Chapter 18

Exponential processes

18.1 Exponential probability distribution

Suppose we have a certain event that occurs (or arrives) at a random time T. The event could be a text message to your phone, a wasp arriving at your barbeque, a shower of rain, your computer disk crashing, a flood, for example. In many cases the time of the event (its *arrival time*) can be modelled by exponential-type distributions, although this is an over-simplification.

The pure exponential distribution has the probability density function:

$$f_T(t) = \begin{cases} 0 & t < 0 \\ \lambda \mathrm{e}^{-\lambda t} & t \geqslant 0 \end{cases} \quad . \tag{18.1}$$

There is only a single parameter for this distribution, $\lambda > 0$. The figure (18.1) displays the shape of the distribution for $\lambda = 0.5, 1.0, 2.0$. The larger the value of λ, the steeper the curve.

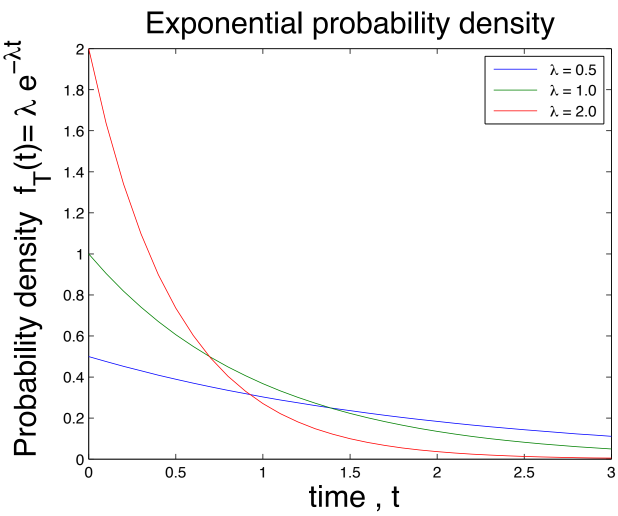

Figure 18.1: Probability density for the exponential distribution for $\lambda = 0.5, 1.0, 2.0$.

Then, the probability distribution has the form:

$$P(T \leqslant t) = F_T(t) = 1 - \mathrm{e}^{-\lambda t} \quad . \tag{18.2}$$

In previous lectures, it was shown that:

$$\boxed{\mathbb{E}(T) = \frac{1}{\lambda}} \quad , \quad \boxed{\mathrm{var}(T) = \frac{1}{\lambda^2}} \quad . \tag{18.3}$$

These relations should be committed to memory.

Similarly the probability that the event has not yet occurred is:

$$P(T > t) = 1 - P(T \leqslant t) = \mathrm{e}^{-\lambda t} \tag{18.4}$$

and this is sometimes called the *survival probability*, particularly in the field of medicine in which the event is a life-threatening condition. So we can use the equivalent notation:

$$\boxed{S_T(t) \equiv P(T > t)} \quad . \tag{18.5}$$

It is a unique property of the exponential distribution that it has *no memory.* This property is expressed by the relation:

$$\boxed{P(T > s + t | T > t) = P(T > s)} \qquad \text{for all } s, t > 0 \tag{18.6}$$

Let $T \geqslant 0$ be the time at which your laptop dies. Then this equation says that, (LHS) given the device has survived beyond a time t, the probability of it surviving an additional time s is the same as the probability that it survives for a time s starting from $t = 0$. That is, if the device is still working at time t, it 'forgets' that it has already been working that long, and 'thinks' it is starting from $t = 0$. That 'forgetfulness' is why we refer to the *no memory* property.

By definition of conditional probability:

$$P(X > s + t | X > t) = \frac{P(X > s + t \text{ and } X > t)}{P(X > t)}$$

The intersection of the intervals $T > s + t$ and $T > t$, that is the common value of time is, $T > s + t$:

if no 'memory'

$$P(X > s + t | X > t) \; = \frac{P(X > s + t)}{P(X > t)} \; = P(X > s)$$

Then the equivalent form of the 'no memory' equation is:

$$\boxed{P(T > s + t) = P(T > s) \cdot P(T > t)}$$

This is true for the exponential distribution, and no other distribution:

$$\begin{aligned} P(T > s) &= \mathrm{e}^{-\lambda s} \\ P(T > s + t) &= \mathrm{e}^{-\lambda(s+t)} = \mathrm{e}^{-\lambda s}\mathrm{e}^{\lambda t} \\ &= P(T > s) \cdot P(T > t) \end{aligned}$$

EXAMPLE Suppose the waiting time on a phone call to a call centre is an exponential distribution with

$\lambda = 0.1$ per minute.

$$\underset{\substack{\downarrow \\ \text{waiting time}}}{P(T > t)} = \mathrm{e}^{-\lambda t} = \mathrm{e}^{-0.1t}$$

(a) What is the probability that you have to wait more than 15 minutes?

(b) Given you have waited 10 minutes, what is the probability you need to wait another 15 minutes

(a) $P(T > 15 \text{ minutes}) = \mathrm{e}^{-\lambda t} = \mathrm{e}^{-0.1\times 15} = \mathrm{e}^{-1.5} \approx 0.223$

(b) $P(T > 25 \text{ mins}|T > 10 \text{ mins}) = \dfrac{P(T > 25)}{P(T > 10)} = \dfrac{\mathrm{e}^{-2.5}}{\mathrm{e}^{-1.0}} = \mathrm{e}^{-1.5} \approx 0.223$ (again)

That is, the call centre has 'no memory' of you having already waited 10 minutes! Therefore, the waiting process restarts from zero!

18.2 Hazards and survival

Suppose we extend the exponential distribution to a more general form. That is, we consider the *survival function* to be of the form:

$$\boxed{S_T(t) = P(T > t) = e^{-\int_0^t \lambda(\tau)d\tau}} \quad . \tag{18.7}$$

where $\lambda(\tau)$ is now a function of time. Note that we require:

$$\lim_{t\to\infty} S_T(t) = 0 \tag{18.8}$$

and thus

$$\lim_{t\to\infty} t\lambda(t) = +\infty. \tag{18.9}$$

If $\lambda(\tau) = \lambda_0$ (a constant, then we have the usual exponential distribution. The corresponding probability distribution function will then be:

$$\boxed{F_T(t) = P(T \le t) = 1 - e^{-\int_0^t \lambda(\tau)d\tau}} \quad . \tag{18.10}$$

The associated probability density then follows as:

$$f_T(t) = \frac{d}{dt}F_T(t) = e^{-\int_0^t \lambda(\tau)d\tau}\frac{d}{dt}\left(\int_0^t \lambda(\tau)d\tau\right) \quad . \tag{18.11}$$

That is,

$$\boxed{f_T(t) = \lambda(t)e^{-\int_0^t \lambda(\tau)d\tau}} \quad . \tag{18.12}$$

The function $\lambda(t)$, in this context is known as a *hazard function* or *failure rate*. We can understand the *meaning* of the *hazard function* by considering a conditional probability density:

$$P(t \le T \le t + dt|T > t) = \frac{f_T(t)dt}{S_T} = \lambda(t)dt \quad . \tag{18.13}$$

That is, the hazard function is proportional to the probability that the event will occur in the interval of time $[t, t+dt]$ *given* that has not occurred before t. In the context of life insurance, the hazard rate bears the rather morbid name *force of mortality*, since it represents the instantaneous rate of death given that you have survived up to that point!

The exponential distribution is a unique example of a constant hazard function, but this is an exceptional case. In general, one would expect the hazard function to vary in time. For example. the risk of a car breakdown might be expected to increase as the car gets older. Conversely, one might expect that the risk of a driver having an accident decreases with time as they become more experienced.

The expected time for the arrival of this event is given by:

$$\mathbb{E}\left(T\right) = \int_0^{+\infty} t.\ f_T(t)dt = \int_0^{+\infty} t.\ \lambda(t)e^{-\int_0^t \lambda(\tau)d\tau}dt \tag{18.14}$$

Integrating by parts, noting that,

$$\lim_{t\to\infty} te^{-\int_0^t \lambda(\tau)d\tau} = 0 \quad ,$$

according to (18.9). Then this gives the result for the expected time of arrival of the event as:

$$\boxed{\mathbb{E}\left(T\right) = \int_0^{+\infty} e^{-\int_0^t \lambda(\tau)d\tau}dt} \quad . \tag{18.15}$$

A very simple extension is to take the linear function:

$$h_T(t) = \lambda(t) = a + bt \quad . \tag{18.16}$$

With $a > 0$ and $b > 0$ are constants, representing a hazard rate increasing over time. In this case we have

$$\int_0^t \lambda(\tau)d\tau = \int_0^t (a + b\tau)d\tau = at + \frac{1}{2}bt^2 \quad . \tag{18.17}$$

Then the corresponding survival function is given by:

$$S_T(t) = e^{-at-\frac{1}{2}bt^2} \quad . \tag{18.18}$$

This gives the expected time for the event arrival:

$$\mathbb{E}\left(T\right) = \sqrt{\frac{2\pi}{b}}e^{a^2/(2b)}\left[1 - \Phi\left(\frac{a}{\sqrt{b}}\right)\right] \quad , \tag{18.19}$$

with Φ the standard normal probability distribution function.

Another simple analytic form is given by the *Weibull distribution*:

$$\lambda(t) = \lambda_0\gamma t^{\gamma-1} \tag{18.20}$$

where $\lambda_0, \gamma > 0$, and $t \geq 0$. Again, this function is used two describe a hazard rate that either increases $(\gamma > 1)$ or decreases $(0 < \gamma < 1)$ with time. Then this gives the survival function:

$$S_T(t) = \exp\left(-\lambda_0 t^\gamma\right) \quad . \tag{18.21}$$

The *expected time* for the event is given by (18.15),

$$\mathbb{E}\left(T\right) = \int_0^{+\infty} \exp\left(-\lambda_0 t^\gamma\right) dt \quad . \tag{18.22}$$

Changing variable, $z = \lambda_0 t^\gamma$, we have:

$$\mathbb{E}(T) = \frac{1}{\gamma \lambda_0^{1/\gamma}} \int_0^{+\infty} e^{-z} z^{1/\gamma - 1} dz \qquad . \tag{18.23}$$

The integral is now in the form of the well-known gamma function defined by:

$$\Gamma(u) \equiv \int_0^{+\infty} e^{-z} z^{u-1} dz \qquad , \quad u > 0 \qquad . \tag{18.24}$$

Finallly we have the following expression for the expected time for the event as:

$$\boxed{\mathbb{E}(T) = \lambda_0^{-1/\gamma} \Gamma(1 + 1/\gamma)} \qquad . \tag{18.25}$$

where we have used the relation: $z\Gamma(z) = \Gamma(1 + z)$.

Of course, for the special case $\gamma = 1$, this is the same as the exponential distribution. What we wish to consider are the implications of (a) $\gamma < 1$, and (b) $\gamma > 1$. In the graph below (figure 18.2) we plot the hazard function for $\gamma = 0.5, 1.0$ and 1.5 for comparison. We can think of the hazard as it applies to some system (a machine or piece of software for example) which is prone to breakdown/failure at some point.

- For $\gamma < 1$ the hazard function decreases with increasing time. This could be used to model a system for which breakdown/failure is more likely in the early stages of use, for example a new piece of software, or a pilot learning to fly.
- $\gamma = 1$, that is the exponential distribution, in which failure can occur at anytime and with the same likelihood, for which there is no memory of the total time of operation.
- $\gamma > 1$, there is a steady increase of $\lambda(t)$. This might correspond to wear and tear, so that the older the system is, the more likely it is to fail.

There is a similar analogy for medical conditions - those which increase/decrease in their hazard level with the age of the patients, paediatric conditions and geriatric conditions.

A sketch of the Weibull densities for $\lambda_0 = 1$, and $\gamma = 0.5, 1.0, 1.5, 2.0$ is given in figure (18.3)

18.2.1 Radioactivity

There is one very important process in nature that follows the exponential distribution with astonishing accuracy. The process of nuclear radioactive decay which drives nuclear energy generation. It is also of great importance in medical physics, in which radioactive tracers are used to track the movement of a substance through the body.

For example, the element Iodine has many different isotopes (over 20 in fact). The isotope called I-123 has a radioactive decay (through a process called electron capture in which an associated x-ray is emitted) used to study the function of the thyroid gland in particular. As the substance passes through the thyroid, the x-ray emissions can be viewed by a gamma-ray camera and thus provide a moving real-time image of the organ. The waste products from nuclear reactors are potentially harmful as they are still radioactive. An example is the α-decay of Plutonium. This decay process produces α particles which are potentially harmful.

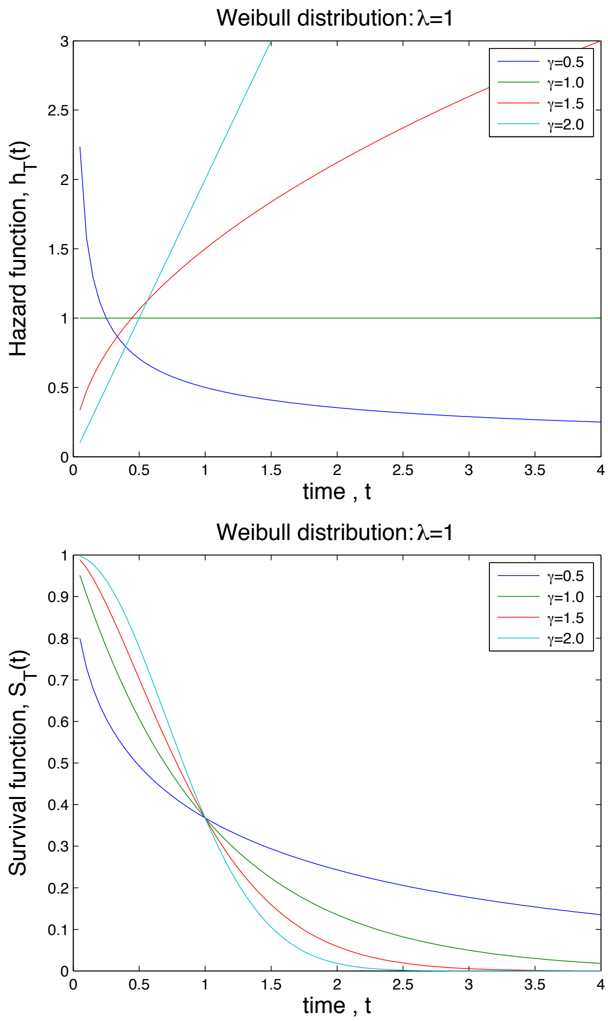

Figure 18.2: Top: Hazard function, $h_T(t) = \lambda(t) = \lambda_0 \gamma t^{\gamma-1}$, for the Weibull distribution for $\lambda_0 = 1$ and $\gamma = 0.5, 1, 0, 1.5, 2.0$. Bottom: the corresponding survival functions.

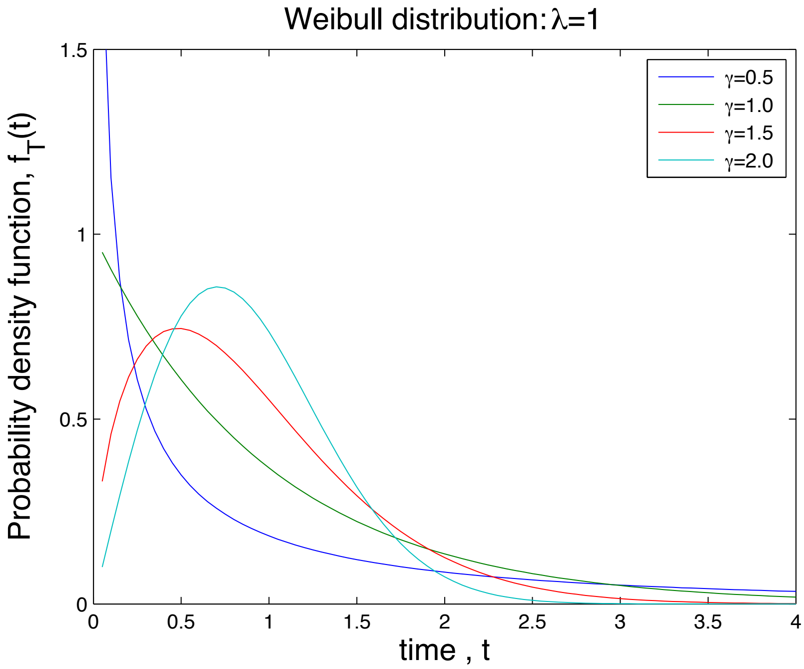

Figure 18.3: The probability density function for the Weibull distribution: $f_T(t) = \lambda_0\gamma t^{\gamma-1}e^{-\lambda_0 t^{\gamma-1}}$. for $\lambda = 1$ and $\gamma = 0.5, 1, 0, 1.5.2.0$.

In all cases of radioactive decay, the decay process, for each nucleus, proceeds according to an exponential distribution. One defines the *half-life* $T_{1/2}$ of a radioactive isotope as the amount of time it takes for half the radioactive atoms in a sample to decay. Thus one has:

$$P(T > T_{1/2}) = e^{-\lambda T_{1/2}} = \frac{1}{2} \qquad . \tag{18.26}$$

That is, we end up with the formula known to every nuclear physicist:

$$\lambda = \frac{\ln 2}{T_{1/2}} \qquad . \tag{18.27}$$

Now, we recall that λ is the hazard rate, that is the rate of decays in a fixed time, so a shorter half-life means a higher rate of emissions.

For the Iodine isotope we have, $T_{1/2} = 13.22$ hours. That is one expects half the atoms to decay within 13.22 hours. This means a short-term *high* hazard rate, λ. It is good for the body that the substance rapidly decays, but this means that the radioactivity rate is high, and thus the correct dosage is important. It also makes it difficult for suppliers, since this isotope needs to be made on-demand. One cannot store/deep freeze a radioactive isotope, and for this reason, some hospitals have their own mini-nuclear reactors.

On the other hand, the Plutonium half-life is 87.7 years! While this might seem more hazardous, in fact the the radioactive decay rate λ is much much less over the same time-scale, albeit α rays are more harmful than γ rays. But that's a subject for another day.

18.2.2 Example

Suppose that an exponentially distributed event, has a rate λ. One can ask what the value of the following conditional expectation,$\mathbb{E}\left(T|T>c\right)$, would be. That is, what is the expected time of an event *given* it has not yet happened on or before a time c.

There are a few ways to answer the problem, as usual, we will present the most difficult first, and then the simple answer.

Tackling this problem head on, one recognises that an expected value for a continuous random variable can be obtained by integration.

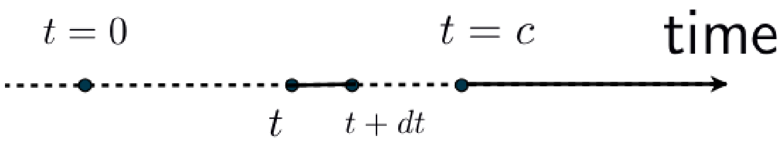

Figure 18.4: The time intervals $T > c$ and $t \leq T \leq t + dt$, when $t < c$, do *not* overlap.

$$\mathbb{E}\left(T|T>c\right) = \int_0^{+\infty} t f_{T|T>c}(t) dt \tag{18.28}$$

The problem is finding the conditional density: $f_{T|T>c}(t)$.

This can be derived as follows:

$$f_{T|T>c}(t)dt = P(t \leq T \leq t + dt|T > c) = \frac{P(t \leq T \leq t + dt \text{ and } T > c)}{P(T > c)} \quad . \tag{18.29}$$

But,

$$P(t \leq T \leq t + dt \text{ and } T > c) = 0 \quad \text{when} \quad t < c \quad . \tag{18.30}$$

that is, it is impossible for the event to occur after c and before c. There is no interaction between these time intervals; see figure 18.4. While, when there is overlap between the two intervals (figure 18.4), and this only arises when $t \geq c$ (figure **??**), and then:

$$P(t \leq T \leq t + dt \text{ and } T > c) = P(t \leq T \leq t + dt) = \lambda e^{-\lambda t} dt \quad \text{when} \quad t \geq c \quad . \tag{18.31}$$

and since, $P(T > c) = e^{-\lambda c}$, thus:

$$f_{T|T>c}(t) = \begin{cases} 0 & t < c \\ \lambda e^{-\lambda t} e^{\lambda c} & t \geq c \end{cases} \quad . \tag{18.32}$$

$$\mathbb{E}\left(T|T>c\right) = \int_c^{+\infty} t \lambda e^{-\lambda t} e^{\lambda c} dt \tag{18.33}$$

After a bit of algebra, integration by parts:

$$\mathbb{E}\left(T|T>c\right) = \frac{1}{\lambda} + c \quad . \tag{18.34}$$

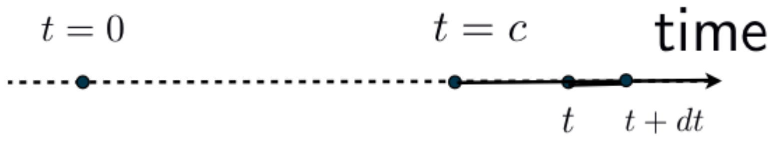

Figure 18.5: The time intervals $T > c$ and $t \leq T \leq t + dt$, when $t > c$, and thus the overlap/intersection is $t \leq T \leq t + dt$.

Now the short answer to the question: $\mathbb{E}\left(T|T > c\right) =?$. We are told the event has not happened by a time c and then we want to know how much longer it will take to occur. In the exponential process recall that it has no memory. So if nothing has happened up to a time c, the waiting starts again from time zero. We know that the average waiting time for an exponential process is $1/\lambda$, starting from time zero. Thus, we need to wait, on average, a further time $1/\lambda$. Then we have

$$\mathbb{E}\left(T|T > c\right) = c + \frac{1}{\lambda} \qquad , \tag{18.35}$$

which agrees perfectly with the answer obtained previously (18.34).