Homework solutions
The goal of physical chemistry - Chapter 1

Homework Problem O1.1: For the function $y=t$, calculate the average $\bar{y}$, if $P(t)=k\cdot e^{-kt}$ per unit time over the range $t=0$ to ∞. Since t is time, this gives the average time for a molecule with rate coefficient k to decay. How does that compare with the half-life of the molecule? [Hint: $y=t$ is a continuous function, so you need to use the integral formula, not the summation formula.]

Solution: Use the formul**a** $\bar{y} = \int dx y(x) P(x)$, where our independent variable 'x' is now called 't' and $y(x)$ is thus $y(t)$:

$$\bar{y} = \int_0^\infty dt\ t\ ke^{-kt}.$$

You can do this by integration by parts, or just go to an online integrator such as Wolfram Alpha: https://www.wolframalpha.com/calculators/integral-calculator/ Try it out right now by typing the formula x * k * exp(-k*x) into the integral and hitting equal on the right. You get

$$-\frac{e^{-k\,x}\,(1 + k\,x)}{k}$$

Evaluating at ∞, where $e^{-kt}(1+kt)$ goes to 0, and evaluating at $t=0$, where we get -(1/k), the result is

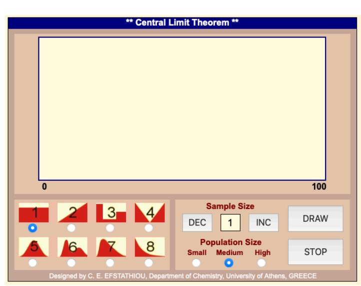
$$\bar{y} = 0 - \left(-\frac{1}{k}\right) = \frac{1}{k} = \bar{t}.$$

Thus the average time for the reaction is the inverse of the rate coefficient k, and since the half-life $t_{1/2} = \ln(2)/k$, the average time of reaction $\bar{t} = t_{1/2}/\ln(2)$.

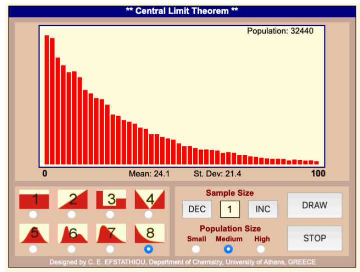
Why is $P(t)=k\cdot e^{-kt}$ multiplied by the k out front, and not just $P(t)=e^{-kt}$? The reason is that the probability must be normalized to 1, i.e. the chance the reaction occurs at some time between $t=0$ and ∞ must be 100% (or 1). But $\int_0^\infty e^{-kt} dt = 1/k$, not 1, so we multiply by k to normalize the probability.

Homework Problem O1.2: Go to the applet
http://195.134.76.37/applets/AppletCentralLimit/Appl_CentralLimit2.html ,
which shows you how starting with 1 random variable X_1 with distribution $P(X_1)$, then adding a second, and so forth, eventually produces a Gaussian. Click on "8" = exponential decay, Population size 'Medium', and draw for sample size 1. Then INCrease sample size to 2 (two random variables, both exponentially distributed added up), 5 and 10. Does it begin to look like a Gaussian for 10 random exponential variables added up?"

Solution: Going to the website, you'll see a window



Let's show the first example here, a single random variable, but don't forget to turn in the results for the sum of 2, 5 and 10 random variables as well for full credit! Click on "8" to select an exponential decay $P(t)=k \cdot e^{-kt}$ and then click "DRAW". You'll get, with an x-axis going from 0 to 100



,

the expected exponential distribution. But if you set "Sample Size" to 2, 5 or 10, the distribution will lose probability at 0 and start looking like a Gaussian. The reason is that while getting "0" is the most likely value to get (highest probability $P(t)$), getting "0" twice in a row is still not as likely as getting, say 10=0+10 or 1+9 or 2+8 or lots of other possibilities. By the time you add up 10 random variables, the chance of getting "0" ten times in a row is negligible, and you get a near-Gaussian centered at the average, about 24 in the online app.

Homework problem O1.3**:** Consider the molecular DNA sequence CCAGCAC_GGCAGCAT, with a base missing. What would be a good guess for that base? To start, what is the probability P(A) that adenine occurs in that known sequence? The probability P(C)? The probability that whenever there is an A, there's a C before it? (that is P(C, given A is in the next position). Therefore what is P(A|C) = P(A, given C is in the previous position)? So what odds do you give an A being in the blank space?

Solution: The known sequence has 15 entries and there are 4 adenines, so P(A)=4/15. Similarly, P(C)=6/15. The probability there is a C at position N, given there is an A at position N+1 is P(C|A)=1 (100%). Thus

$$P(A|C)=P(C|A)\ P(A)/P(C) = 1*(4/15)/(6/15) = 2/3.$$

For guanine = G, the answer would have been

$$P(G|C)=P(C|G)\ P(G)/P(C) = 0*(4/15)/(6/15) = 0$$

since there is no prior occurrence of a G after a C in the sequence. So A would be a good guess, but Bayes is not giving it a '1': there are extra Cs in the sequence, such as the very first one, that do not have an A after them. If the actual answer came out differently, you can recalculate the probabilities, and the one for A will come out smaller.

Note that such techniques are used to reconstruct phylogenetic trees (ancient DNA sequences). So take the answers with a grain of salt: they are probabilistically reasonable, but by no means guaranteed right.

Homework problem O1.4: Write the number i in exponential notation (what is its angle φ?). Now take the natural logarithm ln(i). What do you get? Can you devise a simple procedure that extracts the magnitude 'r' and angle 'φ' of any complex number by distributing them into a real and imaginary piece?

Solution: Since $x=r[\cos(\varphi)+i\sin(\varphi)]$, we must have r=1 because the magnitude of $|i|=1$. $\cos(\varphi)=0$ must also be true, since i has no real part. Thus $\varphi = 90°$, or $\pi/2$ in radians units. We therefore have

$$i=e^{i\pi/2}$$

in exponential notation. Taking the natural log, we have

$$\ln(i)=\ln(e^{i\pi/2})=(i\pi/2)\ln(e)= i\pi/2$$

So, the log can extract the magnitude and phase angle of a number as follows: if $x=re^{i\varphi}$, then $\ln(x) = \ln(r)+i\varphi$. The real part will be the log of the magnitude, and the imaginary part will be the angle. This little trick, like the idea that multiplying to complex numbers means you multiply their r's and add their angles to get the new number, make it useful to sometimes thing of complex numbers in polar coordinates, sometimes in terms of Cartesian coordinates. They are related as usual $x = r\cos(\varphi)$ and $y=r\sin(\varphi)$, and vice-versa $r=(x^2+y^2)^{1/2}$ and $\varphi=\tan^{-1}(y/x)$. On exams, you don't need to memorize formulas like this. They'll be given at the end of the exam, or they will be in the notes (exams are open notes, no electronic devices).