# Last week's quiz

**1** Suppose you have a box of radioactive material. At 9:00am, you have 2500 atoms. You come back after an hour and count 2460 atoms.

(a) What is the probability an atom decays over 1 hr?

$$N(t+\delta t) = 2460 = N(t)(1-\delta) = 2500\cdot(1-\delta)$$

$$\Longrightarrow \delta = 0.016$$

(b) Write down the difference equation describing this system

$$N(t+\delta t) = N(t)(1-0.016) = 0.984 \times N(t)$$

(c) Write down the ODE that models this system (approximately)

$$\frac{dN}{dt} = -\underbrace{\left[\frac{\delta}{dt}\right]}_{\text{decay rate}} N(t) = -0.016\, N(t)$$

time has units?

$\star$ /hr

decay probability

has units!

**NOT** $N(t) = N_0 \exp(-\tilde{\delta} t) \Leftarrow$ not a differential equation!

# Models & Data

1) What you want to do / know does not always align with the tools at your disposal

--> models are built on assumptions (when are they true?)

--> data is built on what you can measure (and it has error!)

2) When given data, one goal is to understand: **do the assumptions of my model seem to hold here?**

3) You answer this question by visualizing your data

--> make a chart or a plot of your data

--> Note units, axes limits, context

4) For the exponential decay problem, what do you notice about this data?

--> It does not decay to zero

5) Does the observation it does not decay to zero match our assumption that all atoms decay?

--> **No!**

6) What could be the mismatch?

--> Maybe some atoms are not decaying

--> Let's estimate how many are not decaying (this is not always easy! But we can do it here)

7) Ok, subtracting off the 'inert' atoms, how can we estimate decay rate?

--> There are many options!

Difference Eqn

$$N(t+\delta t) = N(t)(\underbrace{1-\delta}_{\text{probability atom survives}})$$

Fraction that survives:

$$\eta(t) = \frac{N(t+\delta t)}{N(t)}$$

ideally, $\eta(t) = 1-\delta$

Q When is $\eta(t)$ a good measure of $1-\delta$?

$\Rightarrow$ when $N(t)$ & $N(t+\delta t)$ are sufficiently large

→ $\eta(t)$ toward end of data is not reliable!

$\boxed{Q}$ How can we estimate "true" $\eta$?

Option (A)

$$\bar{\eta} = \frac{\eta(0)+\eta(1)+\eta(2)+\dots+\eta(k)}{k}$$

$\boxed{Q}$ how big can we take $k$?

Option (B)

"clump" first $k$ days together:

$$N(t+k\delta t) = N(t)\,(1-\delta)^k$$

$$\rightsquigarrow \eta \approx \left(\frac{N(t+k\delta t)}{N(t)}\right)^{1/k}$$

$\boxed{Q}$ How do you know your $\eta$ estimate is "good"?

# Maximum likelihood estimate

probability!

We want an $\eta$ that maximizes the likelihood that, if you used it in the model, you would get the data you observe.

(let $dt = 1$ time unit)

joint probability function

$$p(\text{data} \mid \eta) = \text{“}L(\eta)\text{”}$$

$$= \prod_{t=0}^{T-1} \underbrace{\binom{N(t)}{N(t+1)}}_{\text{choice of atoms that survive}} \left( \underbrace{\eta^{N(t+1)}}_{\text{probability they survive}} \cdot \underbrace{(1-\eta)^{N(t)-N(t+1)}}_{\text{probability others decay}} \right)$$

Note: $\eta = 1 - \delta$ ← decay probability

We want

$$\frac{d}{d\eta}\left(L(\eta)\right) = 0$$

identically, $\frac{d}{d\delta}\left(L(1-\delta)\right) = 0$

First, <u>much</u> easier to work with $\log(p)$:

$$\log(p) = \sum_{t=0}^{T-1} \log\binom{N(t)}{N(t+1)} + \log\left((1-\delta)^{N(t+1)}\right) + \log\left(\delta^{N(t)-N(t+1)}\right)$$

$$= \sum \text{“} \qquad \text{”} + N(t+1)\log(1-\delta) + (N(t)-N(t+1))\log(\delta)$$

$$\frac{d}{d\delta}\log(p) = \frac{1}{p}\frac{dp}{d\delta} \longleftarrow \text{want } = 0$$

$$\left(\text{and also } \frac{d^2p}{d\delta^2} < 0\right)$$

$$0 = \sum N(t+1)\left(\frac{-1}{1-\delta}\right) + \sum \frac{(N(t)-N(t+1))}{\delta}$$

$$\Rightarrow \frac{\sum N(t+1)}{1-\delta} = \frac{\sum N(t) - N(t+1)}{\delta}$$

$$\Rightarrow \left[\sum N(t+1)\right]\delta = \left[\sum N(t) - N(t+1)\right](1-\delta)$$

$$\Rightarrow +\left[\sum N(t)\right]\delta = \sum N(t) - N(t+1)$$

$$\Rightarrow \quad \delta = \frac{\sum_{t=0}^{T-1} N(t) - N(t+1)}{\sum_{t=0}^{T-1} N(t)}$$

$$\Rightarrow \quad \boxed{\delta^* = \frac{N(0) - N(T)}{\sum_{t=0}^{T-1} N(t)}}$$

maximum likelihood estimate (MLE)

Notice: as $dt \to 0$,

$$\delta^* \to \frac{\text{total difference}}{\text{total integral}}$$

Unlike options A & B, MLE is robust to the overall length of the experiment

⇒ multiple zero estimates at the end does not ruin this!

⇒ you can also extend this to account for multiple short independent experiments

⇒ this strategy can be also extended to incorporate more sophisticated models

e.g.

$$p(\text{data} \mid \eta, M)$$

# inert atoms, unknown?

then need

$$\partial_\eta p = 0 \quad \underline{\text{and}} \quad \partial_M p = 0$$

i.e. $\nabla p = 0$

# Discussing HW

plot the data

$$jA + kB \underset{\beta}{\overset{\alpha}{\rightleftharpoons}} \ell C$$

$$\frac{d[A]}{dt} = -j\alpha[A]^j[B]^k + j\beta[C]^\ell$$

$$\frac{d[C]}{dt} = +\ell\alpha[A]^j[B]^k - \ell\beta[C]^\ell$$

☆ Consider the case where you have mostly the backward reaction

$[A], [B] << 1$

$$\Rightarrow -\frac{d[C]}{dt} \approx \ell\beta[C]^\ell$$

$$\downarrow$$

$$\log\left(-\frac{d[C]}{dt}\right) = \underbrace{\log(\ell\beta)}_{\text{unknown}} + \underbrace{\ell}_{\text{unknown}}\log([C])$$

$$y = b + mx \; !$$

$\Rightarrow$ $l$ is the slope of the line

$$\{ \log([c]), \log(-\frac{d[c]}{dt}) \}$$

(approxmately)

$[c] << 1$

☆ Consider when it is mostly forward

$$\hookrightarrow -\frac{d[A]}{dt} \approx \alpha_j [A]^j [B]^k \quad \text{(time } t\text{)}$$

$$\underbrace{\log\left(-\frac{d[A]}{dt}\right)}_{\substack{\text{known} \\ \text{(approxmately)}}} = \underbrace{\log(\alpha_j)}_{\substack{\text{unknown} \\ \text{"}c_1\text{"}}} + \underbrace{j}\,\underbrace{\log([A])}_{\text{known}} + \underbrace{k}\,\underbrace{\log([B])}_{\text{known}}$$

$\leadsto$ you should be able to estimate $c_1$, $j$, $k$ from this

# Systems of ODEs

$$\begin{cases} \dfrac{dX}{dt} = f(X,Y) \\[2mm] \dfrac{dY}{dt} = g(X,Y) \end{cases}$$

Define: steady state // fixed points

$\Rightarrow$ $X(t)$ and $Y(t)$ do not change in time

$\Rightarrow$ i.e. $\dfrac{dX}{dt} = 0$ and $\dfrac{dY}{dt} = 0$

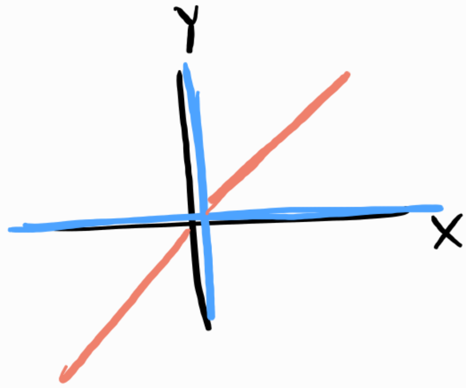
$\rightsquigarrow$ $\{X_1, Y_1\}$ s.t. $f(X_1, Y_1) = 0$

$\rightsquigarrow$ $\{X_2, Y_2\}$ s.t. $g(X_2, Y_2) = 0$

X nullclines    Y nullclines

e.g. if $f(X,Y) = XY$ ★

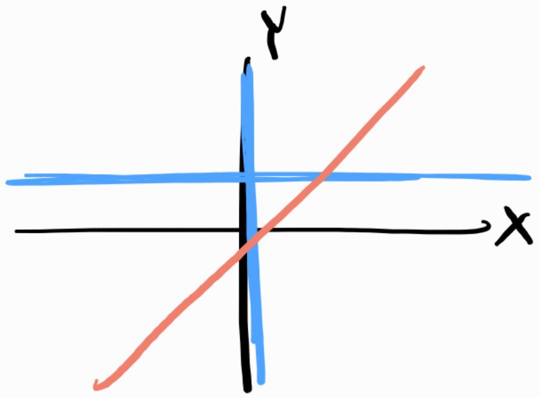
$g(X,Y) = X - Y$ ★



$\boxed{Q}$ What are the fixed points here?

Just one: $(x^*, y^*) = (0,0)$

$\boxed{Q}$ What if $f(x,y) = x(y-1)$?




Two fixed points:

$(x^*, y^*) = (0,0)$

$(x^*, y^*) = (1,1)$

## Approximate solution to this ODE

We will use our intuition to come up with a numerical scheme which is just the Euler method

$$\frac{dx}{dt} = f(x,t)$$

$$\| $$

$$\lim_{dt \to 0} \frac{x(t+dt) - x(t)}{dt} \overset{\text{if } dt \text{ is small}}{\approx} \frac{x(t+dt) - x(t)}{dt} \quad (+\text{error})$$

$$\Rightarrow \underline{\underline{x(t+dt) = x(t) + dt \cdot f(x,t)}}$$

Be smart about choosing:

→ initial condition $x(0)$

→ timestep size $dt$

(more accurate if $dt$ is <u>small</u>)

# Quiz

1. List two assumptions inherent to the rate equation "model" for chemical kinetics

2. Suppose $A + 2B \xrightarrow{\alpha} 3C$

and you have data

| | 0 min | 3 min | 7 min |
|---|---|---|---|
| [A] | 1.2 | 1.12 | 1.05 |
| [B] | 1.3 | 1.14 | 1.00 |
| [C] | 0.1 | 0.33 | 0.55 |

(a) What, approximately, might be the rate constant $\alpha$? Show how you find it.

(b) In your opinion, is your answer in (a) a "good" estimate of the <u>true</u> $\alpha$? Why or why not?
(Hint: is your data "good"?)