Reminder:

Last class we started looking at a model of radioactive decay. For one C-14 atom, we found that the time-to-decay T was a random variable w/ cdf and pdf

(cdf) $F_T(t) = 1 - e^{-\lambda t}$

(pdf) $f_T(t) = \lambda e^{-\lambda t}$

In particular, this means that the probability of seeing a decay event by time t is

$$Pr[T \leq t] = 1 - e^{-\lambda t}$$ ← probability of decay before t

and the probability of seeing a decay event after time t is

$$Pr[T > t] = e^{-\lambda t}$$ ← probability of decay after t

We used this to look at a related (and more interesting) random variable: The number of undecayed atoms at time t, $N(t)$.

When we started w/ 1 atom, so $N(0)=1$, we found that

$$Pr[N(t)=0] = 1-e^{-\lambda t} \text{ and}$$
$$Pr[N(t)=1] = e^{-\lambda t},$$

so $\langle N \rangle = e^{-\lambda t}$.

Next, we found that when $N(0)=2$, we had

$$Pr[N(t)=0] = (1-e^{-\lambda t})^2 = 1-2e^{-\lambda t}+e^{-2\lambda t},$$
$$Pr[N(t)=1] = 2\cdot(1-e^{-\lambda t})\cdot e^{-\lambda t} = 2e^{-\lambda t}-2e^{-2\lambda t},$$
$$Pr[N(t)=2] = (e^{-\lambda t})^2 = e^{-2\lambda t}.$$

Notice that these add up to 1, which is a good check on our algebra.

We therefore have

$$\langle N \rangle = 0\cdot(1-2e^{-\lambda t}+e^{-2\lambda t}) + 1\cdot(2e^{-\lambda t}-2e^{-2\lambda t}) + 2\cdot e^{-2\lambda t}$$
$$= 2e^{-\lambda t}.$$

Finally, let's look at an arbitrary number of starting atoms, $N(0) = N_0$.

We want to find $\Pr[N(t) = n]$ for any $t \geq 0$ and any $0 \leq n \leq N_0$.

Notice that if we have n atoms left, then $N_0 - n$ of them decayed. If, for example, the first n had not decayed yet and the last $N_0 - n$ had decayed, we would have

$$\Pr[\text{first } n \text{ not decayed and last } N_0 - n \text{ decayed}]$$

$$= \Pr[T_1 > t \text{ and } \cdots \text{ and } T_n > t \text{ and } T_{n+1} \leq t \text{ and } \cdots \text{ and } T_{N_0} \leq t]$$

$$= \Pr[T_1 > t] \cdots \Pr[T_n > t] \cdot \Pr[T_{n+1} \leq t] \cdots \Pr[T_{N_0} \leq t]$$

$$= \underbrace{\left(e^{-\lambda t}\right) \cdots \left(e^{-\lambda t}\right)}_{n \text{ times}} \cdot \underbrace{\left(1 - e^{-\lambda t}\right) \cdots \left(1 - e^{-\lambda t}\right)}_{N_0 - n \text{ times}}$$

$$= \left(e^{-\lambda t}\right)^n \cdot \left(1 - e^{-\lambda t}\right)^{N_0 - n}.$$

We chose the first n to decay, but there are many other possibilities, each with the same probability. In particular, there are

$$\binom{N_0}{n} = \frac{N_0!}{n!(N_0 - n)!}$$ ways to choose n atoms from N_0.

We therefore have

$$\Pr[N(t)=n] = \frac{N_0!}{n!\,(N_0-n)!} \cdot \left(1-e^{-\lambda t}\right)^{N_0-n} \cdot \left(e^{-\lambda t}\right)^{n}$$

This is called a binomial distribution. For convenience, let's let $e^{-\lambda t} = p$, so we have

$$P_n(t) = \frac{N_0!}{n!\,(N_0-n)!} (1-p)^{N_0-n} p^n.$$

We also will need the "binomial theorem":

$$(a+b)^m = \sum_{n=0}^{m} \frac{m!}{n!(m-n)!} a^n b^{m-n}.$$

In particular, this means that

$$\sum_{n=0}^{N_0} P_n(t) = \sum_{n=0}^{N_0} \frac{N_0!}{n!\,(N_0-n)!} (1-p)^{N_0-n} p^n$$

$$= (1-p+p)^{N_0}$$

$$= 1^{N_0}$$

$$= 1,$$

So $P_n(t)$ really is a probability distribution.

Now we want to find the expected value of N,

$$\langle N \rangle = \sum_{n=0}^{N_0} n \cdot \frac{N_0!}{n!(N_0-n)!} (1-p)^{N_0-n} p^n$$

$$= \sum_{n=1}^{N_0} n \cdot \frac{N_0!}{n!(N_0-n)!} (1-p)^{N_0-n} p^n \quad \leftarrow n=0 \text{ term is zero}$$

$$= \sum_{n=1}^{N_0} \frac{N_0!}{(n-1)!(N_0-n)!} (1-p)^{N_0-n} p^n.$$

Now let $N_0 = M_0 + 1$ and $n = m+1$. We have

$$\langle N \rangle = \sum_{m=0}^{M_0} \frac{(M_0+1)!}{m!(M_0-m)!} \cdot (1-p)^{M_0-m} p^{m+1}$$

$$= (M_0+1)p \underbrace{\sum_{m=0}^{M_0} \frac{M_0!}{m!(M_0-m)!} \cdot (1-p)^{M_0-m} p^m}_{=1}$$

$$= (M_0+1)p = N_0 p,$$

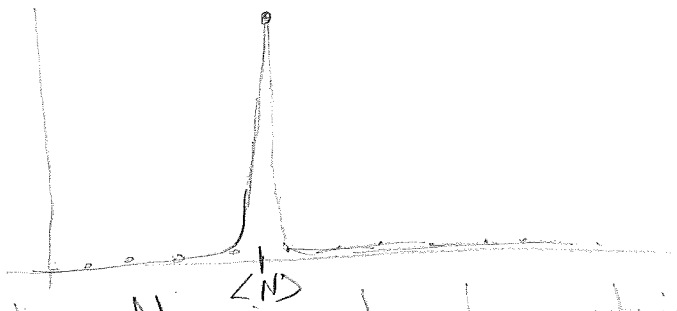
So $\boxed{\langle N(t) \rangle = N_0 e^{-\lambda t}}$.

This means that the usual "deterministic" model $N(t) = N_0 e^{-\lambda t}$ should really be thought of as a formula for the average number of remaining atoms.

Is this good enough? Should we just use the average value, or do we need to track more information about N?

Examples:

- Suppose $P_r[N=n]$ had a distribution like this:

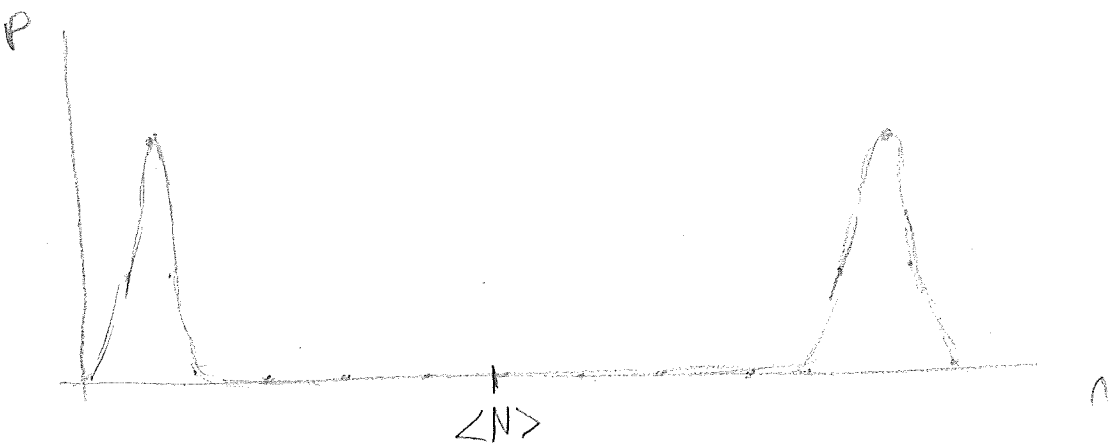


so that N is almost certainly one value. In this case, N is almost always the same as $\langle N \rangle$, so $\langle N \rangle$ is a useful measure of # of atoms.

- Now what if $Pr[N=n]$ looks like:

In this case, we wouldn't be surprised if N were very different than $\langle N \rangle$ (although $\langle N \rangle$ is still the most common value).

- Finally, what if $Pr[N=n]$ looks like:



Now N is almost never close to $\langle N \rangle$, and the most common values of N are very far from $\langle N \rangle$.

We need a way to decide how "spread out" a distribution is. We will use the standard deviation: $\text{std}[N(t)] = \sqrt{\text{Var}[N(t)]} = \sqrt{\langle N^2 \rangle - \langle N \rangle^2}$

Roughly speaking, you would not be surprised to find a value of N within ± 1 standard deviation of the average value.

(Details depend on the distribution, but for large N_0
$\approx 68\%$ of the time N is within one std. dev.
$\approx 95\%$ of the time N is within two std. dev.
$\approx 99.7\%$ of the time N is within three std. dev.)

What is the standard deviation of the binomial distribution? To start, we will calculate

$$\langle N \cdot (N-1) \rangle = \sum_{n=0}^{N_0} n \cdot (n-1) \frac{N_0!}{n!(N_0-n)!} (1-p)^{N_0-n} p^n$$

$$= \sum_{n=2}^{N_0} n \cdot (n-1) \frac{N_0!}{n!(N_0-n)!} (1-p)^{N_0-n} p^n$$

$$= \sum_{n=2}^{N_0} \frac{N_0!}{(n-2)!(N_0-n)!} (1-p)^{N_0-n} p^n$$

If we let $N_0 = M_0 + 2$ and $n = m + 2$, then

$$= \sum_{m=0}^{M_0} \frac{(M_0+2)!}{m!(M_0-m)!} (1-p)^{M_0-m} p^{m+2}$$

$$= (M_0+2)(M_0+1)p^2 \cdot \underbrace{\sum_{m=0}^{M_0} \frac{M_0!}{m!(M_0-m)!} (1-p)^{M_0-m} p^m}_{=1}$$

$$= N_0 \cdot (N_0 - 1) \cdot p^2.$$

We have:

$$\mathrm{Std}[N(t)] = \sqrt{\langle N^2 \rangle - \langle N \rangle^2}$$

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

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

$$= \sqrt{N_0(N_0-1)p^2 + N_0 p - (N_0 p)^2}$$

$$= \sqrt{N_0^2 p^2 - N_0 p^2 + N_0 p - N_0^2 p^2}$$

$$= \sqrt{N_0(p - p^2)}$$

$$= \sqrt{N_0 p (1-p)}.$$

This means that

$$\text{Std}[N(t)] = \sqrt{N_0 e^{-\lambda t}(1 - e^{-\lambda t})}.$$

Unfortunately, this gets very large when N_0 does.

For radioactive isotopes, N_0 might be $\approx 10^{20}$,

so $\text{Std}[N(t)] \sim 10^{10}$, which seems huge.

Why do we use the "deterministic model"?

10^{10} sounds at large in the abstract, but we should really compare it to something. Typically we compare the standard deviation to the expected value. This is called the coefficient of variation:

$$C_v = \frac{\text{Std}[N(t)]}{\langle N(t) \rangle}.$$

It is, roughly speaking, a measure of relative error.

In our case,

$$C_V = \frac{\sqrt{N_0 e^{-\lambda t}(1-e^{-\lambda t})}}{N_0 e^{-\lambda t}} = \frac{\sqrt{e^{\lambda t}(1-e^{-\lambda t})}}{\sqrt{N_0}}$$

$$= \frac{\sqrt{e^{\lambda t}-1}}{\sqrt{N_0}}$$

In particular, for any fixed time, the C_V goes down as N_0 goes up. A typical sample would therefore have a relative error on the order of $\sim 1/\sqrt{N_0}$.

The C_V also depends on time, though. If we wait longer, the error goes up. We have

$$C_V\sqrt{N_0} = \sqrt{e^{\lambda t}-1}$$

$$\Rightarrow C_V^2 N_0 + 1 = e^{\lambda t}$$

$$\Rightarrow \frac{1}{\lambda}\ln\left(C_V^2 N_0 + 1\right) = t.$$

If you wait long enough, the C_V can get arbitrarily large. It takes

$$t = \frac{1}{\lambda} \ln\left(C_V^2 N_0 + 1\right)$$

time to get a given C_V. Should we be worried?

A "bad" C_V is in the ballpark of 1.

This would mean

$$\begin{aligned} N(t) &= N_0 e^{-\lambda t} \\ &= N_0 e^{-\lambda \cdot \left(\frac{1}{\lambda} \ln(N_0+1)\right)} \\ &= N_0 e^{\ln\left(\frac{1}{N_0+1}\right)} \\ &= \frac{N_0}{N_0+1}, \end{aligned}$$

so the error doesn't get too bad until we expect to have ≈ 1 atom left.