Chapter 1

Appendices

1.1 Probability Basis

1.1.1 Random Variables

Extreme Value (EV) distributions

Let $\boldsymbol{X} = \left[X_1, X_2, \cdots, X_n\right]^T$ denote a random vector, with its smallest value denoted by $X_{\min}$ and its largest value denoted by $X_{\max}$. If the distribution of $\boldsymbol{X}$ is known, one can also find the distirbution of $X_{\min}$ and $X_{\max}$.

For the special case when X_is are iid, one can obtain simple expressions for the distribution of $X_{\min}$ and $X_{\max}$.

Let $f_X(x)$ and $F_X(x)$ denote the PDF and CDF of the X_is, the minimum value is larger than x means than $\{X_{\min} > x\} = \{X_1 > x \cap X_2 > x \cap \cdots \cap X_n > x\}$ To derive the CDF of $X_{\min}$

$$
\begin{aligned}
F_{X_{\min}}(x) &= \Pr(X_{\min} \leqslant x) && (1.1)\\
&= 1 - \Pr(X_{\min} > x) && (1.2)\\
&= 1 - \Pr(X_1 > x \cap X_2 > x \cap \cdots \cap X_n > x) && (1.3)\\
&= 1 - \prod_{i=1}^{n} \Pr(X_i > x) && (1.4)\\
&= 1 - [1 - F_X(x)]^n && (1.5)
\end{aligned}
$$

The PDF of $X_{\min}$ is

$$
\begin{aligned}
f_{X_{\min}}(x) &= \frac{dF_{X_{\min}}(x)}{dx} && (1.6)\\
&= n f_X(x)[1 - F_X(x)]^{n-1} && (1.7)
\end{aligned}
$$

To derive the CDF of $X_{\max}$, we could start from the fact that $\{X_{\max} \leqslant x\} = \{X_1 \leqslant x \cap \cdots \cap X_n \leqslant x\}$:

$$\begin{aligned}
F_{X_{\max}}(x) &= \Pr(X_{\max} \leqslant x) && (1.8)\\
&= \Pr(X_1 \leqslant x \cap \cdots \cap X_n \leqslant x) && (1.9)\\
&= \prod_{i=1}^{n} \Pr(X_i \leqslant x) && (1.10)\\
&= [F_X(x)]^n && (1.11)
\end{aligned}$$

The PDF of $X_{\max}$ is

$$f_{X_{\max}}(x) = \frac{dF_X(x)}{dx} = n f_X(x)[F_X(x)]^{n-1} \tag{1.12}$$

The results can be generalized to the k^{th} largest value in $\boldsymbol{X}$ denoted by Y_k: the CDF of Y_k is

$$F_{Y_k}(x) = \sum_{j=k}^{n} C_n^j (F_X(x))^j (1 - F_X(x))^{n-j} \tag{1.13}$$

The underlying X_is are often unknown, and the above solutions are therefore not directly applicable. However, extreme value distributions that arise as asymptotic solutions of the above for $n \to \infty$ provide a useful model in these cases.

It can be shown that the solutions given above in Eqs Eq. 1.1, Eq. 1.6, Eq. 1.8, Eq. 1.12 converge to one of three distribution types, depending on the distribution of the X_is.

The type I EV distribution for maxima is named Gumbel distribution.

1.1.2 Random Process

Bernoulli Process

A Bernoulli process is a finite or infinite seqeunce of independent random variables $X_1, X_2, \cdots$, such that

- for each i, the value of X_i is either 0 or 1;
- for all values of i, the probability p is the same.

In other words, Bernoullic process is a sequence of independent identically distributed Bernoulli trials. The summary of Bernoulli process and its associated distributions are shonw in Figure 1.2

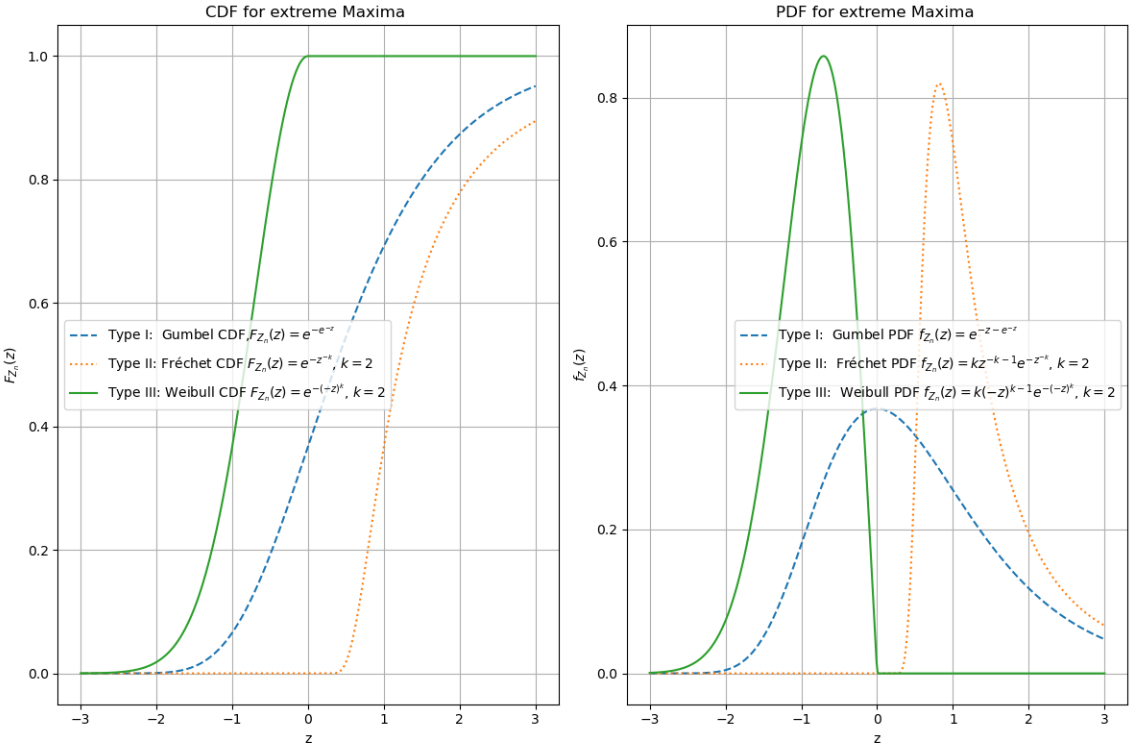


Figure 1.1: Extreme Value distribution for Maximum

Poisson Process

The homogeneous Poisson process can be defined as a counting process with rate $\lambda > 0$ and follows the three properties

- Independent occurences: In two non-overlapping intervals, the corresponding numbers of occurences must be statistically independent of each other
- Oribability of occurences proportional to duration: In an interval $(t, t+\Delta t)$, the probability of exactly one occurence is asymptotically proportional to the interval length Δt as $\Delta t \to 0$
- Occurences do not coincide: The probability of two or more occurences within a sufficiently small interval $(t.t + \Delta t)$ must be orders of magnitude lower than the probability of one occurence

For the homogeneous Poisson Process, the mean rate of occurence λ is constant with tme or locations, whereas for nonhomogeneous case, λ is the function of time or location.

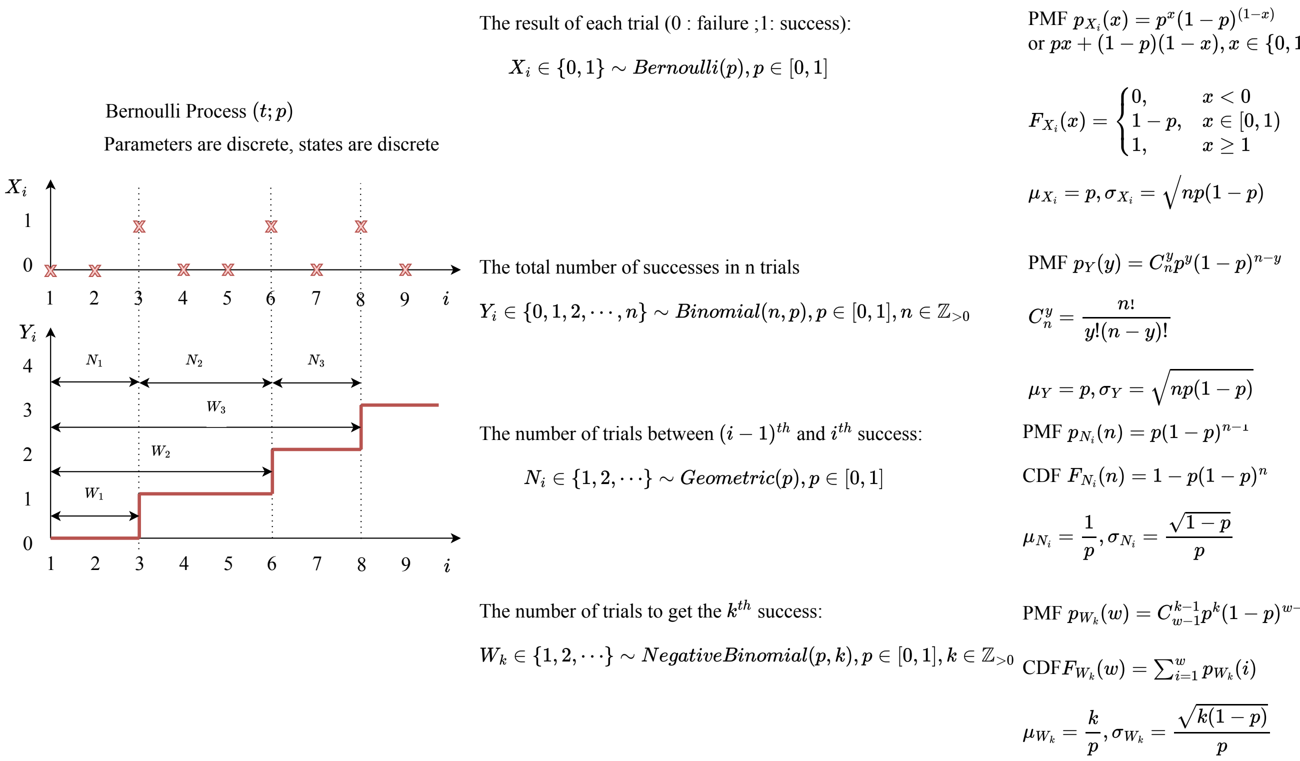


Figure 1.2: Bernoulli Process and its associated Probability Distribution

Compound Poisson Process

A compound Poisson process with rate $\gamma > 0$ abd jump size distribution is a continuous-time stochastic process $\{S(t) : t \geqslant 0\}$ given by

$$S(t) = \sum_{i=1}^{Y(t)} X_i, \tag{1.14}$$

where the sum is by convention equal to zero as long as $Y(t) = 0$. Here $\{Y(t), t \geqslant 0\}\}$ is a Poisson process with rate γ and $\{X_i; i \geqslant 1\}\}$ are independent and identically distributed random variables with distribution function f_{X_i}, which is also independent of $\{Y(t), t \geqslant 0\}\}$. The summary of Compound Poisson Process in shown in Figure 1.4 Compound poisson Process builds on the Poisson process by adding randomness to the event magnitudes. The Poisson porcess models the timing of the events, while the compound Poisson Process accounts for both the timing and the accumulated effect of events.

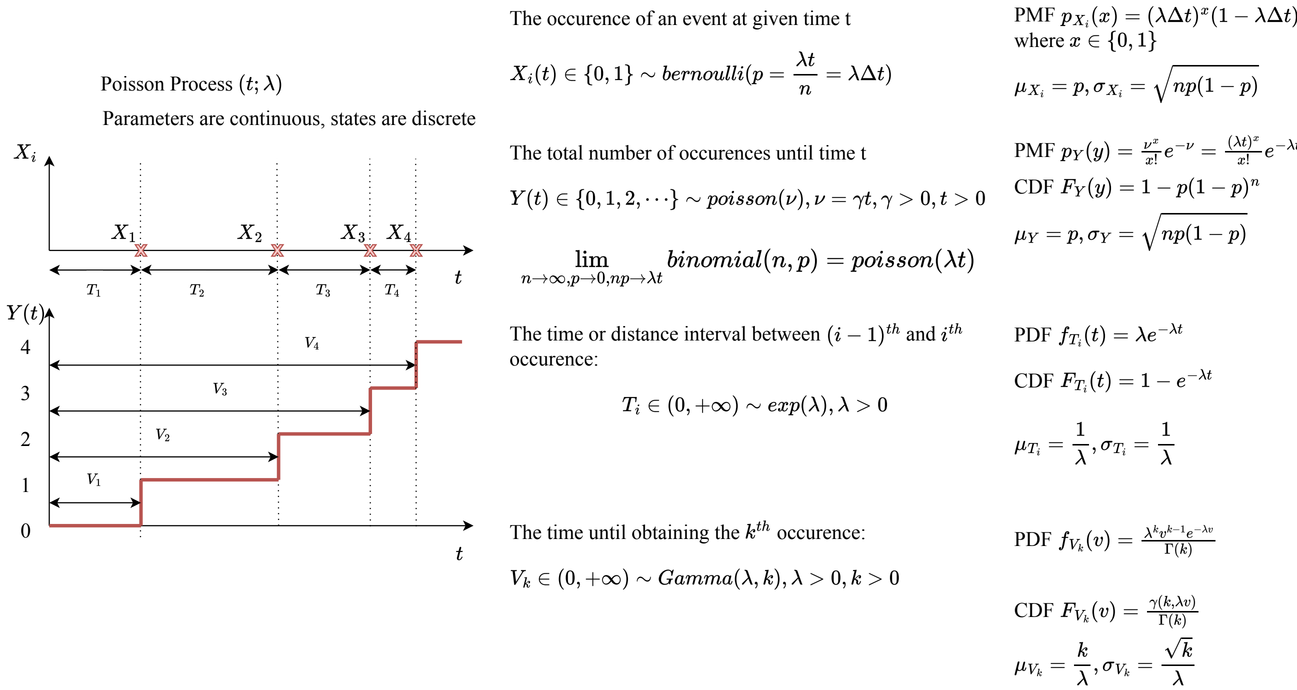


Figure 1.3: Poisson Process and its associated Probability Distribution

Gamma Process

The gamma process $\Gamma(t;\gamma,\lambda)$ is a process which measures the number of occurences of independent gamma-distributed variables over a span of time.

The gamma distribution is a two-parameter continuous parobability distributions family. The exponential, Erlang and chi-squared distributions are special cases of the gamma distributions. The two parameters are the shape parameter $\gamma > 0$ and the rate parameter $\lambda > 0$ (or equivalently the scale parameter $\theta = \frac{1}{\lambda}$). Assume the random variable $X \sim \Gamma(\gamma,\lambda)$, then the probability density function is

$$f_X(x;\gamma,\lambda) = \frac{x^{\gamma-1}e^{-\lambda x}\lambda^{\gamma}}{\Gamma(\gamma)}, \text{ for } x > 0 \tag{1.15}$$

The gamma function $\Gamma(\cdot)$ is the extension of factorial function to complex numbers. It is defined for all complex numbers z except nno-positive integers.

$$\Gamma(z) = \begin{cases} (z-1)!, & z \in \mathbb{Z}_{>0} \\ \int_0^\infty t^{z-1}e^{-t}dt, & \mathbf{R}(z) > 0 \end{cases} \tag{1.16}$$

For the gamma process: the increment distribution is following the gamma distribution.

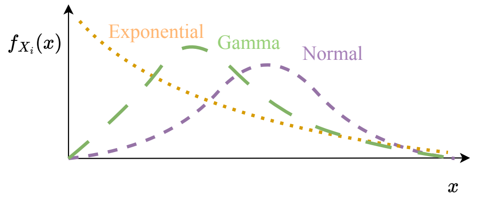
$$S(t+\Delta t) - S(t) \sim \Gamma(\gamma\Delta t, \lambda) \tag{1.17}$$

It is used to model phenomena where the cumulative total grows smoothly over time, e.g. Modelling degradation and aging in systems due to wear and tear.

In contrast the compound Poisson Process models the phenomena where discrete random events contribute to a cumulative sum, e.g. total rainfall, aggregate claims in insurance, financial losses.

The gamma process is a random process consisting of independently distributed gamma distributions where $Y(t)$ represents the number of event occurences from time 0 to time t. The gamma distribution has shape parameter γ and rate parameter λ, often wirtten as $\Gamma(\gamma, \lambda)$, $\gamma > 0$ and $\lambda > 0$. The gamma process is often written as $\Gamma(t; \gamma, \lambda)$ where t represents the time from 0. The process is a pure-jump increasing Levy proess with intensity measure $\nu(x) = \gamma x^{-1} e^{(-\lambda x)}$ for all positive x. Thus jumps whose size lies in the interval $[x, x+d)$ occur as a Poisson process with intensity $\nu(x)dx$. The parameter γ controls the rate of jump arrivals and the scaling parameter γ inversely controls the jump size. It is assumed the process starts from a value 0 at t=0 , i.e. $Y(0) = 0$

Compound Poisson Process
$(t; \lambda, f_{X_i})$
Parameters are continuous, states are discrete



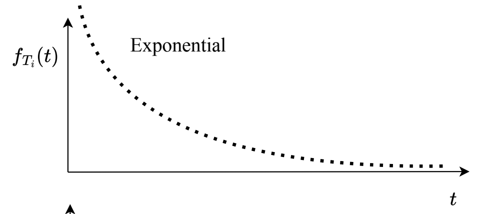
The magnitude of i^{th} event $X_i \in (0, +\infty)$ could follow any specified distribution, e.g. continuous dist.: normal, gamma (exponential is a special case of gamma distribution); or discrete dist.

If X_i is gamma distribution with shape parameter a and rate parameter b

$$X_i \in (0, +\infty) \sim \Gamma(a, b), a > 0, b > 0$$

PDF: $f_X(x) = \frac{x^{a-1}e^{-x/b}}{b^a\Gamma(a)}$
CDF: regularized incomplete gamma function $F_X(x) = \frac{\gamma(a, x/b)}{\Gamma(a)}$
lower incomplete gamma function
$\gamma(a, x/b) = \int_0^{x/b} t^{a-1}e^{-t}dt$
gamma function
$\Gamma(a) = \int_0^{\infty} t^{a-1}e^{-t}dt$

$\mu_X = ab, \sigma_X = \sqrt{a}b$



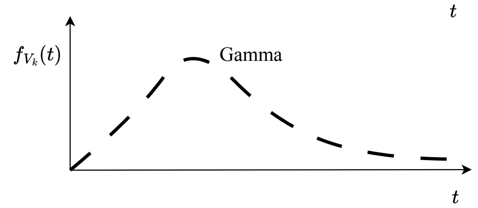
The time or distance between $(i-1)^{th}$ and i^{th} success:

$$T_i \in (0, +\infty) \sim exp(\lambda), \lambda > 0$$

PDF $f_{T_i}(t) = \lambda e^{-\lambda t}$

CDF $F_{T_i}(t) = 1 - e^{-\lambda t}$

$$\mu_{T_i} = \frac{1}{\lambda}, \sigma_{T_i} = \frac{1}{\lambda}$$



The time until obtaining the k^{th} success:Erlang dist.

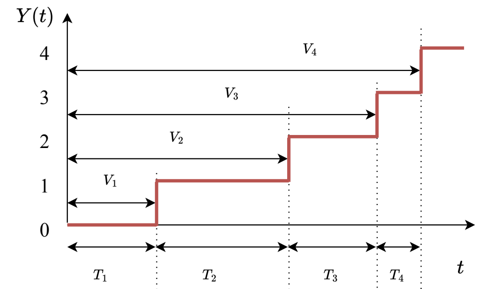
$$V_k \in (0, +\infty) \sim \Gamma(k, \lambda), \lambda > 0, k = 1, 2, \cdots$$

In numpy
V_k = np.random.gamma(shape = k,scale = 1/lambda)

PDF $f_{V_k}(t) = \frac{\lambda^k t^{k-1}e^{-\lambda t}}{\Gamma(k)} = \frac{\lambda^k t^{k-1}e^{-\lambda t}}{(k-1)!}$

CDF $F_{V_k}(t) = \frac{\gamma(k, \lambda t)}{\Gamma(k)} = \frac{\int_0^{\lambda t} \tau^{k-1}e^{-\tau}d\tau}{(k-1)!}$

$$\mu_{V_k} = \frac{k}{\lambda}, \sigma_{V_k} = \frac{\sqrt{k}}{\lambda}$$

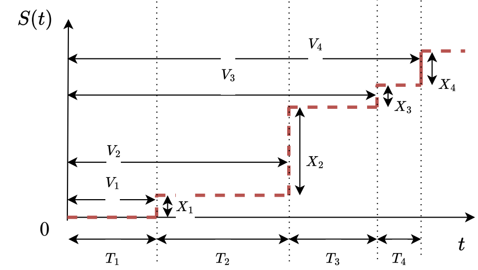


The total number of successes until time t

$$Y(t) \in \{0, 1, 2, \cdots\} \sim poisson(\nu), \nu = \lambda t, \lambda > 0, t > 0$$

$$\lim_{y \to \infty, p \to 0, yp \to \lambda t} binomial(y, p) = poisson(\lambda t)$$

PMF $p_Y(y) = \frac{\nu^y}{y!}e^{-\nu} = \frac{(\lambda t)^y}{y!}e^{-\lambda t}$
CDF $F_Y(y) = 1 - p(1-p)^y$
$\mu_Y = \nu = \lambda t, \sigma_Y = \sqrt{\nu} = \sqrt{\lambda t}$



The accumulated total amount until time t

$$S(t) = \sum_{i=1}^{Y(t)} X_i \in (0, +\infty) \sim CPP(t; \nu, F_X(x)), \nu = \lambda t, \lambda > 0, t > 0$$

$$\mathbb{E}[S(t)] = \mathbb{E}[\mathbb{E}[S(t)|Y(t)]] = \mathbb{E}[Y(t)\mathbb{E}[X(t)]] = \lambda t \mathbb{E}[X_i]$$

$$\mathbb{D}[S(t)] = \mathbb{E}[\mathbb{D}[S(t)|Y(t)]] + \mathbb{D}[\mathbb{E}[S(t)|Y(t)]] = \mathbb{E}[\sigma_X^2 Y(t)] + \mathbb{D}[\mu_X Y(t)] = \sigma_X^2\mu_Y + \mu_X^2\sigma_Y^2 = \lambda t$$

$$S|_{Y(t)=y} = \sum_{i=1}^{y} X_i \sim \Gamma(ya, b) \qquad f_{S|_{Y=y}}(x) = (f_X * \cdots f_X)(x) = \frac{b^{ya}x^{ya-1}e^{-bx}}{\Gamma(ya)}$$

$$f_S(x) = \sum_{y=0}^{\infty} p_Y(y) f_{S|_{Y=y}}(x) = e^{-\lambda t} + \sum_{y=1}^{\infty} \frac{(\lambda t)^y}{y!}e^{-\lambda t}\frac{b^{ya}x^{ya-1}e^{-bx}}{\Gamma(ya)} = e^{-\lambda t} + \sum_{y=1}^{\infty} \frac{(\lambda t)^y b^{ya} x^{ya-1}}{y!\Gamma(ya)} e$$

$$W(x, z) = \sum_{y=1}^{\infty} \frac{x^y}{y!\Gamma(zy)} \qquad f_S(x) = e^{-\lambda t} + \frac{e^{-\lambda t - bx}}{x} \sum_{y=1}^{\infty} \frac{(\lambda t (bx)^a)^y}{y!\Gamma(ya)} = e^{-\lambda t} + \frac{e^{-\lambda t - bx}}{x} W(\lambda t(bx$$

$$F_S(s) = \int_0^s f_S(x)dx = e^{-\lambda t} + \int_0^s \frac{e^{-\lambda t - bx}}{x} \sum_{y=1}^{\infty} \frac{(\lambda t(bx)^a)^y}{y!\Gamma(ya)} dx$$

Figure 1.4: Compound Process Process and its associated Probability Distribution

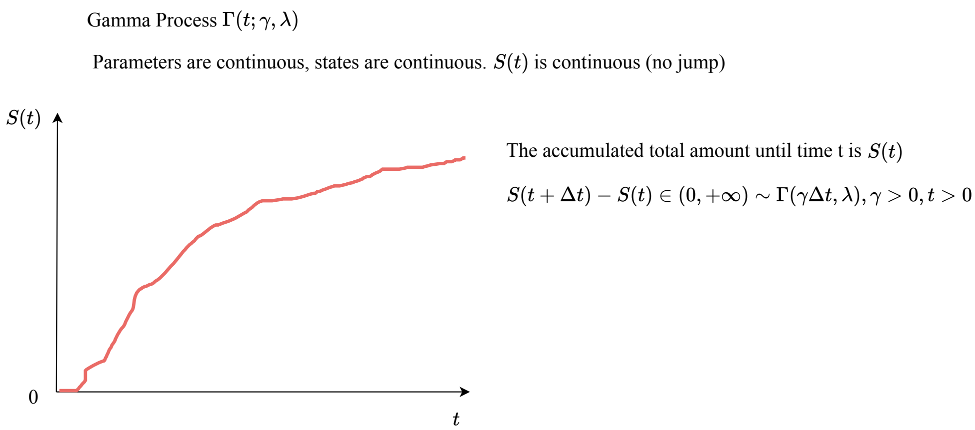


Figure 1.5: Gamma Process and its associated Probability Distribution