Appendix: Elements of probability theory

see the book of Christophe and also math lecture

-> we have seen the necessity of principles to use a stat. description

- practical necessity.
- Absence of conserved qtity -> dynamics becomes chaotic, i.e. dependent on exact initial conditions that we do not know!
- Actually, we do not know the exact Hamiltonian either ~> isolated system do not really exist.
- Something I did not mention, is that we shall also use Q. Mechanics to describe our system -> inherent probabilist descript°

I) Random variables

1) Discrete one

We consider a variable X taking the value $X^{(m)}$ if the event e_m is realized. X is a random variable. We can define the probabilities $\{P_m\}$ such that $P(x) \equiv \text{Proba}(X = X^m) = \sum_{m=1}^{M} P_m \, \delta_{x, X^{(m)}}$

Here we assumed the space of events $\mathcal{E} = \{e_1, \ldots, e_M\}$ M possibles values

Toss a coin -> head or queue -> $M = 2$

consider N realizat° of the experiment. We obtain N_m the event e_m

$N_m \leq N$ is a random

$$P_m \stackrel{def}{=} \lim_{N \to \infty} \frac{N_m}{N} \Rightarrow \sum_{m=1}^{M} P_m = 1$$

Ex: Jackpot – $\boxed{\alpha}\boxed{\beta}\boxed{\gamma}$ with $\alpha, \beta, \gamma \in [0, 9]$

1000 events $P_{\alpha\beta\gamma} = \frac{1}{1000}$

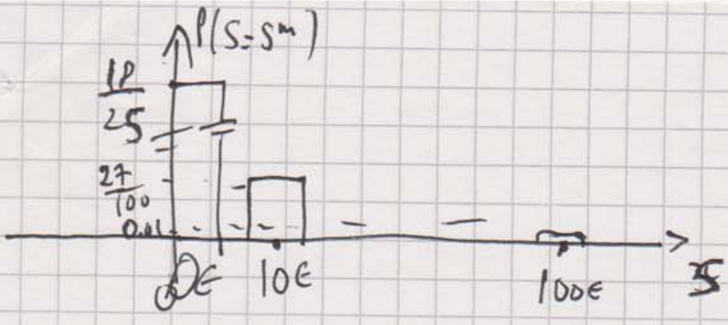
Assume If $\alpha = \beta = \gamma$ -> win $100€ = S_1$

$\alpha = \beta$ or $\alpha = \gamma$ or $\gamma = \beta$ — $10€ = S_2$

else $0 = S_3$

S = sum of gain => random variable -> 3 values

- $P(S_1) = \frac{10}{1000}$ for $S_1 = 100€$
- $P(S_2) = \frac{270}{1000}$ for $S_2 = 10€$
- $P(S_3) = \frac{720}{1000} = \frac{18}{25}$ for $S_3 = 0€$



2) Continuous variable

In the previous example $S \in \{S^1, S^2, S^3\} \rightarrow$ discrete set

In some case the random variable can be $\underline{C}$

ex : brownian motion of a microscopic object at the surface of a liquid $\rightarrow$ Jean Perrin, 1908

$\vec{r} = (x, y)$ where both x, y are random positi°

The extension to a $\underline{C}$ variable is as follows:

we introduce the probability density $w(x)dx$ that the random variable X takes the value x within dx :

$$\boxed{\begin{gathered} w(x)\,dx \equiv \text{Proba}\{X \in [x, x+dx]\} \\ \int w(x)\,dx = 1 \quad \text{normalisat°} \end{gathered}}$$

change of variable: If Y random variable such that $X = f(Y)$

$P(y)\,dy = ?$ $\qquad$ $\underline{f \text{ bijective}}$

$$P(y)dy = \text{Proba}\{X \in [f(y), f(y) + f'(y)\,dy]\}$$

$$= w(f(y))\,f'(y)\,dy$$

Thus $P(y) = w(f(y))\,f'(y)$

Let us check the normalisat° $\int P(y)\,dy = \int w(f(y))\,f'(y) = \int w(x)\,dx = 1$

Ex $\underset{\substack{\downarrow \\ X}}{E} = \frac{p^2}{2m}$, Y $\qquad$ suppose we know the distribut° of p say $\mathcal{P}(p)$

~~then $P(E)\,dE = \mathcal{P}(p)\,dp$~~

~~$P(E)\,\frac{p}{m}\,dp = \mathcal{P}(p)\,dp \Rightarrow P(E) =$~~

$$\mathcal{P}(p)\,dp = w(E)\,dE \Leftrightarrow \mathcal{P}(\sqrt{2mE})\sqrt{\frac{m}{2E}}\,dE = w(E)\,dE$$

$$\boxed{w(E) = \mathcal{P}(\sqrt{2mE})\sqrt{\frac{m}{2E}}}$$

If we have N random variables $w(x_1, \ldots, x_N)\,dx_1 \ldots dx_N = \text{Proba}\{X_1 \in [x_1, x_1+dx], \ldots X_N \in [x_N, x_N+dx_N]\}$

II) Average, variance, moments, cumulants & generating function

We consider the probabilistic average. (this assumes that the number of realizat° of an experiment, say N, is long $\Rightarrow N \to \infty$)

$$\langle X \rangle = \sum_{m=1}^{M} P_m X^{(m)} \quad \text{for discrete variables}$$

noted $E(X)$ in math

In our jackpot game: $M = 1000$, However S can take only 3 values $= V$

$$\langle x \rangle = \sum_{v=1}^{V} P_v x_v \quad \text{here the sum.}$$

$$\langle S \rangle = 100 \times \frac{1}{100} + 10 \cdot \frac{27}{100} + 0 \times \frac{18}{25} = 3.70\,€$$

The casino has to make sure that the cost of the try is > 3.70 to avoid bankrup.

Continuous variable $\boxed{\langle X \rangle = \int dx\, w(x)\, x}$

and $\langle f(X) \rangle = \int dx\, w(x)\, f(x)$

Variance $\quad \text{Var}(X) = \langle (X - \langle X \rangle)^2 \rangle$

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

standard deviation $\boxed{\sigma_X = \sqrt{\text{Var}(X)}}$

caracterizes the typical standard deviation from a random variable around its mean value

A4

If $w(x) = \frac{1}{a} e^{-x/a}$ for $x > 0$

$$\langle X \rangle = \int_0^{+\infty} dx \, \frac{x}{a} e^{-x/a} = a \int_0^{\infty} dy \, y \, e^{-y}$$

$$\langle x^2 \rangle = 2a^2 \qquad = a\left[-y e^{-y}\right]_0^{+\infty} + a \int_0^{+\infty} dy \, e^{-y} = a$$

Add $\quad \sigma_x = a \quad \rightarrow$ Correlat° (SIDE Page)

We can extend these definitions to higher moments

$$\mu_n \equiv \langle X^n \rangle = \int dx \, x^n \, w(x)$$

RQ: If we know all $\mu_n \; \forall n$, then $w(x)$ is known provided $w(x)$ is non pathological.

If $\sigma_x \ll |\langle x \rangle|$ then $\langle x^2 \rangle = \langle x \rangle^2 + \mathrm{var} X \approx \langle x \rangle^2$

Basically the second moment is given by the square of $\mu_1 \Rightarrow$ the variance brings some extra information compared to μ_1

We also call the variance $\left| \begin{array}{l} \mu_1 = c_1 \\ \sigma_x^2 \equiv c_2 \end{array} \right. \longrightarrow$ second cumulant

We can also extend the notion of cumulants to larger n with the idea that c_n brings new infos not contained in $\{c_{i \leq n}\}$

1) $c_1 \equiv \mu_1 = \langle x \rangle$

2) $c_2 \equiv \mu_2 - \mu_1^2 = \sigma_x^2 = \langle \tilde{x}^2 \rangle$

where $\tilde{x} = X - \langle x \rangle$

$\Rightarrow$ width of the distributi°

$\longrightarrow$ picture.

3) skewness $= \langle (x - \langle x \rangle)^3 \rangle$

$$c_3 = \langle \tilde{x}^3 \rangle = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3$$

$\Rightarrow$ asymmetry of the distributi° function

$c_3 > 0 \rightarrow$ fluctuat° are positive

4) Kurtosis ~~[illegible]~~ (A5)

For a Gaussian distribut° funct° $\langle \tilde{x}^4 \rangle = 3\langle \tilde{x}^2 \rangle^2$

$c_4 = \langle \tilde{x}^4 \rangle - 3\langle \tilde{x}^2 \rangle^2$ ⇒ deviat° from Gaussian.

show plots

$c_4 > 0$ ⇒ large fluctuat° are more favorable

Q: → Is there a systematic way to extend cumulants?

⇒ Generating functions

$$\boxed{g(p) \equiv \langle e^{-px} \rangle = \int dx\, e^{-px} w(x)} = \sum_n \frac{(-p)^n}{n!} \underbrace{\int dx\, x^n}_{\mu_n}$$

Laplace transform of the proba distribut°

$$g(p) = \sum_{n=0}^{\infty} \frac{(-p)^n}{n!} \mu_n \Rightarrow \mu_n = (-1)^n g^{(n)}(p)\Big|_{p=0}$$

$$\boxed{\mu_n = (-1)^n \frac{d^n g(p)}{dp^n}\Big|_{p=0}}$$

⇒ Easier to calculate $\langle e^{-px} \rangle$ than all μ_n!

Let us define the function ~~[illegible]~~ $\boxed{w(p) \equiv \ln g(p)}$

$$w(p) = \sum_{n=1}^{\infty} \frac{(-p)^n}{n!} c_n \Rightarrow \boxed{c_n = (-1)^n \frac{d^n w(p)}{dp^n}\Big|_{p=0}}$$

$$\ln g(p) = \ln\left(1 - \mu_1 p + \frac{1}{2}\mu_2 p^2 + O(p^3)\right)$$

$$= -\mu_1 p + \frac{1}{2}(\mu_2 - \mu_1^2)p^2 + O(p^3) \Rightarrow \begin{cases} c_1 = \mu_1 \\ c_2 = \mu_2 - \mu_1^2 \\ \text{etc} \ldots \end{cases}$$

III Some standard statistical laws

ex: binomial laws

N events → tail with proba p ⇒ what is the proba to get n tails
→ face (= head) —— $1-p$

$$(s = e^{-p}) \qquad \Pi_N(n) = C_N^n\, p^n (1-p)^{N-n}$$

$g_N(s) \equiv \langle s^n \rangle$ ⇒ we find $g_N(s) = [1 + p(s-1)]^N$

indeed $g_N(s) = \sum_n \Pi_N(n) s^n = \sum C_N^n (ps)^n (1-p)^{N-n} = (1-p+ps)^N$

Poissonian law $\Rightarrow$ Applies for independent events

Proba of an event to occur during time $dt = \lambda dt$

$\Rightarrow$ Applies to schottky exp $\to$ proba of extracting an el. under a biais

$$\langle I \rangle = \frac{1}{T}\int_0^T \lambda e \, dt = \lambda e$$

$P_{\lambda T}$ of having n events in a time T

$$\boxed{P_{v=\lambda T \atop =n} = \frac{\mu^n}{n!} e^{-\mu}} \quad \text{with } \mu = \lambda T \qquad \text{Poisson law}$$

$$g_\mu(p) = \exp\left(\mu(e^{-p}-1)\right) \quad \text{and} \quad w_v(p) = \mu(e^{-p}-1)$$

$$c_n = (-1)^n \frac{d^n}{dp^n} w_v \qquad = \mu \sum_{n=1} \frac{(-p)^n}{n!}$$

$$\Rightarrow \underline{c_n = \mu} \quad \forall n$$

$\Rightarrow$ all the cumulants are equal!

Ex: . Emmission of e^- in a tunnel junction $\Rightarrow$ $\begin{cases} \neq \text{bunching (photons)} \\ \neq \text{antibunching} \\ \text{corr } e^- \to \text{Pauli} \end{cases}$

. Radiation decay $\to$ SLIDES

Gaussian law

$$P_{a,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}} \to 2 \text{ parameters } \begin{cases} a \\ \sigma \end{cases}$$

$$g_{a,\sigma}(p) = \int dx \, P_{a,\sigma}(x) e^{-px}$$

$$= \exp\left(-ap + \frac{1}{2}\sigma^2 p^2\right)$$

$$\Rightarrow w_{a,\sigma}(p) = -ap + \frac{1}{2}\sigma^2 p^2 \longrightarrow \left| \begin{array}{l} c_1 = a = \langle x \rangle \\ c_2 = \mathrm{var}(x) = \sigma^2 \end{array} \right.$$

$$c_{n>2} = 0$$

$\Rightarrow$ There might exist pathological distributions

$w(x) \sim |x|^{-\mu-1}$ when $|x| \sim \infty \to \mu_n$ finite iff $n < \mu$

Ex : Cauchy law $w(x) = \dfrac{a/\pi}{a^2+x^2}$ (A7)

here $\mu = 1 \Rightarrow$ all moments are ∞

$$\tilde{g}(k) = \langle e^{-ikx} \rangle = e^{-a|k|} \Rightarrow \text{no Taylor expansion in } k=0$$

IV) Central limit theorem

We consider a set of N random variables $X_1, \ldots, X_N$

Assume.
i) They are statistically independent
ii) They all follow the same distribut° function $p(x)$
iii) We only assume $\langle X_n \rangle$ and $\langle X_n^2 \rangle$ are finite
i.e. the two first moments do exist!

Question : What is the distribution of $S = \sum_{i=1}^{N} X_i$?

RQ : • Up to a prefactor $\frac{1}{N}$ S coincides with the statistical average $\overline{x}^{(N)} = \frac{1}{N}\sum_{i=1}^{N} x_i \Rightarrow$ S therefore controls the statistical error made by a given sample $\Rightarrow$ vote

• Gaz $E_c \gg E_p \Rightarrow E \sim E_c = \sum_{i=1}^{N} \frac{p_i^2}{2m} \Rightarrow$ Assuming p_i R.V E_c ?

Look at the generating function:

$$G_N(p) \equiv \langle e^{-pS} \rangle = \langle \exp(-p\sum_{n=1}^{N} X_n) \rangle$$

$g(p) = \langle e^{-pX} \rangle$
$w(p) = \ln g(p)$

$$= \langle \prod_{n=1}^{N} (e^{-pX_n}) \rangle \quad \text{statistical independence}$$

$$\text{hyp 1} \quad = \prod_{n=1}^{N} \langle e^{-pX_n} \rangle = \underbrace{|g(p)|^N}_{\text{hyp 2}}$$

$$\Rightarrow \underline{G_N(p) = |g(p)|^N}$$

The cumulant generating function $\Omega_N(p) \equiv \ln G_N(p) = N \ln|g(p)|$

$$= \sum \frac{(-p)^k}{k!} \underset{\text{cumulants of } S}{C_k} = \underline{N\, w(p)}$$

If we introduce $\omega(p) = \sum \frac{(-p)^k}{k!} c_k \Rightarrow \boxed{\mathcal{C}_k = N c_k}$ $\mathcal{C}_k^{1/k} = N^{1/k} c_k^{1/k} \sim N^{1/k}$

$$\left| \begin{array}{ll} \mathcal{C}_1 = N \langle x \rangle & \to \text{trivial} \\ \mathcal{C}_2 = N\, var(X) & \to \text{not so obvious} \end{array} \right) \text{hyp 3}$$

$\to$ This implies $\boxed{\sigma_S = \sqrt{N}\,\sigma_X}$ $\to$ typical mean deviation of S behaves as $\sqrt{N}$ i.e. increases with N!

Let us show that the higher cumulants are negligible

from W_S $\hookleftarrow$ • All cumulants increase linearly with $N \Rightarrow \mathcal{C}_k \propto N$

• We have seen above that the 2^{nd} cumulant $\mathcal{C}_2$ contributes to the fluctuat° of S as $\sqrt{\mathcal{C}_2} \to$

• By dimensional analysis, the ($k>2$) cumulant contributes to the fluctuation of S as $(\mathcal{C}_k)^{1/k} \propto N^{1/k}$

Therefore $\left(\frac{\mathcal{C}_k^{1/k}}{\mathcal{C}_2^{1/2}} \right) \propto N^{-\frac{(k-2)}{2k}} \xrightarrow[N\to\infty]{} 0$

$\propto N^{\frac{1}{k}-\frac{1}{2}} = N^{-\frac{(k-2)}{2k}}$

Therefore, in the large N limit

$$G_N(p) \simeq e^{-Nc_1 p + \frac{1}{2} N c_2 p^2}$$

i.e.

$$\boxed{P_N(S) \underset{N\gg 1}{\simeq} \frac{1}{\sqrt{2\pi N\, var X}} \exp\left\{ -\frac{(S - N\langle x \rangle)^2}{2N\, var X} \right\}}$$

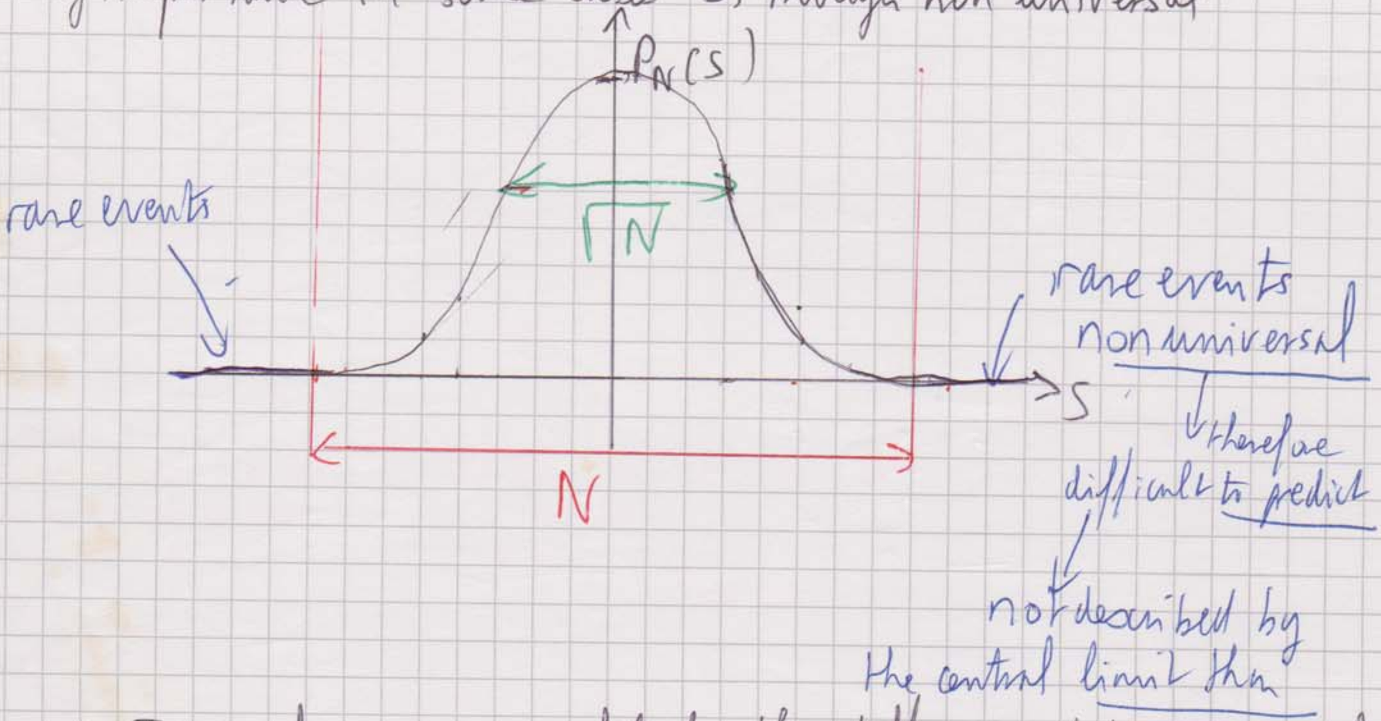
In the large N limit, S tends towards a Gaussian distribut° function of mean $\langle S \rangle = N \langle x \rangle$ and standard deviat° $\sigma_S = \sqrt{N}\,\sigma_X$

$\Rightarrow$ This is remarquable because we only make very few hypothes independent of $p(x)$!

$\Rightarrow$ Emergence of a new universal law in the large N independent of microscopic details

This thm is fundamental and actually perfectly illustrates that the laws governing the macroscopic (thermodynamic) may be completely $\neq$ from the microscopic $\rightarrow$ Depends on only few parameters (two here)

RQ: σ_s characterizes the standard deviations. This does not preclude the existence of rare events (or large deviations) which are also of importance in some cases $\Rightarrow$ though non universal



$\Rightarrow$ To analyze in more details the differences between universal fluctuations and rare events see Annex of Christophe book