determined
  by $P_1(y,t)$ and $G(y_2 y_1 | t_2 t_1)$

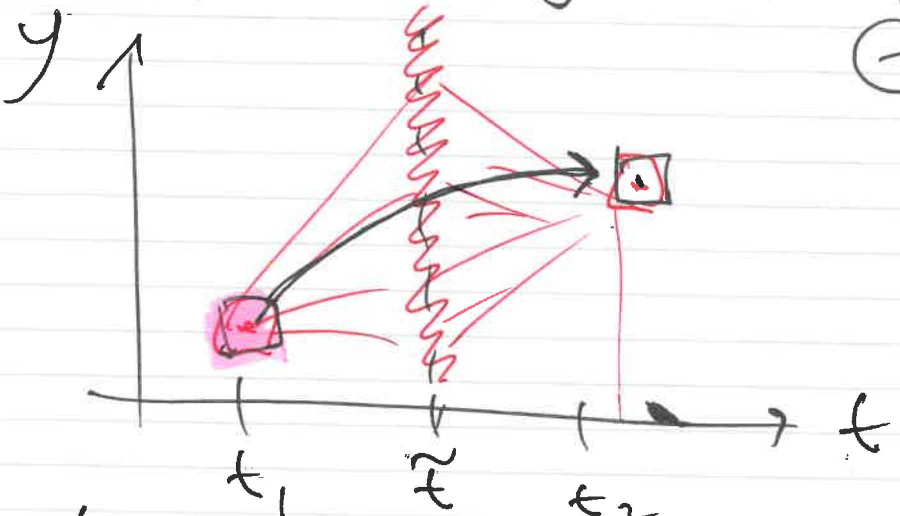
- Evolution relation

$$P_1(y_2, t_2) = \int G[y_2 y_1 | t_2 t_1] P_1(y_1 | t_1) \, dy_1$$



- Kolmogorov – Chapman relation

$$G[y_2 y_1 | t_2 t_1] = \int G[y_2 \tilde{y} | t_2 \tilde{t}] \cdot G[\tilde{y} \, y_1 | \tilde{t} \, t_1] \, d\tilde{y}$$



The difference with "Evolution" is that we retain (track) the initial condition $y_1(t_1)$ here

A Note about C.L.T.

We expand in exponent:

$$\phi_S(k) = \left[\phi_x(k/N)\right]^N$$

$$= \exp\left[1 + (ik)\langle x\rangle - \frac{k^2}{N} c_2(x) + \frac{ik^3}{N^2} c_3 + \dots \frac{k^n}{N^{n-1}} c_n + \dots\right]$$

$$e^{ik\langle x\rangle} \cdot \underbrace{e^{-\frac{k^2}{N}c_2}}_{|k| \sim \sqrt{N}} \cdot e \dots \underbrace{e^{-\frac{k^n}{N^{n-1}}c_n}}_{|k| \sim N^{1-1/N}}$$

almost ...

(wider)

$-\sqrt{N}$ $\qquad$ $+\sqrt{N}$

$e^{\frac{1}{N^*}} \approx 1$

Hence only the k^2 term matters

Gaussian $e^{-\frac{k^2}{2N}c_2}$ is the leading non-trivial term left.