Probability Theory for EOR

Some special continuous random variables I (Exponential)

Some **continuous random variables** are associated very special/ubiquitous **distributions (PDFs)**, they get their own names!

Definition (PDF of continuous real-valued r.v.)

The **probability density function (PDF)** of a **continuous** real-valued r.v. is a non-negative function f_X on the real line such that via the Riemann integral:

$$\int_{-\infty}^{x} f_X(s)ds = P(X \leq x).$$

For a **continuous** random variable with differentiable CDF F_X, conventionally, $f_X(x) = F'_X(x)$.

Exponential distribution

Waiting time

Customers come in according to a Poisson process such that:

Poisson arrivals.
The number of arrivals that occur in an interval of length t is a Pois(λt) r.v.;

Independence condition.
The number of arrivals that occur in disjoint intervals are independent from each other.

What is the distribution of the waiting time until the first customer?

Waiting time

What is the distribution of the waiting time until the first customer?

- Denote N_t the number of arrivals within $[0, t]$, then N_t follows Pois(λt).
- Denote T_1 the time until the first arrival.
- Note that $\{T_1 > t\} = \{N_t = 0\}$.
- $P(T_1 > t) = P(N_t = 0) = e^{-\lambda t}$.
- The CDF of T_1 then

$$F_{T_1}(t) = P(T_1 \leq t) = 1 - P(T_1 > t) = 1 - e^{-\lambda t}.$$

- **CDF has non-zero first derivative for $t > 0$ such that CDF can be rewritten as a Riemann integral of the derivative**

$$f_{T_1}(t) = F'_{T_1}(t) = \lambda e^{-\lambda t}.$$

Exponential distribution: Expo(λ)

- A continuous real-valued r.v. X follows Expo(λ) if its PDF $f_X(x)$ is

$$f_X(x) = \lambda e^{-\lambda t}, t > 0$$

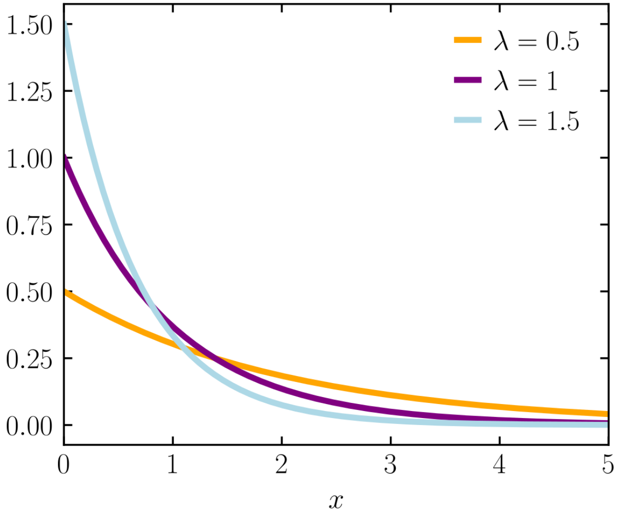
  and zero for $t \leq 0$.

- CDF: $F_X(t) = 1 - e^{-\lambda t}$.

- **Note that if $X \sim$Expo(1), then $Y = \frac{X}{\lambda} \sim$Expo($\lambda$).**

$$P(Y \leq t) = P(X \leq \lambda t) = 1 - e^{-\lambda t}.$$

PDF Expo(λ)



Expectation and Varaince

$X \sim$Expo(1), $Y = \frac{X}{\lambda} \sim$Expo($\lambda$).

- The expectation EX.

$$\mathrm{E}[X] = \int_0^\infty x e^{-x} dx = 1.$$

- The varaince VX.

$$\mathrm{E}[X^2] = \int_0^\infty x^2 e^{-x} dx = 2.$$
$$\mathrm{V}[X] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2 = 1.$$

- $EY = E\left(\frac{X}{\lambda}\right) = 1/\lambda, VY = V\left(\frac{X}{\lambda}\right) = 1/\lambda^2.$

I. From Poisson process to exponential: waiting time between the first and the second arrivals is also Expo

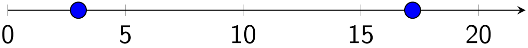
Customers come in according to a Poisson process such that:

Poisson arrivals.
The number of arrivals that occur in an interval of length t is a Pois(λt) r.v.;

Independence condition.
The number of arrivals that occur in disjoint intervals are independent from each other.

Denote T_i the waiting time until the ith arrivals.

The distribution of $T_2 - T_1$.



- Suppose at time t, there is one visit, what is the time of next visitor?
  * We re-start to wait and as if we are waiting for the *first* arrival.
  * $(0, T_1)$ and $(T_1, +\infty)$ are disjoint intervals and thus arrivals in these intervals are independent.

- Therefore, $T_2 - T_1 \sim \text{Expo}(\lambda)$.
  Results can be generalised to the waiting time between the ith and the $(i+1)$th arrivals.
- However, T_2 does not follow exponential distribution: the sum of independent Expo's is not Expo.

II. Memoryless property

Definition

Memoryless property

$$\mathbb{P}(X - s \geq t | X \geq s) = \mathbb{P}(X \geq t)$$

If X is a positive continuous real-valed random variable with the memoryless property, then X has an exponential distribution.

- Let $G(x) = 1 - F(x)$, then memoryless implies

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

- (1) $G(mt) = (G(t))^m$.
- (2) $G(t/n) = (G(t))^{1/n}$.
- (3) $G(m/nt) = (G(t))^{m/n}$.
- (4) $G(xt) = (G(t))^x$ and thus $G(x) = (G(1))^x = e^{-\log(G(1))x}$
- The CDF of X would be the exponential distribution CDF (Expo(log($G(1)$))): $F(x) = 1 - e^{-\log(G(1))x}$.

- One discrete r.v. also satisfies the memoryless property:Geometric!

III. Minimum of independent Expo's is still Expo

A web server receives requests from independent computers $1, \cdots, n$, **the number of requests from the computer i follows Poisson process with unit arrival rate λ_i.**

the waiting time before the first request from computer i, T_i, follows Expo(λ_i).

What is the distribution of $T = \min_{1 \le i \le n} T_i$?

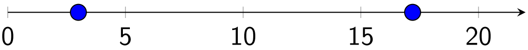
- Note that the number of total requests in a time interval of lenghth t, N_t, follows Poisson distribution with arrival rate $\left(\sum_{1 \le i \le n} \lambda_i\right) t$.
- Therefore, the waiting time before any requests would follow Expo$\left(\left(\sum_{1 \le i \le n} \lambda_i\right) t\right)$.
- $$F_T(t) = 1 - P(T > t) = 1 - P(N_t = 0) = 1 - e^{-\left(\sum_{1 \le i \le n} \lambda_i\right) t}.$$

IV. From Exponential to Poisson

The economy starting from time 0 follows a business cycle which switches between peak and trouph periods, suppose each period lasts for a random time length following i.i.d. Expo(λ).

What is the distribution of the number of transitions from one status to another in the time interval (0,t)?

- It is like the economy would receive a signal, such that receiving the signal would change economic status.



Time duration between each successive signals follows i.i.d. Expo(λ).

- Therefore, **the number of signals follow Pois(λt).**

The waiting time between two Poisson process arrivals follows independent exponential distributions, while if the waiting time between two successive arrivals follows i.i.d. exponential distributions then the number of arrivals follows a Poisson process.