Lecture 14: The Poisson Arrival Process

Relevant textbook passages:
Pitman [13]: Sections 2.4,3.5, 4.2
Larsen–Marx [12]: Sections 4.2, 4.6

14.1 Waiting times and arrivals

In this lecture we discuss the following random experiment: waiting repeatedly in continuous real time for the occurrence of a phenomenon that happens at random times. Each occurrence is called an **arrival** or a **death** and the intervals between successive arrivals are called **waiting times**. Before we can proceed there are a number of preliminaries.

Consider the random experiment of waiting for something to die or fail. For instance, run a disk drive until it fails, or a human body till it fails, or wait until a uranium atom disintegrates. This gives rise to a random variable called a lifetime or duration or waiting time.

Pitman [13]: § 4.2

For a **lifetime** or **duration** T with cumulative distribution function $F(t)$, define the **survival function** by

$$G(t) = P(T > t) = 1 - F(t).$$

When T is the (random) time to failure, the survival function $G(t)$ at epoch t gives the probability of surviving (not failing) until at least t.

Note the convention that the present is time $t = 0$, and durations are measured as times after that.

Aside: If you've ever done any programming involving a calendar, you know the difference between a *point in time* or *date*, called an **epoch** by probabilists, and a *duration*, which is the elapsed time between two epochs.

Pitman [13]: § 4.3

For a lifetime T with a density f on $[0, \infty)$ and cumulative distribution function F, the **hazard rate** $\lambda(t)$ is defined by

$$\lambda(t) = \lim_{h \downarrow 0} \frac{P\big(T \in (t, t+h] \bigm| T > t\big)}{h}.$$

Or

$$\lambda(t) = \frac{f(t)}{G(t)}.$$

Proof: By definition,

$$P\big(T \in (t, t+h] \bigm| T > t\big) = \frac{P(T \in (t, t+h))}{P(T > t)} = \frac{P(T \in (t, t+h])}{G(t)}.$$

Moreover $P(T \in (t, t+h]) = F(t+h) - F(t)$, so the limit is just $F'(t)/G(t) = f(t)/G(t)$. ∎

The hazard rate $f(t)/G(t)$ is often thought of as the "instantaneous" probability of death or failure.

14.2 The Exponential Distribution

The **Exponential(λ)** is widely used to model random durations or times. It is another name for the Gamma$(1, \lambda)$ distribution. (See Section 14.12 below.)

14.2.1 Definition *A random time T has an Exponential(λ) distribution if it has density*

$$f(t) = \lambda e^{-\lambda t} \qquad (t \geqslant 0),$$

and cdf

$$F(t) = 1 - e^{-\lambda t},$$

which gives survival function

$$G(t) = e^{-\lambda t},$$

and hazard rate

$$\lambda(t) = \lambda.$$

See Figure 14.1.

It is the **constant hazard rate** that makes the exponential family a good model for certain kinds of waiting or arrival times.

> The only distribution with a constant hazard rate $\lambda > 0$ is the Exponential(λ) distribution.

To see this, look at the survival function G. If the hazard rate is constant at λ, then G satisfies the differential equation $G'(t)/G(t) = -\lambda$, so $G(t) = ce^{-\lambda t}$ for some constant c, and $c = 1$ is the only constant that satisfies $G(0) = 1$.

> The mean of an Exponential(λ) random variable is given by
>
> $$\int_0^\infty \lambda t e^{-\lambda t}\, dt = \frac{1}{\lambda}.$$

Proof: Use the integration by parts formula:

$$\int h'g = hg - \int g'h,$$

with $h'(t) = \lambda e^{-\lambda t}$ and $g(t) = t$ (so that $h(t) = -e^{-\lambda t}$ and $g'(t) = 1$) to get

$$\begin{aligned}
\boldsymbol{E}\, T &= \int_0^\infty \lambda t e^{-\lambda t}\, dt \\
&= -te^{-\lambda t}\Big|_0^\infty + \int_0^\infty e^{-\lambda t}\, dt \\
&= -te^{-\lambda t}\Big|_0^\infty + \frac{-1}{\lambda} e^{-\lambda t}\Big|_0^\infty \\
&= \frac{-e^{-\lambda t}}{\lambda}\Big|_0^\infty \\
&= \frac{1}{\lambda}.
\end{aligned}$$

∎

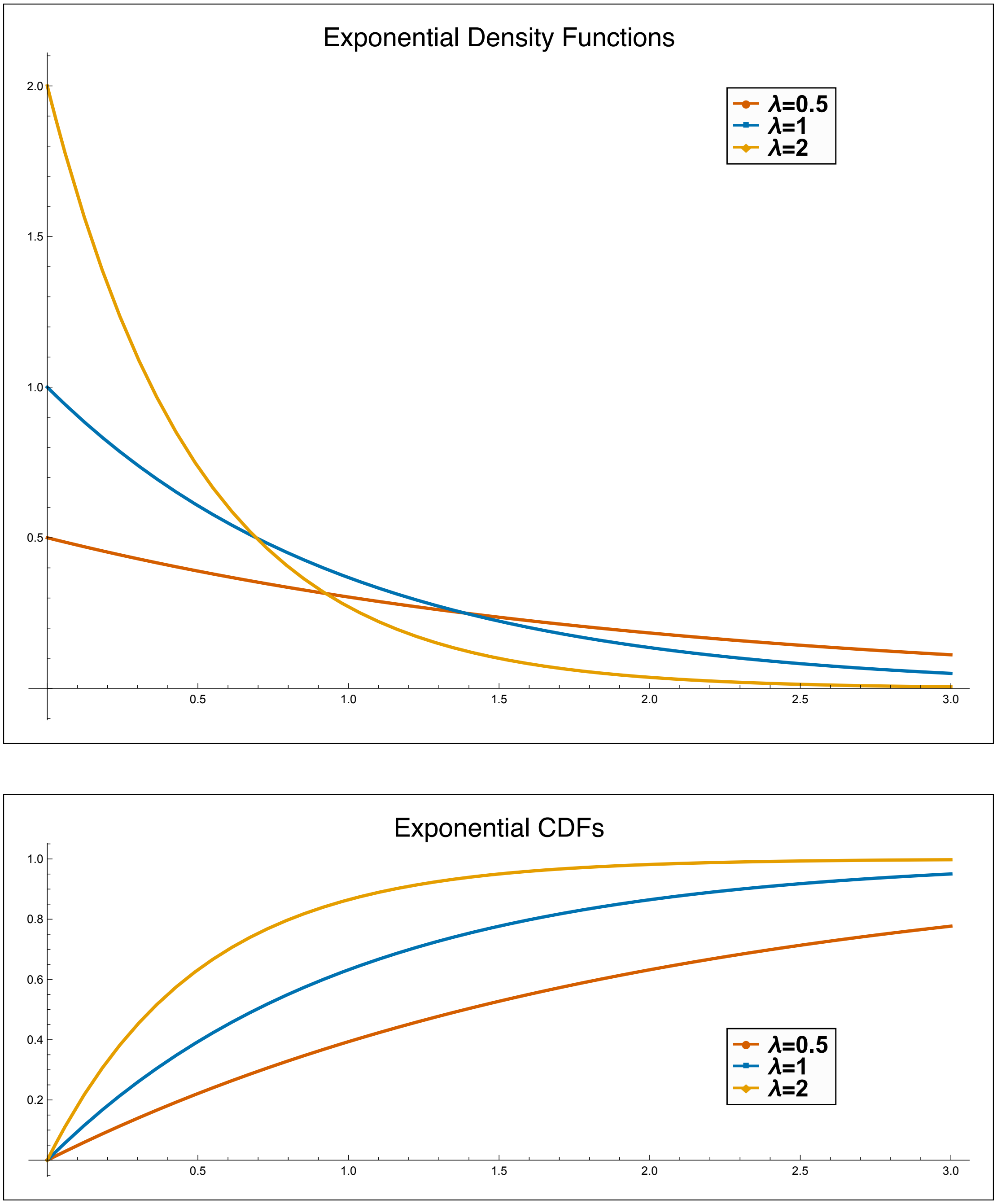

Exponential Density Functions
λ=0.5
λ=1
λ=2
2.0
1.5
1.0
0.5
0.5
1.0
1.5
2.0
2.5
3.0
Exponential CDFs
1.0
0.8
0.6
0.4
0.2
λ=0.5
λ=1
λ=2

Figure 14.1.

The variance of an Exponential(λ) is $\frac{1}{\lambda^2}$.

Proof:

$$\begin{aligned}\boldsymbol{Var} T &= \boldsymbol{E}(T^2) - (\boldsymbol{E}\, T)^2 \\ &= \int_0^\infty t^2 \lambda e^{-\lambda t}\, dt - \frac{1}{\lambda^2}.\end{aligned}$$

Setting $h'(t) = \lambda e^{-\lambda t}$ and $g(t) = t^2$ and integrating by parts, we get

$$\begin{aligned} &= \underbrace{t^2 e^{-\lambda t}\Big|_0^\infty}_{=0} + 2 \underbrace{\int_0^\infty t e^{-\lambda t}\, dt}_{=\boldsymbol{E}\, T/\lambda} - \frac{1}{\lambda^2} \\ &= 0 + \frac{2}{\lambda^2} - \frac{1}{\lambda^2} \\ &= \frac{1}{\lambda^2}. \end{aligned}$$

■

14.3 The Exponential is Memoryless

Pitman [13]: p. 279

A property that is closely related to having a constant hazard rate is that the exponential distribution is **memoryless** in that for an Exponential random variable T,

$$P\big(T > t + s \bigm| T > t\big) = P(T > s), \qquad (s > 0).$$

To see this, recall that by definition,

$$\begin{aligned} P\big(T > t + s \bigm| T > t\big) &= \frac{P\left((T > t + s)(T > t)\right)}{P(T > t)} \\ &= \frac{P(T > t + s)}{P(T > t)} \qquad \text{as } (T > t + s) \subset (T > t) \\ &= \frac{G(t + s)}{G(t)} \\ &= \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} \\ &= e^{-\lambda s} \\ &= G(s) = P(T > s). \end{aligned}$$

In fact, *the only* continuous memoryless distributions are Exponential.

Proof: Rewrite memorylessness as

$$\frac{G(t + s)}{G(t)} = G(s),$$

or

$$G(t + s) = G(t)G(s) \qquad (t, s > 0).$$

It is well known that this last property (plus the assumption of continuity at one point) is enough to prove that G must be an exponential (or identically zero) on the interval $(0, \infty)$. See J. Aczél [1, Theorem 1, p. 30].[1] ∎

14.4 Comparison of independent exponential random variables

Let $X \sim \text{Exponential}(\lambda)$ and $Y \sim \text{Exponential}(\mu)$ be independent. What is

$$P(X < Y)?$$

Pitman [13]: p. 352

The joint density is

$$f(x,y) = \lambda e^{-\lambda x} \mu e^{-\mu y} = \lambda\mu e^{-\lambda x - \mu y},$$

so

$$\begin{aligned}
P(X < Y) &= \int_0^\infty \int_0^y \lambda e^{-\lambda x} \mu e^{-\mu y} \, dx \, dy \\
&= \int_0^\infty \mu e^{-\mu y} \left(\int_0^y \lambda e^{-\lambda x} \, dx \right) dy \\
&= \int_0^\infty \mu e^{-\mu y} \left(-e^{-\lambda x} \Big|_0^y \right) dy \\
&= \int_0^\infty \mu e^{-\mu y} \left(1 - e^{-\lambda y} \right) dy \\
&= \underbrace{\int_0^\infty \mu e^{-\mu y} \, dy}_{=1} - \int_0^\infty \mu e^{-(\lambda+\mu) y} \, dy \\
&= 1 - \frac{\mu}{\lambda+\mu} \underbrace{\int_0^\infty (\lambda+\mu) e^{-(\lambda+\mu) y} \, dy}_{=1} \\
&= 1 - \frac{\mu}{\lambda+\mu} \\
&= \frac{\lambda}{\lambda+\mu}.
\end{aligned}$$

14.5 The sum of i.i.d. exponential random variables

Pitman [13]: pp. 373—375

Let X and Y be independent and identically distributed $\text{Exponential}(\lambda)$ random variables. The density of the sum for $t > 0$ is given by the convolution:

$$\begin{aligned}
f_{X+Y}(t) &= \int_0^\infty f_X(t-y) f_Y(y) \, dy \\
&= \int_0^t \lambda e^{-\lambda(t-y)} \lambda e^{-\lambda y} \, dy \qquad \text{since } f_Y(t-y) = 0 \text{ if } y > t \\
&= \int_0^t \lambda^2 e^{-\lambda t} \, dy \\
&= t\lambda^2 e^{-\lambda t}.
\end{aligned}$$

This is a $\text{Gamma}(2, \lambda)$ distribution. See Section 14.12 below.

[1] Aczél [1] points out that there is another kind of solution to the functional equation when we extend the domain to $[0, \infty)$, namely $G(0) = 1$ and $G(t) = 0$ for $t > 0$.

More generally, the sum of n independent and identically distributed Exponential(λ) random variables has a **Gamma(n, λ) distribution**, given by

$$f(t) = \lambda^n e^{-\lambda t} \frac{t^{n-1}}{(n-1)!}.$$

14.6 The Weibull distribution

The **Weibull distribution** generalizes the exponential distribution as a model of waiting times. It allows for the hazard rate to be age-dependent. There are two parameters of the Weibull distribution, usually denoted λ and α. α is a *shape* parameter and λ is a transformed *scale* parameter. The density is given by

$$f(t) = \alpha\lambda t^{\alpha-1} e^{-\lambda t^\alpha}$$

and the cumulative distribution function is

$$F(t) = 1 - e^{-\lambda t^\alpha},$$

so the survival function is

$$G(t) = e^{-\lambda t^\alpha}$$

and the hazard rate is

$$\lambda(t) = \frac{f(t)}{G(t)} = \alpha\lambda t^{\alpha-1}.$$

Note that when $\alpha = 1$, this reduces to the Exponential(λ) distribution. When $\alpha > 1$, the hazard rate is increasing with age, and when $\alpha < 1$, the hazard rate is decreasing with age. See Figure 14.2. N.B. Some authors refer to the parameters of the Weibull as α and β, where α is the same as I use, but β is a scale parameter related to λ. Mathematica, R, and Jacod and Protter [11, Example 6, p. 43] use β to mean my $1/\lambda^{1/\alpha}$. I am using the terminology used by Pitman [13, Exercise 4, p. 301], and Forbes, et al. [9, p. 193].

If T has a Weibull(λ, α) distribution its moments are given by

$$\boldsymbol{E}\, T^n = \Gamma\left(1 + \frac{n}{\alpha}\right) \lambda^{-n/\alpha}.$$

14.7 Survival functions and moments

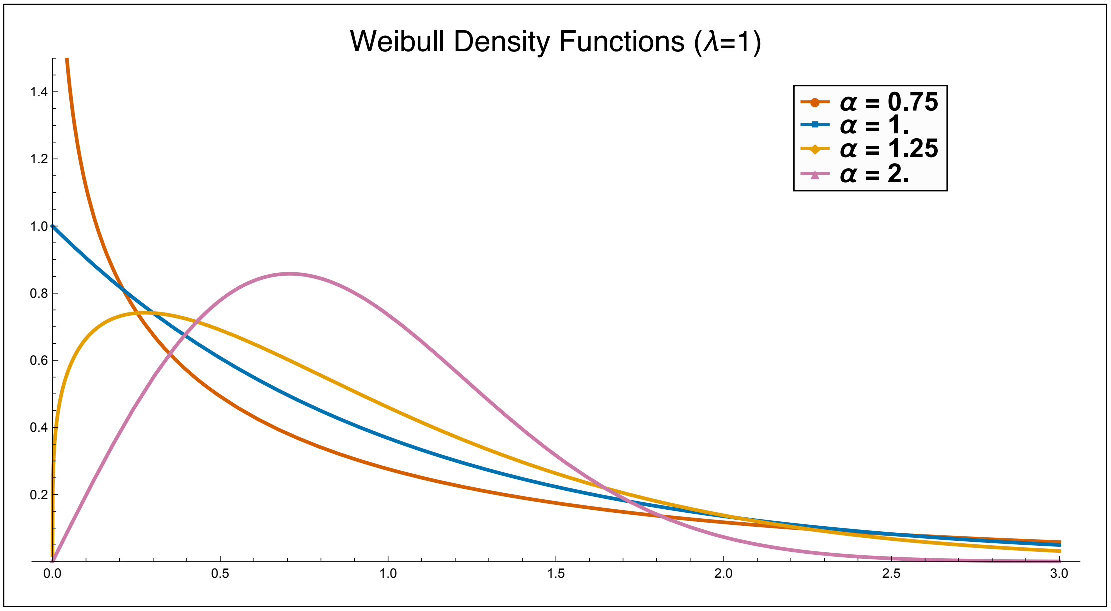
For a nonnegative random variable with a continuous density f, integration by parts allows us to prove the followin generalization of Proposition 6.4.2.

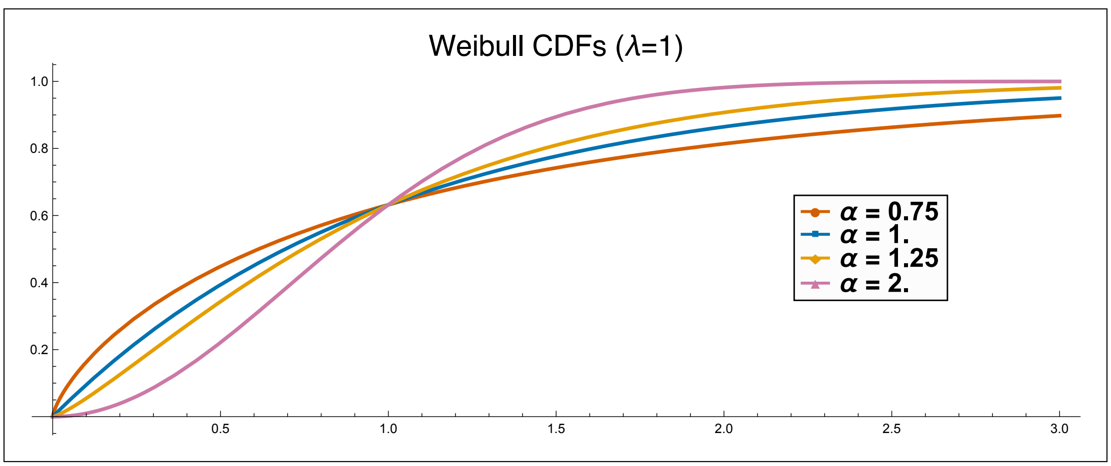
14.7.1 Proposition *Let F be a cumulative distribution function with continuous density f on $[0, \infty)$. Then the n^{th} moment can be calculated as*

$$\int_0^\infty x^n f(x)\,dx = \int_0^\infty n x^{n-1}\big(1 - F(x)\big)\,dx = \int_0^\infty n x^{n-1} G(x)\,dx.$$

Proof: Use the integration by parts formula:

$$\int h'g = hg - \int g'h,$$


Weibull Density Functions (λ=1)
α = 0.75
α = 1.
α = 1.25
α = 2.
0.2
0.4
0.6
0.8
1.0
1.2
1.4
0.0
0.5
1.0
1.5
2.0
2.5
3.0


Weibull CDFs (λ=1)
α = 0.75
α = 1.
α = 1.25
α = 2.
0.2
0.4
0.6
0.8
1.0
0.5
1.0
1.5
2.0
2.5
3.0

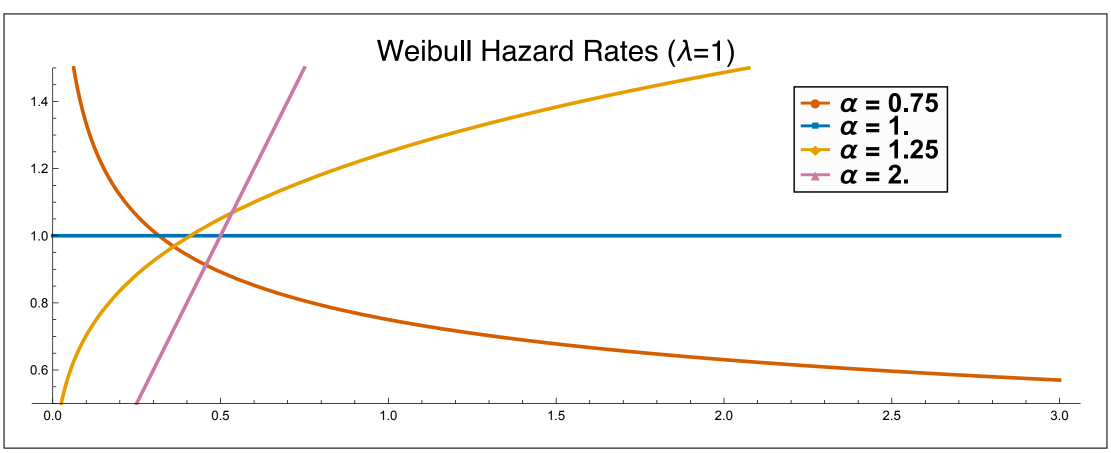

Weibull Hazard Rates (λ=1)
α = 0.75
α = 1.
α = 1.25
α = 2.
0.6
0.8
1.0
1.2
1.4
0.0
0.5
1.0
1.5
2.0
2.5
3.0

Figure 14.2.

with $h'(x) = f(x)$ and $g(x) = x^n$ (so that $h(x) = F(x)$ and $g'(x) = nx^{n-1}$) to get

$$\begin{aligned}\int_0^b x^n f(x)\,dx &= x^n F(x)\Big|_0^b - \int_0^b nx^{n-1}F(x)\,dx \\ &= b^n F(b) - \int_0^b nx^{n-1}F(x)\,dx \\ &= F(b)\int_0^b nx^{n-1}\,dx - \int_0^b nx^{n-1}F(x)\,dx \\ &= \int_0^b nx^{n-1}\big(F(b) - F(x)\big)\,dx,\end{aligned}$$

and let $b \to \infty$. ∎

In particular, the first moment, the mean, is given by the area under the survival function:

$$\boldsymbol{E} = \int_0^\infty \big(1 - F(x)\big)\,dx = \int_0^\infty G(x)\,dx.$$

14.8 The Poisson Arrival Process

Pitman [13]: § 4.2; and pp. 283–285

The “Poisson arrival process” is a mathematical model that is useful in modeling the number of occurrences (called **arrivals**) of a phenomenon over a continuous time period. For instance the number of telephone calls per minute, the number of Google queries in a second, the number of radioactive decays in a minute, the number of earthquakes per year, etc. In these phenomena, the events are rare enough to be counted, and to have measurable delays between them.[2] (Interestingly, the Poisson model is not a good description of LAN traffic, see [3, 14].)

The Poisson arrival process with parameter λ works like this:

> Let $W_1, W_2, \ldots$ be a sequence of independent and identically distributed Exponential(λ) random variables, representing **waiting times** for an **arrival**, on the sample space $(\Omega, \mathcal{F}, P)$. At each $\omega \in \Omega$, the first arrival happens at time $W_1(\omega)$, the second arrival happens a duration $W_2(\omega)$ later, at $W_1(\omega) + W_2(\omega)$. The third arrival happens at $W_1(\omega) + W_2(\omega) + W_3(\omega)$. Define
>
> $$T_n = W_1 + W_2 + \cdots + W_n.$$
>
> This is the epoch when the n^{th} event occurs. The sequence T_n of random variables is a nondecreasing sequence.

For each ω we can associate a step function of time, $N(t)$ defined by

> $$\begin{aligned}N(t) &= \text{the number of arrivals that have occurred at a time} \leqslant t \\ &= \text{the number of indices } n \text{ such that } T_n \leqslant t.\end{aligned}$$

14.8.1 Remark Since the function N depends on ω, I should probably write

$$N(t, \omega) = \text{the number of indices } n \text{ such that } T_n(\omega) \leqslant t.$$

But that is not traditional. Something a little better than no mention of ω that you can find, say in Doob’s book [7] is a notation like $N_t(\omega)$. But most of the time we want to think of N as a random function of time, and putting t in the subscript disguises this.

[2] An exception is the 1952 Kern County earthquake cluster, where the number and size of aftershocks overwhelmed the recording equipment, so an accurate count of aftershocks could not be made [10, pp. 437, 439].

14.8.2 Definition *The random function N is called the **Poisson process** with parameter λ.*

So why is this called a Poisson Process? Because $N(t)$ has a Poisson(λt) distribution. There is nothing special about starting at time $t = 0$. The Poisson process looks the same over every time interval.

The Poisson process has the property that for any interval of length t, the distribution of the number of "arrivals" is Poisson(λt).

14.9 The Poisson Arrival Process is also a Poisson Scatter

Another way to describe the Poisson arrival process is as a one-dimensional Poisson scatter of arrivals over a time interval. Here is an argument to convince you of this, but it has a gap. In a one-dimensional Poisson scatter, the number of hits/arrivals in an interval of length t is a Poisson random variable with parameter λt, and the random counts are stochastically independent for disjoint intervals. We can use this to calculate the distribution of waiting times between hits. If the n^{th} hit occurs at time w, the probability that no hit occurs in the interval $(w, w+t]$ is a Poisson probability with rate λt, namely

$$p_{\lambda t}(0) = e^{-\lambda t}\frac{(\lambda t)^0}{0!} = e^{-\lambda t}.$$

Let W_{n+1} be the $n+1^{\text{st}}$ waiting time. Then $W_{n+1} > t$ if and only if no hit occurs in the interval $(w, w+t]$. We have just seen that the probability of this is

$$P(W_{n+1} > t) = e^{-\lambda t}.$$

But this is just the survival function for an Exponential(λ) distribution. In other words,

A one-dimensional Poisson scatter with intensity λ generates waiting times that are independent with Exponential(λ) distributions. That is, it generates a Poisson arrival process.

Aside: So where is the gap in the argument? I implicitly used something called the Strong Markov Property (which will be discussed in Section 17.7⋆) without justifying it. But you probably didn't notice.

14.10⋆ Renewal processes

****

To be added.

A renewal process is like a Poisson arrival process where the waiting times need not be exponential.

****

14.11⋆ The Inspection Paradox

****

To be added.

If machines are replaced with new one when the old ones fail (a renewal process), then if you select a time at random to inspect a machine, the expected lifetime of the machine you are examining is greater than the expected lifetime of a machine. The intuition is that lifetimes divide the interval from $[0, T]$ into subintervals. If you choose a time at random in $[0, T]$, it is more likely to land in long interval than a short interval. There is some math that goes along with this intuition.

****

14.12 Appendix: The Gamma family of distributions

Pitman [13]: Exercise 4.2.12, p. 294; p. 481

14.12.1 Definition *For* $r, \lambda > 0$, ***Gamma(r, λ) distribution*** *has a density given by*

$$f(t) = \frac{\lambda^r}{\Gamma(r)}\, t^{r-1} e^{-\lambda t} \qquad (t > 0), \tag{1}$$

where

$$\Gamma(r) = \int_0^\infty z^{r-1} e^{-z}\, dz$$

is the Gamma function. (See Appendix 14.13.)

To very that this is indeed a density with support $[0, \infty)$, use the change of variable $z = \lambda t$ to see that $\int_0^\infty \lambda(\lambda t)^{r-1} e^{-\lambda t}\, dt = \int_0^\infty z^{r-1} e^{-z}\, dz = \Gamma(r)$.

The parameter r is referred to as the **shape parameter** or **index** and λ is a **scale parameter**.

Why is it called the scale parameter?

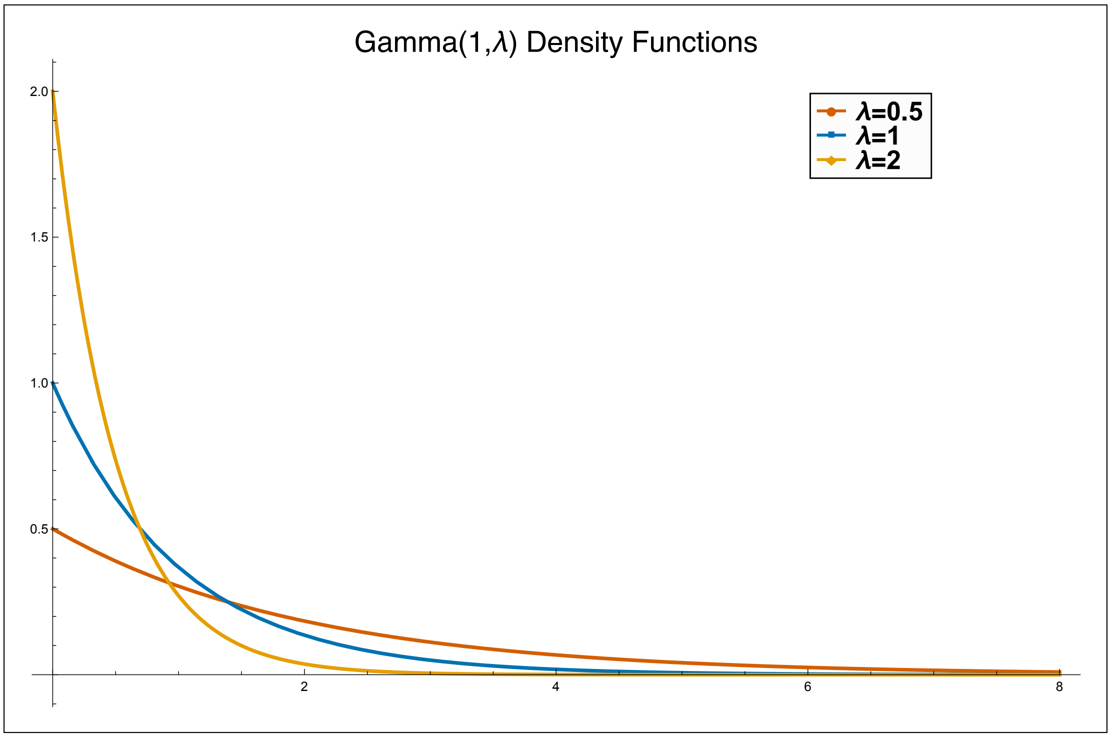
$$T \sim \text{Gamma}(r, \lambda) \iff \lambda T \sim \text{Gamma}(r, 1)$$

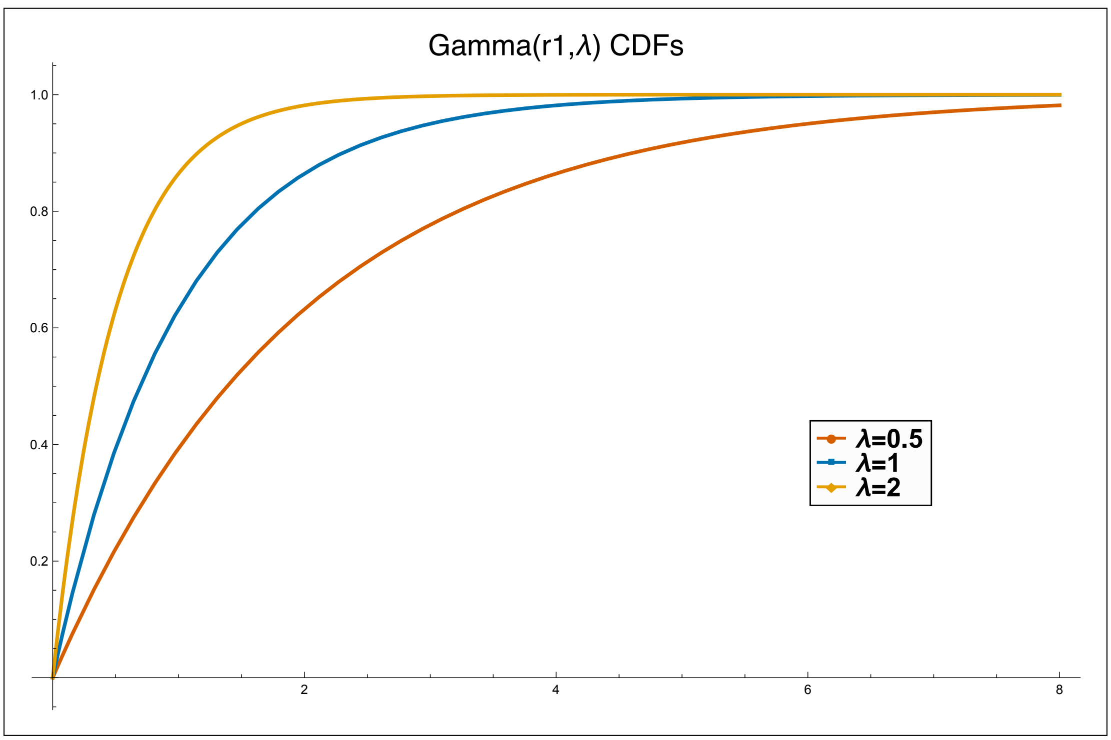
Wait. Shouldn't we call $1/\lambda$ the scale parameter on this basis? Probably, and some people do (see the sidebar on page 14–14). But Pitman [13] and Larsen–Marx [12] are not among them.

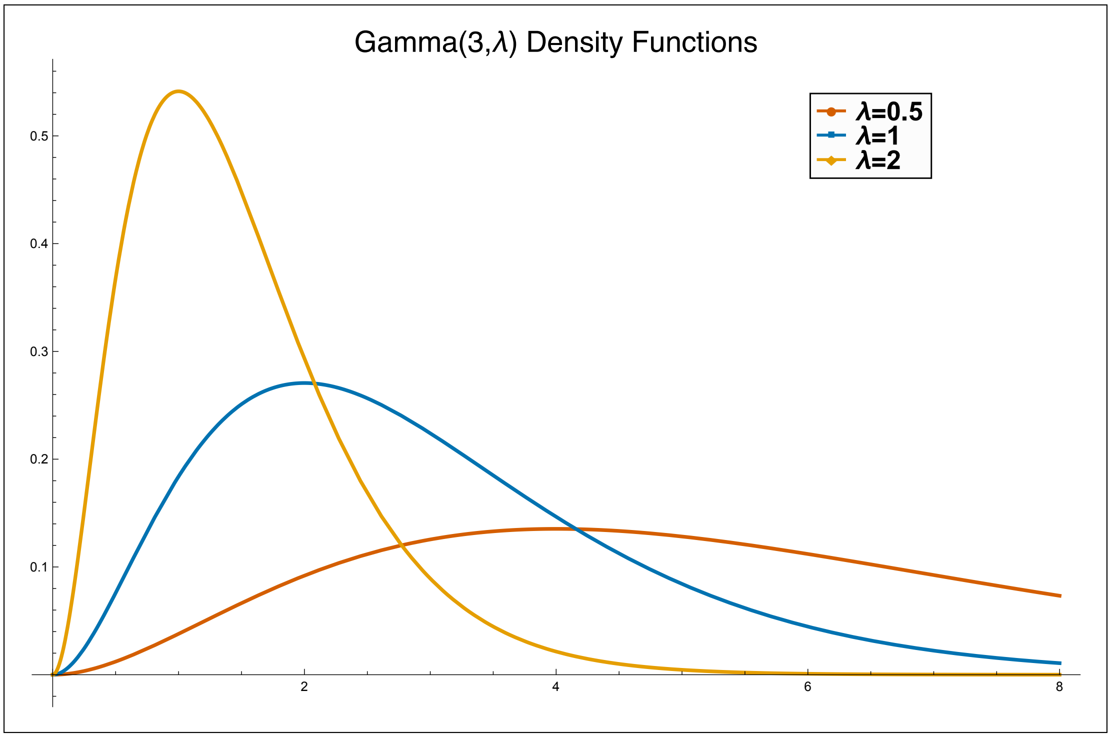
The Gamma(r, λ) distribution has mean and variance given by

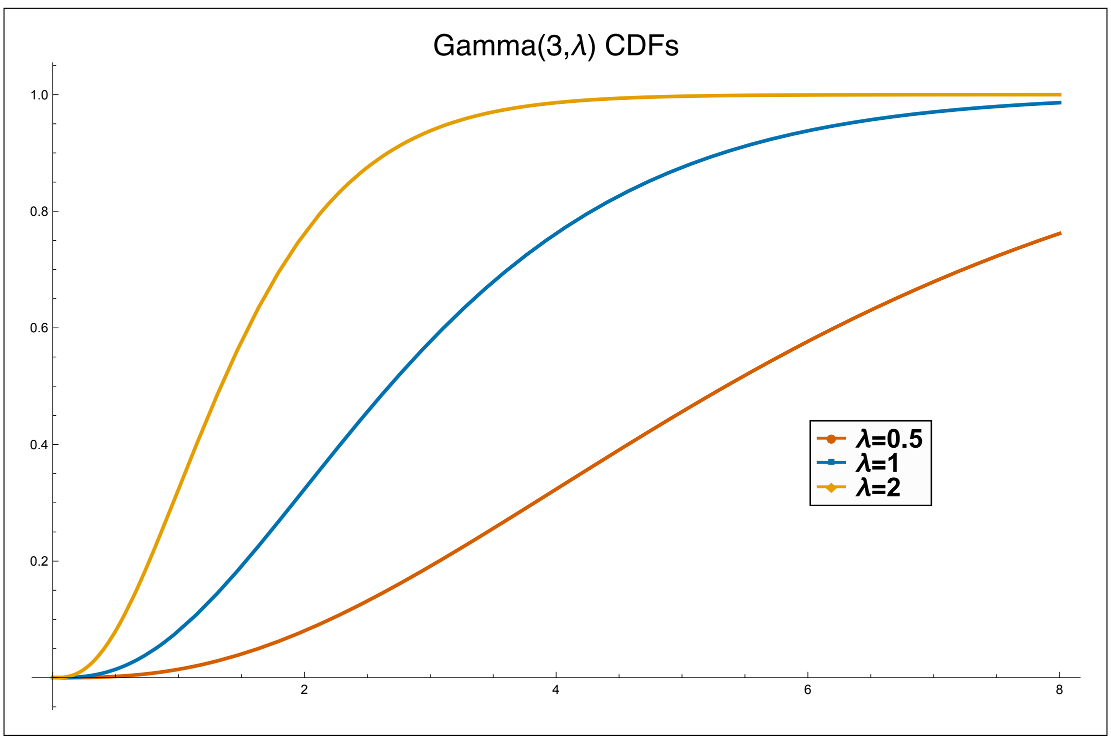
$$\boldsymbol{E}\, X = \frac{r}{\lambda}, \qquad \boldsymbol{Var}\, X = \frac{r}{\lambda^2}.$$

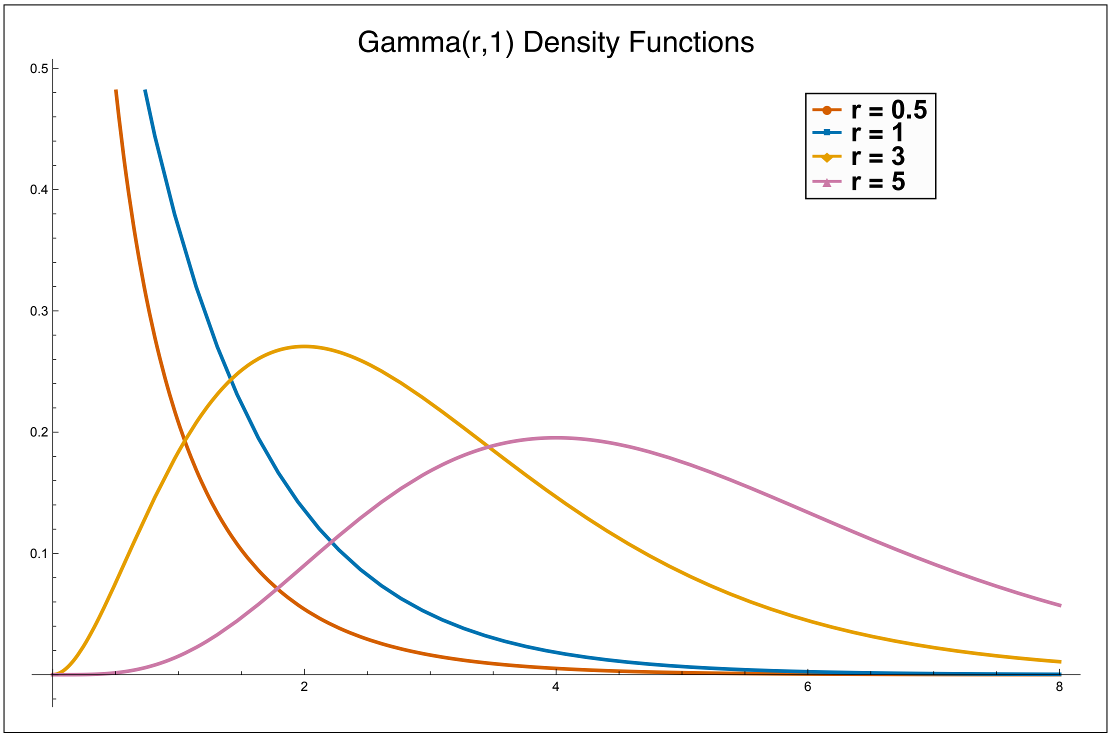
According to Pitman [13, p. 291], "In applications, the distribution of a random variable may be unknown, but reasonably well approximated by some gamma distribution."

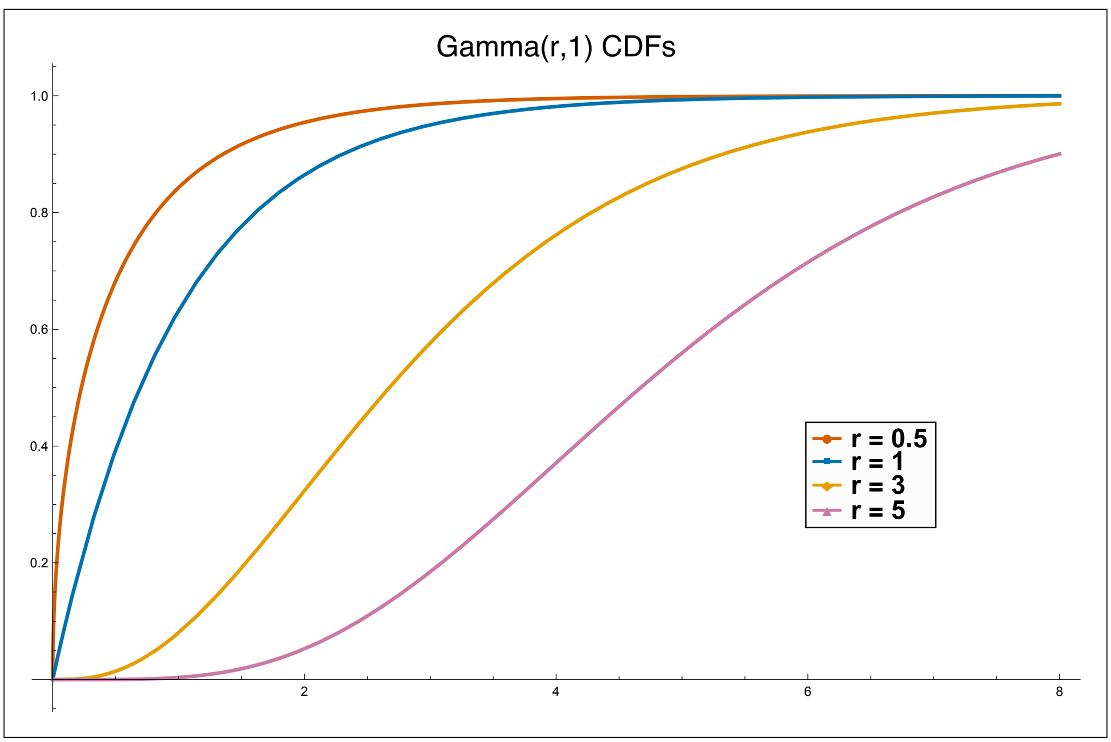

Gamma(1,λ) Density Functions
λ=0.5
λ=1
λ=2
2.0
1.5
1.0
0.5
2
4
6
8


Gamma(r1,λ) CDFs
λ=0.5
λ=1
λ=2
1.0
0.8
0.6
0.4
0.2
2
4
6
8


Gamma(3,λ) Density Functions
λ=0.5
λ=1
λ=2
0.5
0.4
0.3
0.2
0.1
2
4
6
8


Gamma(3,λ) CDFs
1.0
0.8
0.6
0.4
0.2
λ=0.5
λ=1
λ=2
2
4
6
8


Gamma(r,1) Density Functions
r = 0.5
r = 1
r = 3
r = 5
0.5
0.4
0.3
0.2
0.1
2
4
6
8


Gamma(r,1) CDFs
r = 0.5
r = 1
r = 3
r = 5
1.0
0.8
0.6
0.4
0.2
2
4
6
8

Hey: Read Me

There are (at least) three incompatible, but easy to translate, naming conventions for the Gamma distribution.

Pitman [13, p. 286] and Larsen and Marx [12, Defn. 4.6.2, p. 272] refer to their parameters as r and λ, and call the function in equation (1) the Gamma(r, λ) density. Note that the shape parameter is the first parameter and the scale parameter is the second parameter for Pitman and Larsen and Marx. This is the convention that I used above in equation (1).

Feller [p. 47][8] calls the scale parameter α instead of λ, and he calls the shape parameter ν instead of r. Cramér [p. 126][5] also calls the scale parameter α instead of λ, but the shape parameter he calls λ instead of r. Other than that they agree that equation (1) is the Gamma density, but they list the parameters in reverse order. That is, they list the scale parameter first, and the shape parameter second.

Casella and Berger [4, eq. 3.3.6, p. 99] call the scale parameter β and the shape parameter α, and list the shape parameter first and the scale parameter second. But here is the confusing part, their scale parameter β is our $1/\lambda$. [a] Mathematica and R also invert the scale parameter. To get my Gamma(r, λ) density in Mathematica, you have to call `PDF[GammaDistribution[r, 1/λ], t]`; to get it in R, you would call `dgamma(t, r, rate = 1/λ)`.

I feel sorry for you, but it's not my fault. But you do have to be careful to know what naming convention is being used.

[a] That is, C–B write the Gamma(α, β) density as

$$\frac{1}{\Gamma(\alpha)\beta^\alpha} t^{\alpha-1} e^{-t/\beta}.$$

14.13 Appendix: The Gamma Function

Larsen–Marx [12]: Section 4.6, pp. 270–274.
Pitman [13]: p. 291

14.13.1 Definition *The **Gamma function** is defined by*

$$\Gamma(t) = \int_0^\infty z^{t-1} e^{-z}\, dz \qquad \text{for } t > 0.$$

Clearly $\Gamma(t) > 0$ *for* $t > 0$.

See Figure 14.3.

The value at $t = 1$ is relatively easy to compute:

$$\Gamma(1) = \int_0^\infty e^{-z}\, dz = \Big\|_0^1 - e^{-a} = 1.$$

The next result enables us to use this to recursively compute $\Gamma(n)$ for any natural number n.

14.13.2 Proposition *The Gamma function is a continuous version of the factorial, and has the property that*

$$\Gamma(t+1) = t\Gamma(t) \qquad (t > 0).$$

Proof: Let $v(z) = z^t$ and $u(z) = -e^{-z}$. Then integrating by parts yields

$$\Gamma(t) = \int_0^\infty v(z)u'(z)\, dz = uv\Big|_0^\infty - \int_0^\infty u(z)v'(z)\, dz$$
$$= (0-0) + \int_0^\infty t z^{t-1} e^{-z}\, dz = t \int_0^\infty z^{t-1} e^{-z}\, dz = t\Gamma(t-1).$$

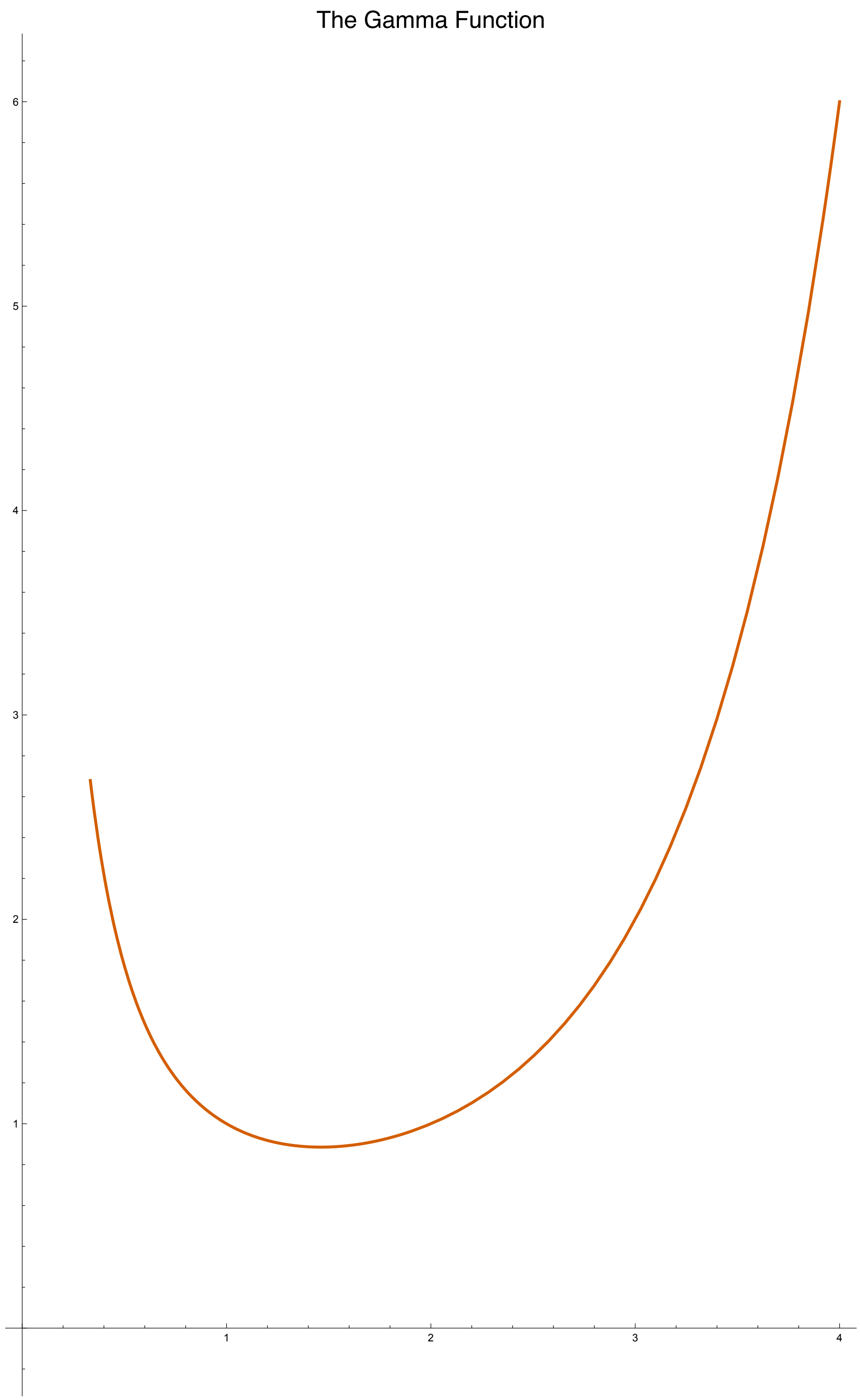


Figure 14.3. The Gamma function.

■

Consequently,

$$\Gamma(2) = \Gamma(1) = 1,$$

and for every natural number m,

$$\Gamma(m) = (m-1)!$$

In light of this, it would seem to make more sense to define a function $f(t) = \int_0^\infty z^t e^{-z}\,dz$, so that $f(m) = m!$. According to Davis [6, fn., p 855] the current definition was formulated by Legendre, while Gauss advocated the alternative. Was this another example of VHS vs. Betamax? (Do you even know what that refers to?)

We also have

$$\Gamma(1/2) = \sqrt{\pi}.$$

To see this, note that

$$\Gamma(1/2) = \int_0^\infty z^{-1/2} e^{-z}\,dz.$$

Make the change of variable $u(z) = \sqrt{2z}$, so $u'(z) = 1/\sqrt{2z}$, and note that

$$\int_0^\infty z^{-1/2} e^{-z}\,dz = \frac{1}{\sqrt{2}} \int_0^\infty e^{-u^2/2}\,du = \sqrt{\pi},$$

where the last equality follows from the normal density integral, see Proposition 10.9.1.

There is no closed form formula for the Gamma function except at integer multiples of 1/2. See, for instance, Pitman [13, pp. 290–291] or Apostol [2, pp. 419–421] for more.

Bibliography

[1] J. D. Aczél. 2006. *Lectures on functional equations and their applications*. Mineola, NY: Dover. Reprint of the 1966 edition originally published by Academic Press. An Errata and Corrigenda list has been added. It was originally published under the title *Vorlesungen über Funktionalgleichungen and ihre Anwendungen*, published by Birkhäuser Verlag, Basel, 1961.

[2] T. M. Apostol. 1967. *Calculus*, 2d. ed., volume 1. Waltham, Massachusetts: Blaisdell.

[3] J. Beran, R. P. Sherman, M. S. Taqqu, and W. Willinger. 1995. Variable-bit-rate video traffic and long-range dependence. *IEEE Transactions on Communications* 43(2/3/4):1566–1579. DOI: 10.1109/26.380206

[4] G. Casella and R. L. Berger. 2002. *Statistical inference*, 2d. ed. Belmont, California: Brooks/Cole Cengage Learning.

[5] H. Cramér. 1946. *Mathematical methods of statistics*. Number 34 in Princeton Mathematical Series. Princeton, New Jersey: Princeton University Press. Reprinted 1974.

[6] P. J. Davis. 1959. Leonhard Euler's integral: A historical profile of the Gamma function: In memoriam: Milton Abramowitz. *American Mathematical Monthly* 66(10):849–869. http://www.jstor.org/stable/2309786

[7] J. L. Doob. 1953. *Stochastic processes*. New York: Wiley.

[8] W. Feller. 1971. *An introduction to probability theory and its applications*, 2d. ed., volume 2. New York: Wiley.

[9] C. Forbes, M. Evans, N. Hastings, and B. Peacock. 2011. *Statistical distributions*, 4th. ed. Hoboken, New Jersey: John Wiley & Sons.

[10] K. Hutton, J. Woessner, and E. Hauksson. 2010. Earthquake monitoring in Southern California for seventy-seven years (1932–2008). *Bulletin of the Seismological Society of America* 100(2):423–446. DOI: 10.1785/0120090130

[11] J. Jacod and P. Protter. 2004. *Probability essentials*, 2d. ed. Berlin and Heidelberg: Springer.

[12] R. J. Larsen and M. L. Marx. 2012. *An introduction to mathematical statistics and its applications*, fifth ed. Boston: Prentice Hall.

[13] J. Pitman. 1993. *Probability*. Springer Texts in Statistics. New York, Berlin, and Heidelberg: Springer.

[14] W. Willinger, M. S. Taqqu, R. P. Sherman, and D. V. Wilson. 1997. Self-similarity through high variability: Statistical analysis of ethernet LAN traffic at the source level (extended version). *IEEE/ACM Transactions on Networking* 5(1):71–86. DOI: 10.1109/90.554723