Foundations

www.tcd.ie/physics/quantumtech

Why probability theory?

* probability/statistics underpins all science
* quantum mechanics is inherently probabilistic
* fascinating mathematics & philosophy

[Lectures based on textbooks by van Kampen & Jaynes]

What do probabilities mean?

* frequentist – P(heads) is the fraction of observed heads in many identical trials
* Bayesian – P(heads) quantifies the degree of belief we have in observing heads in each trial

Plausibility vs. deduction

Propositions e.g. A = it is raining

B = it is cloudy

C = I am hungry ....

Logical product: AB = A AND B are true

Logical sum: $A+B$ = A OR B is true

[e.g. think about binary variables]

Negation: $\bar{A}$ = A is NOT true

Implication: $A \Rightarrow B$ (A implies B)

What about the converse: $B \Rightarrow A$?

B does not imply A but it does make it more <u>plausible</u>.

We want to quantify plausibility, <u>conditional</u> on some info

$A|B$ = A given/conditioned on B

$A|BC$ etc.

<u>Cox's theorem</u> (see Jaynes' textbook)

1. Represent plausibility by real numbers (increasing, continuous)

$P(A|C) > P(B|C) \Rightarrow$ A more plausible than B given C

2. Consistency with common sense, e.g.

$$\begin{aligned} P(A|C') &> P(A|C) \\ P(B|AC') &= P(B|AC) \end{aligned} \quad \Rightarrow \quad \begin{aligned} P(\bar{A}|C') &< P(\bar{A}|C) \\ P(AB|C') &\geq P(AB|C) \end{aligned}$$

3. Mathematical self-consistency

From (1-3) we can obtain the following rules

* $P(AB|C) = P(A|BC)\,P(B|C) = P(B|AC)\,P(A|C)$
* $P(A|B) + P(\bar{A}|B) = 1$
* If $C = B \Rightarrow A$ then $P(A|BC) = 1$
* $P(A+B|C) = P(A|C) + P(B|C) - P(AB|C)$
* If $\{A_1, \ldots, A_N\}$ are mutually exclusive & exhaustive, i.e. $P(A_iA_j|B) = P(A_i|B)\,\delta_{ij}$ & $\sum_i P(A_i|B) = 1$, and if B does not favour one A_i over any other, then $P(A_i|B) = \frac{1}{N}$ (principle of indifference)

Simple example

B = an urn contains 2 red (R) balls & 8 green (G) ones

What is the prob. that we draw a red one?

background indifferent

Label the balls: 1,2 are red, 3–10 are green

A_i = the ith ball is drawn

$$P(A_i|B) = 1/10$$

$$P(R|B) = P(A_1 + A_2 | B)$$

$$= P(A_1|B) + P(A_2|B) - \cancelto{0}{P(A_1 A_2|B)}$$

$$= 2/10$$

more generally, if B = M red balls, $N-M$ green

$$P(R|B) = P(A_1 + A_2 + \dots + A_M | B) = \sum_{i=1}^{M} P(A_i|B) = \frac{M}{N}$$

Interesting example (Litt 2024)

Suppose an urn has M red balls, $N-M$ green, $N = 100$ balls total, M is chosen by picking a number from $0 \to 100$ out of a hat. You choose a ball and it's red (R_1). Is the next ball more likely to be red (R_2) or green?

$$P(R_2|R_1) = \sum_{m=0}^{N} P(R_2, M=m | R_1)$$

$$= \sum_{m=0}^{N} P(R_2 | M=m, R_1)\, P(M=m | R_1)$$

$P(M=m|R_1) = \ ?$

$P(M=m, R_1) = P(R_1|M=m)\,P(M=m) = P(M=m|R_1)\,P(R_1)$

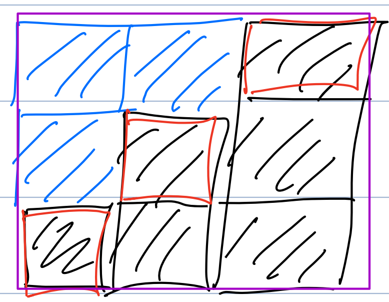
$\Rightarrow P(M=m|R_1) = \dfrac{P(R_1|M=m)\,P(M=m)}{P(R_1)}$ Bayes' rule

$P(R_1) = \sum_{m=0}^{N} P(R_1|M=m)\,P(M=m)$

$P(R_1|M=m) = \dfrac{m}{N}$, $P(M=m) = \dfrac{1}{101} = \dfrac{1}{N+1}$

$\Rightarrow P(R_1) = \dfrac{1}{N(N+1)} \sum_{m=0}^{N} m$

$= \dfrac{1}{2}$



[makes sense if you think about it! principle of indifference]

$\sum_{m=0}^{N} m = N^2 - \frac{1}{2}[N^2 - N]$
$= \frac{1}{2}(N^2 + N) = \frac{1}{2}N(N+1)$

$P(M=m|R_1) = \dfrac{P(R_1|M=m)\,P(M=m)}{P(R_1)}$

$= \dfrac{m}{\frac{1}{2}N(N+1)}$ (length-biased sampling)

$P(R_2|M=m, R_1) = \begin{cases} \dfrac{m-1}{N-1} & m \geq 1 \\ 0 & m = 0 \end{cases}$

$\Rightarrow P(R_2|R_1) = \sum_{m=1}^{N} \dfrac{m-1}{N-1} \cdot \dfrac{m}{\frac{1}{2}N(N+1)} = \dfrac{2}{3}$

$\sum_{m=0}^{N} m(m-1) = \frac{1}{3}(N-1)N(N+1)$

Applications

Probabilities & random variables

Random variable X specified by:

i) range of values, e.g. discrete $X = \{x_1, x_2, \dots, x_N\}$

or continuous $X \in [a, b]$

ii) probability distribution

$$P(X = x) = p_x$$

$$P(X \in [x, x+dx]) = p(x)\,dx$$

normalised $\quad \sum_x p_x = 1 \quad$ or $\quad \int dx\, p(x) = 1$

positive $\quad P(X) \geq 0$

Expectations:

$$E[X] = \sum P(X=x)\, x \quad \text{the "centre of mass"}$$

$$E[f(X)] = \sum P(X=x)\, f(x)$$

e.g. moments $\quad E[X^m]$

variance $\quad Var[X] = E[(X - E[X])^2]$

$$= E[X^2] - E[X]^2 \quad \text{(exercise)}$$

Characteristic function: (very useful for proving various results)

$$G_X(u) = E[e^{iuX}] = \sum_x P(X=x)\, e^{iux}$$

$$G_X(0) = \sum_x P(X=x) = 1, \quad P(X=x) = \int \frac{du}{2\pi}\, e^{-iux}\, G_X(u)$$

$$\frac{d}{du} G_X(u)\Big|_{u=0} = \sum_x P(X=x)\, ix = i\,E[X]$$

$$(-i)^m \frac{d^m}{du^m} G_X(u)\Big|_{u=0} = E[X^m]$$

aka moment generating function $G_X(u) = \sum_{m=0}^{\infty} \frac{(iu)^m}{m!} E[X^m]$

Cumulant-generating function (CGF) $K_X(u) = \ln G_X(u)$

$$(-i)^m \frac{d^m}{du^m} K_X(u)\Big|_{u=0} = m^{th} \text{ cumulant } \kappa_m$$

$\kappa_1 = E[X]$, $\kappa_2 = Var[X]$, ... (exercise)

or $K_X(u) = \sum_{m=1}^{\infty} \frac{(iu)^m}{m!} \kappa_m$

<u>Multi-variate distributions</u>

Two random variables X_1, X_2

$$P(X_1=x_1, X_2=x_2) = P(X_1=x_1 \mid X_2=x_2)\, P(X_2=x_2)$$

<u>Independence</u> means that

$$P(X_1=x_1 \mid X_2=x_2) = P(X_1=x_1) \quad (\forall x_2)$$

$$\Rightarrow \quad P(X_1=x_1, X_2=x_2) = P(X_1=x_1)\, P(X_2=x_2)$$

Correlation / covariance :

$$Cov[X_1, X_2] = E[X_1 X_2] - E[X_1]E[X_2]$$

if independent : $E[X_1 X_2] = \sum_{x_1, x_2} P(X_1=x_1) P(X_2=x_2)\, x_1 x_2$

$$= E[X_1]\, E[X_2]$$

$$\Rightarrow \quad Cov[X_1, X_2] = 0.$$

BUT $Cov[X_1, X_2] = 0$ does not imply independence.

Sums of indpt. random variables

$$S = \sum_{n=1}^{N} X_n$$

$$G_S(u) = E\left[e^{iu\sum_{n=1}^{N} X_n}\right] = E\left[\prod_{n=1}^{N} e^{iuX_n}\right] = \prod_{n=1}^{N} E\left[e^{iuX_n}\right]$$

$$= \prod_n G_{X_n}(u)$$

$$K_S(u) = \sum_n \ln G_{X_n}(u) = \sum_n K_{X_n}(u)$$

$$\Rightarrow E[S] = \sum_n (-i) K'_{X_n}(0) = \sum_n E[X_n]$$

$$Var[S] = \sum_n (-i)^2 K''_{X_n}(0) = \sum_n Var[X_n]$$

$$\vdots$$

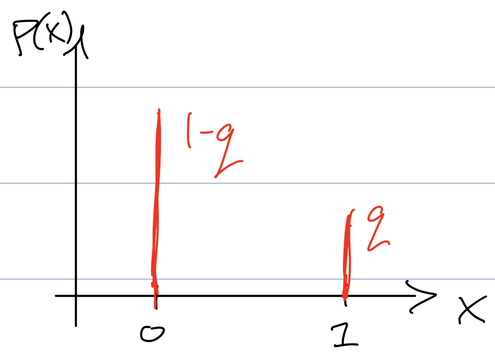
[Means & variances of <u>independent</u> random variables are additive!]

If $\{X_n\}$ are i.i.d. then $E[S] = N E[X]$, $Var[S] = N Var(X)$ etc.

Bernoulli & binomial distributions

Bernoulli distⁿ:

$X = \{0, 1\}$, $P(X=1) = q$ [e.g. biased coin flip, $q = 2/3$, $1-q = 1/3$]



$E[X] = q$, $E[X^m] = q$

$Var[X] = q(1-q)$

$G_X(u) = 1 + q(e^{iu} - 1)$

Binomial distⁿ: repeat N times & sum: $S = \sum_{n=1}^{N} X_n$

e.g. $S = 2$, $N = 5$,

01001
00110
10100
$\vdots$

$\binom{5}{2}$ outcomes each with prob. $q^2(1-q)^3$

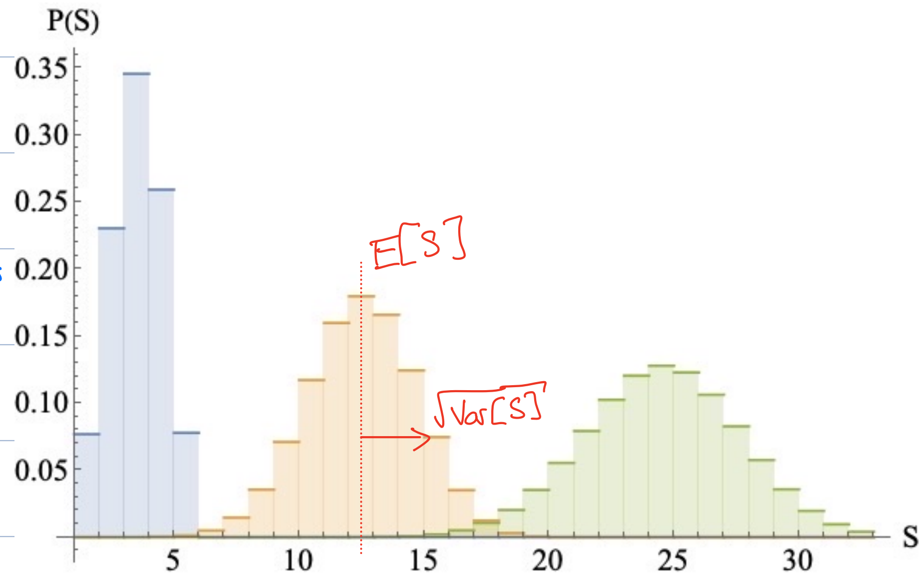
$$P(S=s) = \binom{N}{s} q^{s}(1-q)^{N-s}$$

$$E[S] = Nq$$

$$Var[S] = Nq(1-q)$$

this

then this



$$\frac{\sqrt{Var[S]}}{E[S]} = \sqrt{\frac{1-q}{Nq}}$$

Central limit theorem

$$E[X] = \mu \quad , \quad Var[X] = \sigma^2$$

$$\tilde{S} = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} (X_n - \mu)$$

$$G_{\tilde{S}}(u) = E\left[e^{iu\tilde{S}}\right] = \left[e^{iu(X-\mu)/\sqrt{N}}\right]^N$$

$$= \left[e^{-iu\mu/\sqrt{N}}\, G_X(u/\sqrt{N})\right]^N$$

$$G_X(u) = e^{K_X(u)} = e^{\underbrace{iu\mu}_{\kappa_1} - \frac{1}{2}u^2\underbrace{\sigma^2}_{\kappa_2} + \dots}$$

$$\Rightarrow G_{\tilde{S}}(u) = \left[e^{-iu\mu/\sqrt{N}}\, e^{iu\mu/\sqrt{N} - \frac{1}{2}u^2\sigma^2/N + O(1/N^{3/2})}\right]^N$$

$$\simeq \left[1 - \frac{\frac{1}{2}u^2\sigma^2}{N}\right]^N \qquad (N \gg 1)$$

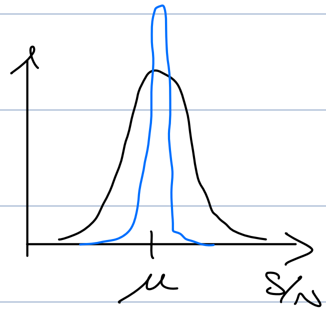
$$\xrightarrow[N\to\infty]{} e^{-\frac{1}{2}u^2\sigma^2}$$

$$\Rightarrow P(\tilde{S} = \tilde{s}) = \int \frac{du}{2\pi}\, e^{-\frac{1}{2}u^2\sigma^2}\, e^{-iu\tilde{s}} = \frac{e^{-\tilde{s}^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}$$

Gaussian distⁿ

Consider sample mean $\frac{S}{N} = \frac{1}{N}\sum_{n=1}^{N} X_n$:

$$\Rightarrow P(S/N = x) \propto \exp\left[-\frac{N}{2\sigma^2}(x-\mu)^2\right]$$



Holds for any sum of i.i.d. random variables so long as cumulants $\kappa_m < \infty$.

As $N \to \infty$, the sample mean $S/N = \frac{1}{N}\sum_{n=1}^{N} X_n$ converges to the expectation value $\frac{S}{N} \to E[X]$ with probability $\to 1$. [frequentist interp$^{\underline{n}}$]

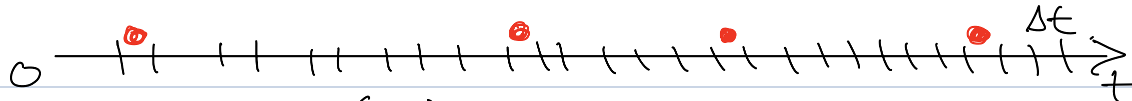
Fluctuations of the sample mean are $O(1/\sqrt{N})$

i.e. $$E[S/N] = \frac{1}{N}\sum_{n=1}^{N} E[X_n] = E[X]$$

$$\mathrm{Var}[S/N] = E\left[\left(\frac{S}{N} - E[X]\right)^2\right] = \frac{\mathrm{Var}[X]}{N}$$

Poisson distribution

Consider binomial distribution where # trials $N \to \infty$ and "success prob" $q \to 0$, s.t. $Nq = \alpha = \text{const}$.



eg. photodetection probability q in each time step Δt

$$P(S=s) = \binom{N}{s} q^s (1-q)^{N-s}$$

$$= \frac{N!}{s!\,(N-s)!}\left(\frac{\alpha}{N}\right)^s \left(1 - \alpha/N\right)^{N-s}$$

Stirling's approx: $\ln n! \simeq n \ln n$ for $n \gg 1$

$$\ln P(S=s) \simeq N\ln N - \ln s! - (N-s)\ln(N-s)$$
$$+ s\ln\alpha - s\ln N + (N-s)\ln\left(1-\frac{\alpha}{N}\right)$$
$$\simeq N\ln N - \ln s! - (N-s)\ln N + s\ln\alpha$$
$$- s\ln N + (N-s)\left(-\alpha/N\right)$$
$$\simeq s\ln\alpha - \ln s! - \alpha$$

$$\Rightarrow P(S=s) = \frac{\alpha^s}{s!}e^{-\alpha}$$ Poisson distribution

$$G_S(u) = E[e^{iuS}] = \sum_{s=0}^{\infty} e^{ius}\frac{\alpha^s}{s!}e^{-\alpha}$$
$$= e^{-\alpha}\sum_{s=0}^{\infty}\frac{(\alpha e^{iu})^s}{s!}$$
$$= e^{-\alpha}e^{\alpha e^{iu}}$$
$$= \exp[\alpha(e^{iu}-1)]$$

or $K_S(u) = \alpha(e^{iu}-1)$

$\Rightarrow E[S] = -iK_S'(0) = \alpha$

$\mathrm{Var}[S] = -K_S''(0) = \alpha = E[S]$

all cumulants $\kappa_m = E[S] = \alpha$!

Poisson statistics, e.g. photons from a laser.