Probabilistic Models for Radioactive Decay

Addendum to: https://imsc.uni-graz.at/keeling/skripten/modnsc.pdf

A central principle of radioactive decay is that radioactive particles have no memory. Specifically, observations show that the probability that a particle will decay in a forthcoming time interval is independent of how long it has survived up to the beginning of that time interval without decaying. Yet observations also show of course that the number of particles which have not decayed decreases exponentially on time scales comparable to the half-life of the radioactive substance. In the following, three approaches to the modelling of radioactive decay are considered.

The first and most widely accepted approach is to see radioactive decay as a homogeneous Poisson process, in which decay events are so rare, at least on a short time scale, that instead of seeing particles decaying next to each other in space one may as well see one particle decaying after another in time, with the same waiting time between decay events, since they have no memory and are independent of each other. The disadvantage of this approach is that it does not produce an expected number of undecayed particles which decreases exponentially, rather only linearly along a tangent line at very beginning of the exponential curve!

The second and not so widely accepted approach is to see radioactive decay as a nonhomogenous Poisson process, in which the exponential curve observed over longer time scales is approximated by ever more tangent lines, each with a different slope, where the changing slopes correspond to a higher incentive for the particles to decay when there are more particles. Beyond the rather ad hoc nature of this approach, there is also the underlying issue here of there being infinitely many particles: How can half of them decay at the half-life time?

The final and apparently most natural approach is to see radioactive decay as a Bernoulli process, as if many coins were flipping randomly, independently and without memory. Then the decay of exactly one particle is more probable when there are more particles, simply because it could be any one of them. It will be seen below that this approach has the advantages of preserving the memoryless property, preserving the independence of decays and producing the exponential decay of the expected number of undecayed particles for time scales on the order of the half-life.

1 Homogeneous Poisson Process

A homogeneous Poisson process is one in which the long-term average event rate is constant. Here the waiting time T between events satisfies

$$P(T > t + s \mid T > s) = P(T > t).$$

In words this property implies that in case $(T > s)$ holds, or that no event occurs during a waiting time s, then the conditional probability $P(T > t + s \mid T > s)$ that no event occurs within the forthcoming time t is the same $P(T > t)$ as if there had been no waiting time s. As a result of this memoryless condition it will be seen that the waiting time T is necessarily exponentially distributed, $P(T > t) = e^{-\lambda t}$.

Since $T > t + s$ implies that $T > s$, it follows with Bayes' rule,

$$P(T > t + s \mid T > s) = \frac{P(T > t + s \;\&\; T > s)}{P(T > s)} = \frac{P(T > t + s)}{P(T > s)}.$$

According to the memoryless property, the left side is $P(T > t)$, and the result is

$$P(T > t + s) = P(T > t)P(T > s).$$

With the survival function $S(t) = P(T > t)$ the memoryless property takes the form,

$$S(t + s) = S(t)S(s).$$

Observing

$$S(0) = P(T > 0) = 1$$

gives the following as a result of the memoryless property,

$$\frac{S(t + dt) - S(t)}{dt} = \frac{S(t)S(dt) - S(t)S(0)}{dt} = S(t)\frac{S(dt) - S(0)}{dt}.$$

After setting

$$\lambda = -\frac{S'(0)}{S(0)} = -S'(0)$$

and letting $dt \to 0$ gives

$$S'(t) = -\lambda S(t), \quad S(0) = 1.$$

Solving this initial value problem shows that the random variable T has the exponential distribution,

$$S(t) = P(T > t) = e^{-\lambda t}.$$

The probability density $f(t)$ is given by

$$P(T > t) = \int_t^\infty f(t)dt \quad \text{or} \quad f(t) = \lambda e^{-\lambda t}$$

und hence,

$$P(T \in [t_1, t_2]) = e^{-\lambda t}|_{t_1}^{t_2} = e^{\lambda t_1} - e^{\lambda t_2}.$$

Furthermore, with the Bayes' rule,

$$\begin{aligned} P(T \in [t_1, t_2] \mid T > t_1) &= \frac{P(T \in [t_1, t_2] \ \& \ T > t_1)}{P(T > t_1)} = \frac{P(T \in [t_1, t_2])}{P(T > t_1)} \\ &= 1 - e^{-\lambda(t_2 - t_1)} = P(T < (t_2 - t_1)). \end{aligned}$$

Now let the random variable $N(\tau)$ denote the number of events in a homogeneous Poisson process with parameter λ that occur in a time interval of length τ. As a result of the last calculation, the conditional probability

$$\begin{aligned} P(N(t + dt) = k + 1 \mid N(t) = k) &= P(T \in [t, t + dt] \mid T > t) \\ &= P(T < dt) = 1 - e^{-\lambda dt} = \lambda dt + o(dt) \end{aligned}$$

is independent of t because of the memoryless property, but it is also independent of k! Hence, the probability that a given particle will decay in a forthcoming time interval $[t, t + dt]$ is independent of the behavior of any other particles. The probability distribution of $N(\tau)$ is given by determining $p_k(\tau) = P(N(\tau) = k)$, $k \in \mathbb{N}_0$. For there to be k events in the time interval $[0, \tau]$, there must be

- $k - 1$ events in an interval $[0, s]$, $s \in (0, \tau)$ with probability

$$P(N(s) = k - 1) = p_{k-1}(s)$$

- 1 event in an interval $[s, s+ds]$ with probability

$$\begin{gathered}P(N(s+ds)=k \mid N(s)=k-1)=P(T\in[s,s+ds]\mid T>s)=\\=1-e^{-\lambda ds}=\lambda ds+o(ds)\end{gathered}$$

and

- no events in an interval $[s+ds, \tau]$ with probability

$$\begin{aligned}&P(N(\tau)=k \mid N(s+ds)=k)=P(T>\tau\mid T>s+ds)\\=&\frac{P(T>\tau)}{P(T>s+ds)}=P(T>(\tau-s-ds))=e^{-\lambda(\tau-s)}+o(ds).\end{aligned}$$

Since these events are independent, the probability of all three is the product of the three, integrated over all possible intermediate times s,

$$p_k(\tau)=\int_0^\tau p_{k-1}(s)(\lambda ds)e^{-\lambda(\tau-s)}=e^{-\lambda\tau}\int_0^\tau p_{k-1}(s)\lambda e^{\lambda s}ds,\quad k\in\mathbb{N}.$$

After differentiating,

$$p_k'(\tau)=-\lambda e^{-\lambda\tau}\int_0^\tau p_{k-1}(s)\lambda e^{\lambda s}ds+e^{-\lambda\tau}\lambda p_{k-1}(\tau)e^{+\lambda\tau}\lambda$$

or

$$p_k'(\tau)=-\lambda p_k(\tau)+\lambda p_{k-1}(\tau),\quad k\in\mathbb{N}.$$

For the case $k=0$,

$$p_0(\tau)=P(N(\tau)=0 \text{ in } [0,\tau])=P(T>\tau)=e^{-\lambda\tau},\quad\text{or}\quad p_0'(\tau)=-\lambda p_0(\tau).$$

Since no events can occur in no time, it holds for $\tau=0$,

$$p_0(0)=1,\quad p_k(0)=0,\quad k\in\mathbb{N}.$$

The solution to this system is

$$P(N(\tau)=k)=p_k(\tau)=\frac{(\lambda\tau)^k}{k!}e^{-\lambda\tau},\quad k\in\mathbb{N}_0.$$

The expected number $n(\tau)=\mathbb{E}(N(\tau))$ of events that occur in a time interval of length τ in a homogeneous Poisson process with parameter λ is given by:

$$\begin{aligned}n(\tau)&=\sum_{k=0}^\infty kP(N(\tau)=k)=\sum_{k=1}^\infty kp_k(\tau)=\int_0^\tau\sum_{k=1}^\infty kp_{k-1}(s)\lambda e^{-\lambda(\tau-s)}ds\\&=\int_0^\tau\underbrace{\sum_{k=1}^\infty(k-1)p_{k-1}(s)}_{=n(s)}\lambda e^{-\lambda(\tau-s)}ds+\int_0^\tau\underbrace{\sum_{k=1}^\infty p_{k-1}(s)}_{=1}\lambda e^{-\lambda(\tau-s)}ds\end{aligned}$$

or

$$n(\tau)=e^{-\lambda\tau}\left\{\int_0^\tau n(s)\lambda e^{\lambda s}ds+\int_0^\tau\lambda e^{\lambda s}ds\right\}.$$

After differentiating,

$$n'(\tau) = -\lambda e^{-\lambda\tau}\left\{\int_0^\tau n(s)\lambda e^{\lambda s}ds + \int_0^\tau \lambda e^{\lambda s}ds\right\} + e^{-\lambda\tau}\left\{n(\tau)\lambda e^{\lambda\tau} + \lambda e^{\lambda\tau}\right\}$$

or

$$n'(\tau) = -\lambda n(\tau) + \lambda n(\tau) + \lambda, \quad n(0) = 0.$$

The solution is

$$n(\tau) = \lambda\tau.$$

Note that this linear growth property is only adequate when τ is very small, e.g., in relation to the half-life $\ln(2)/\lambda$ of the radioactive material. The expected growth pattern is of the form $1 - e^{-\mu\tau}$, where $e^{-\mu\tau}$ is the fraction of material not yet decayed after time τ.

2 Nonhomogeneous Poisson Process

To obtain an exponential decay pattern, a nonhomogeneuous Poisson process must be considered, in which the average event rate changes with time. For this purpose let T now denote the random variable representing the waiting time from $t = 0$ until a decay event occurs, and suppose that the function $S(t) = P(T > t)$ satisfies

$$\lambda(t) = -\frac{S'(t)}{S(t)} = -\frac{d}{dt}\ln(S(t))$$

for a time dependent rate $\lambda(t)$. For $t_2 > t_1 \geq 0$,

$$\ln\left(\frac{S(t_2)}{S(t_1)}\right) = \ln(S(t_2)) - \ln(S(t_1)) = -\int_{t_1}^{t_2} \lambda(t)dt$$

or

$$S(t_2) = S(t_1)\exp\left[-\int_{t_1}^{t_2} \lambda(t)dt\right].$$

With $t_2 = t > 0$, $t_1 = 0$ and $S(0) = 1$,

$$S(t) = \exp\left[-\int_0^t \lambda(t)dt\right].$$

Since

$$S(t_2 - t_1) = \exp\left[-\int_0^{t_2-t_1} \lambda(t)dt\right] \neq \exp\left[-\int_{t_1}^{t_2} \lambda(t)dt\right] = S(t_2)/S(t_1)$$

holds, the memoryless property is not satisfied! The probability density $f(t)$ of T is given by

$$P(T > t) = \int_t^\infty f(t)dt \quad \text{or} \quad f(t) = \lambda(t)\exp\left[-\int_0^t \lambda(s)ds\right].$$

and thus

$$\begin{aligned} P(T \in [t_1, t_2]) &= \int_{t_1}^{t_2} \lambda(t)\exp\left[-\int_0^t \lambda(s)ds\right]ds = -\exp\left[-\int_0^t \lambda(s)ds\right]\Bigg|_{t_1}^{t_2} \\ &= \exp\left[-\int_0^{t_1} \lambda(s)ds\right]\left\{1 - \exp\left[-\int_{t_1}^{t_2} \lambda(s)ds\right]\right\}. \end{aligned}$$

Furthermore, with the Bayes' rule,

$$P(T \in [t_1, t_2] \mid T > t_1) = \frac{P(T \in [t_1, t_2] \ \&\ T > t_1)}{P(T > t_1)} = \frac{P(T \in [t_1, t_2])}{P(T > t_1)}$$
$$= 1 - \exp\left[-\int_{t_1}^{t_2} \lambda(s) ds\right].$$

Now let the random variable $N(\tau)$ denote the number of events in a nonhomogeneous Poisson process with rate $\lambda(t)$ that occur in the time interval $[0, \tau]$. As a result of the last calculation, the conditional probability

$$P(N(t + dt) = k + 1 \mid N(t) = k) = P(T \in [t, t + dt] \mid T > t)$$
$$= 1 - \exp\left[-\int_{t}^{t+dt} \lambda(r) dr\right] = \lambda(t) dt + o(dt)$$

depends upon t because a particle possess memory in the present context, but the conditional probability is independent of k! Hence, the probability that a given particle will decay in a forthcoming time interval $[t, t + dt]$ is independent of the behavior of any other particles. The probability distribution of $N(\tau)$ is given by determining $p_k(\tau) = P(N(\tau) = k)$, $k \in \mathbb{N}_0$. For there to be k events in the time interval $[0, \tau]$, there must be

- $k - 1$ events in an interval $[0, s]$, $s \in (0, \tau)$ with probability

  $$P(N(s) = k - 1) = p_{k-1}(s)$$

- 1 event in an interval $[s, s + ds]$ with probability

  $$P(N(s + ds) = k \mid N(s) = k - 1) = P(T \in [s, s + ds] \mid T > s)$$
  $$= 1 - \exp\left[-\int_{s}^{s+ds} \lambda(r) dr\right] = \lambda(s) ds + o(ds)$$

  and

- no events in an interval $[s + ds, \tau]$ with probability

  $$P(N(\tau) = k \mid N(s + ds) = k) = P(T > \tau \mid T > s + ds) = \frac{P(T > \tau)}{P(T > s + ds)}$$
  $$= \exp\left[-\int_0^{\tau} \lambda(r) dr\right] / \exp\left[-\int_0^{s+ds} \lambda(r) dr\right] = \exp\left[-\int_s^{\tau} \lambda(r) dr\right] + o(ds).$$

Since these events are independent, the probability of all three is the product of the three, integrated over all possible intermediate times s,

$$p_k(\tau) = \int_0^{\tau} p_{k-1}(s)(\lambda(s) ds) \exp\left[-\int_s^{\tau} \lambda(r) dr\right]$$
$$= \exp\left[-\int_0^{\tau} \lambda(r) dr\right] \int_0^{\tau} p_{k-1}(s) \lambda(s) \exp\left[\int_0^{s} \lambda(r) dr\right] ds, \quad k \in \mathbb{N}.$$

After differentiating,

$$p_k'(\tau) = -\lambda(\tau) \exp\left[-\int_0^{\tau} \lambda(r) dr\right] \int_0^{\tau} p_{k-1}(s) \lambda(s) \exp\left[\int_0^{s} \lambda(r) dr\right]$$
$$+ \exp\left[-\int_0^{\tau} \lambda(r) dr\right] p_{k-1}(\tau) \lambda(\tau) \exp\left[\int_0^{\tau} \lambda(r) dr\right]$$

or

$$p_k'(\tau) = -\lambda(\tau)p_k(\tau) + \lambda(\tau)p_{k-1}(\tau), \quad k \in \mathbb{N}.$$

For the case $k = 0$,

$$p_0(\tau) = P(N(\tau) = 0 \text{ in } [0,\tau]) = P(T > \tau) = \exp\left[-\int_0^\tau \lambda(s)ds\right], \quad \text{or} \quad p_0'(\tau) = -\lambda(\tau)p_0(\tau).$$

Since no events can occur in no time, it holds for $\tau = 0$,

$$p_0(0) = 1, \quad p_k(0) = 0, \quad k \in \mathbb{N}.$$

The solution to this system is

$$P(N(\tau) = k) = p_k(\tau) = \exp\left[-\int_0^\tau \lambda(s)ds\right] \times$$
$$\left\{\int_0^\tau ds_1\lambda(s_1)\int_0^{s_1} ds_2\lambda(s_2)\cdots\int_0^{s_{k-1}} ds_k\lambda(s_k)\right\}, \quad k \in \mathbb{N}_0.$$

The expected number $n(\tau) = \mathbb{E}(N(\tau))$ of events that occur in a time interval of length τ in a homogeneous Poisson process with parameter λ is given by:

$$n(\tau) = \sum_{k=0}^{\infty} kP(N(\tau) = k) = \sum_{k=1}^{\infty} kp_k(\tau) = \int_0^\tau \sum_{k=1}^{\infty} kp_{k-1}(s)\lambda(s)\exp\left[-\int_s^\tau \lambda(r)dr\right]ds$$
$$= \int_0^\tau \underbrace{\sum_{k=1}^{\infty}(k-1)p_{k-1}(s)}_{=n(s)}\lambda(s)\exp\left[-\int_s^\tau \lambda(r)dr\right]ds + \int_0^\tau \underbrace{\sum_{k=1}^{\infty} p_{k-1}(s)}_{=1}\lambda(s)\exp\left[-\int_s^\tau \lambda(r)dr\right]ds$$

or

$$n(\tau) = \exp\left[-\int_0^\tau \lambda(r)dr\right]\left\{\int_0^\tau n(s)\lambda(s)\exp\left[\int_0^s \lambda(r)dr\right]ds + \int_0^\tau \lambda(s)\exp\left[\int_0^s \lambda(r)dr\right]ds\right\}.$$

After differentiating,

$$n'(\tau) = -\lambda(\tau)\exp\left[-\int_0^\tau \lambda(r)dr\right]\left\{\int_0^\tau n(s)\lambda(s)\exp\left[\int_0^s \lambda(r)dr\right]ds + \int_0^\tau \lambda(s)\exp\left[\int_0^s \lambda(r)dr\right]ds\right\}$$
$$+\exp\left[-\int_0^\tau \lambda(r)dr\right]\left\{n(\tau)\lambda(\tau)\exp\left[\int_0^\tau \lambda(r)dr\right] + \lambda(\tau)\exp\left[\int_0^\tau \lambda(r)dr\right]\right\}$$

or

$$n'(\tau) = -\lambda(\tau)n(\tau) + \lambda(\tau)n(\tau) + \lambda(\tau), \quad n(0) = 0.$$

The solution is

$$n(\tau) = \int_0^\tau \lambda(s)ds.$$

The anticipated form for the inhomogeneous Poisson process is

$$1 - e^{-\mu\tau} = n(\tau)/n(\infty) = \int_0^\tau \lambda(s)ds/\int_0^\infty \lambda(s)ds.$$

After differentialting,

$$\mu e^{-\mu\tau} = \lambda(\tau)/\int_0^\infty \lambda(s)ds$$

or

$$\lambda(t) = \nu\mu e^{-\mu t}, \quad \nu = \int_0^\infty \lambda(s)ds = n(\infty).$$

Thus, ν can be interpreted as the total number of particles to decay. Yet a Poisson process involves infinitely many particles,

$$P(N(\tau) = k) > 0, \quad \forall\tau > 0, \quad \forall k \in \mathbb{N}_0.$$

3 Bernoulli Process

It appears then to be more natural to model radioactive decay as a Bernoilli process. Based upon discussions of the survival function above, let $X(t) : [0, \infty) \to \{0, 1\}$ be a random variable representing the decay of a single radioactive particle with

$$X(t) = \begin{cases} 1, & \text{decay in } [0, t] \\ 0, & \text{otherwise} \end{cases}$$

and

$$P(X(t) = 0) = e^{-\lambda t}, \quad P(X(t) = 1) = 1 - e^{-\lambda t}, \quad t \geq 0.$$

Then the conditional probability $P(X(s + t) = 0 \mid X(s) = 0)$ satisfies

$$\begin{aligned} P(X(s+t) = 0 \mid X(s) = 0) &= \frac{P(X(s+t) = 0 \;\&\; X(s) = 0)}{P(X(s) = 0)} = \frac{P(X(s+t) = 0)}{P(X(s) = 0)} \\ &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P(X(t) = 0). \end{aligned}$$

Thus, a single particle satisfies the memoryless condition. Now let $X_k(t)$ be a random variable representing the decay of the kth particle, $k = 1, \ldots, n$. Assume that all random variables X_k are independent and identically distributed with the same distribution as $X(t)$. Let

$$N(t) = \sum_{k=1}^{n} X_k(t)$$

be a random variable denoting the number of particles which decay in the time interval $[0, t]$. Then, as is well-known for a Bernoulli process, $N(t)$ has the binomial distribution

$$P(N(t) = k) = \binom{n}{k} (1 - e^{-\lambda t})^k (e^{-\lambda t})^{n-k}, \quad t \geq 0$$

the expected value

$$\mathbb{E}(N(t)) = n(1 - e^{-\lambda t})$$

and the variance

$$\mathbb{E}((N(t) - \mathbb{E}(N(t)))^2) = ne^{-\lambda t}(1 - e^{-\lambda t}).$$

Note that the expected value $\mathbb{E}(N(t)) = n(1 - e^{-\lambda t})$ corresponds precisely to the observation of exponential decay for radioactive substances.

To compare with the Poisson processes, set

$$p_k(t) = P(N(t) = k) = \binom{n}{k} (1 - e^{-\lambda t})^k (e^{-\lambda t})^{n-k}, \quad t \geq 0$$

for which a system of differential equations is constructed as follows. For $k=0$,

$$p_0(t) = e^{-\lambda nt}, \quad p_0'(t) = -\lambda n e^{-\lambda nt} = -\lambda n p_0(t).$$

For $k=n$,

$$p_n(t) = (1-e^{-\lambda t})^n, \quad p_n'(t) = n(1-e^{-\lambda t})^{n-1}(+\lambda e^{-\lambda t}) = \lambda p_{n-1}(t)$$

where

$$p_{n-1}(t) = \left(\begin{array}{c} n \\ n-1 \end{array}\right)(1-e^{-\lambda t})^{n-1}(e^{-\lambda t})^{n-(n-1)} = n(1-e^{-\lambda t})^{n-1}(e^{-\lambda t}).$$

For $k=1,\ldots,n-1$,

$$\begin{aligned} p_k'(t) &= \left(\begin{array}{c} n \\ k \end{array}\right)\left[k(1-e^{-\lambda t})^{k-1}(+\lambda e^{-\lambda t})(e^{-\lambda t})^{n-k} + (n-k)(e^{-\lambda t})^{n-k-1}(-\lambda e^{-\lambda t})(1-e^{-\lambda t})^k\right] \\ &= \left(\begin{array}{c} n \\ k \end{array}\right)\left[\lambda(n-k+1)\frac{k}{n-k+1}(1-e^{-\lambda t})^{k-1}(e^{-\lambda t})^{n-k+1} - \lambda(n-k)(1-e^{-\lambda t})^k(e^{-\lambda t})^{n-k}\right] \end{aligned}$$

and using

$$\begin{aligned} p_{k-1}(t) &= \left(\begin{array}{c} n \\ k-1 \end{array}\right)(1-e^{-\lambda t})^{k-1}(e^{-\lambda t})^{n-k+1} \\ &= \left(\begin{array}{c} n \\ k \end{array}\right)\frac{k}{(n-k+1)}(1-e^{-\lambda t})^{k-1}(e^{-\lambda t})^{n-k+1} \end{aligned}$$

it follows finally that

$$p_k'(t) = \lambda(n-k+1)p_{k-1}(t) - \lambda(n-k)p_k(t), \quad k=0,\ldots,n.$$

Once this system is written in integral form,

$$p_k(t) = \int_0^t p_{k-1}(s)(\lambda(n-k+1))e^{-\lambda(n-k)(t-s)}ds, \quad k=1,\ldots,n, \quad p_0(t) = e^{-\lambda t}$$

the following can be deduced. For there to be k events in the time interval $[0,\tau]$, there must be

- $k-1$ events in an interval $[0,s]$, $s \in (0,\tau)$ with probability

$$P(N(s)=k-1) = p_{k-1}(s)$$

- 1 event in an interval $[s, s+ds]$ with probability

$$P(N(s+ds)=k \mid N(s)=k-1) = (n-k+1)\lambda ds + o(ds)$$

  and

- no events in an interval $[s+ds,\tau]$ with probability

$$P(N(t)=k \mid N(s+ds)=k) = e^{-\lambda(n-k)(t-s)} + o(ds).$$

Hence, the probability of some decay in an interval $[s, s+ds]$ is ever higher, the more particles there are which have not yet decayed at time s. If the number of particles is extremely large, and the time interval ds is extremely small, then the initial decays can be approximated by a homogeneous Poisson process.

4 Further Reading

- Cory Simon, *The memoryless Poisson process and volcano insurance.*
- Robert DeSerio, *Statistical Analysis of Data, Exponential Decay and Poisson Processes*
- Jem Corcoran, *Stochastic Simulation, The Non-Homogeneous (Non-Stationary) Poisson Process*
- Robert Gallager, *Stochastic Processes, Theory for Applications, Chapter 2: Poisson Processes*
- Wiki, Radioactive Decay