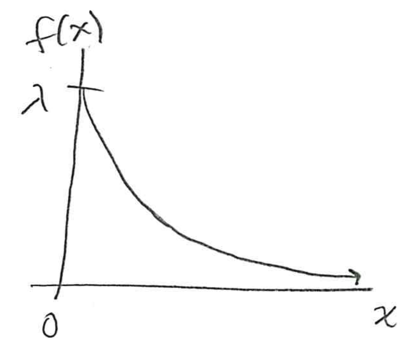
Exponential Distribution

[Ref: Models §5.1-5.2]

def: A continuous random variable X has an exponential distribution with parameter λ $(\lambda > 0)$ if its PDF is

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$



CDF :

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Mean : $E[X] = \frac{1}{\lambda}$

Variance : $Var(X) = \frac{1}{\lambda^2}$

Properties

① Memoryless: $\boxed{P(X > s+t \mid X > t) = P(X > s)}$

for all $s,t \geq 0$

Interpretation: Let X = lifetime of an instrument (device)

If the instrument works at time t, then the future lifetime has the same distribution as the original lifetime.

Note: Exponential is the only continuous distribution with this property!

⇒ This property makes the exponential distn easy to analyze & useful for modeling

[e.g. An item that has been in use for 10 hrs is as good as a new item in regards to the amount of time remaining until the item fails.]

② If $X_1, \ldots, X_n$ are indep. exponential RVs with rates $\lambda_1, \ldots, \lambda_n$, then

$X = \min(X_1, \ldots, X_n)$ is exponential with rate $\lambda = \lambda_1 + \cdots + \lambda_n$

Random Sum of Exponentials

Let $X_1, \dots, X_n$ be i.i.d. exponential RVs with rate $\beta > 0$.

PDF of X_i (for all i) : $f(x) = \beta e^{-\beta x}$, $x > 0$.

The sum $X = X_1 + \cdots + X_n$ is a gamma distributed RV with parameters $\alpha = n$ and β.

$X \sim \text{gamma}(n, \beta)$

PDF of X : $f_X(x) = \frac{\beta^n}{\Gamma(n)} x^{n-1} e^{-\beta x}$, $x > 0$

Q. What if the X_i's are NOT i.i.d. ?

Suppose $X_i \sim \text{exponential}(\lambda_i)$

Assume all rates are different : $\lambda_i \neq \lambda_j$ for $i \neq j$.

$X = X_1 + \cdots + X_n$ is a hypoexponential RV

PDF: $f_X(x) = \sum_{i=1}^{n} C_{i,n} \lambda_i e^{-\lambda_i x}$

* Use convolutions to derive this

where $C_{i,n} = \prod_{j \neq i} \frac{\lambda_i}{\lambda_j - \lambda_i}$

Note: This PDF is not a mixture of exponential RVs since the $C_{i,n}$'s can be negative! (some will be)

but $\sum_{i=1}^{n} \prod_{j \neq i} C_{i,n} = 1$

Mixtures: If X_i has CDF F_i $(i = 1, \dots, n)$

and $p_i > 0$ s.t. $p_1 + \dots + p_n = 1$, ← mixing probabilities

then RV X with CDF

$$F(x) = p_1 F_1(x) + \dots + p_n F_n(x)$$

$$= \sum_{i=1}^{n} p_i F_i(x)$$ is called a mixture of the F_i's.

Note: PDF of the mixture

$$\frac{d}{dx} F(x) = \sum_{i=1}^{n} p_i f_i(x)$$ where f_i = PDF of X_i

def: A mixture of exponential RVs is called a hyperexponential RV.

Hazard Rate

Let X be a positive, continuous RV with PDF f and CDF F

$$S(x) = P(X > x) = 1 - P(X \le x) = 1 - F(x)$$

↑ survival function of X

$$\boxed{r(t) = \frac{f(t)}{S(t)}} = \text{hazard rate function}$$

(aka failure rate, mortality rate)

Interpretation: $r(t)$ represents the conditional probability density of a t-year old item to fail.

↑ or whatever unit of time you're using

For the exponential distribution:

$$\boxed{r(t) = \lambda = \frac{1}{\text{mean}}}$$

Why is this true?

Details : $f(x) = \lambda e^{-\lambda x}$, $x > 0$

$$S(x) = 1 - F(x) = 1 - (1 - e^{-\lambda x}) = e^{-\lambda x}$$

$$\Rightarrow \quad r(t) = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda \quad ✓$$

Poisson Process

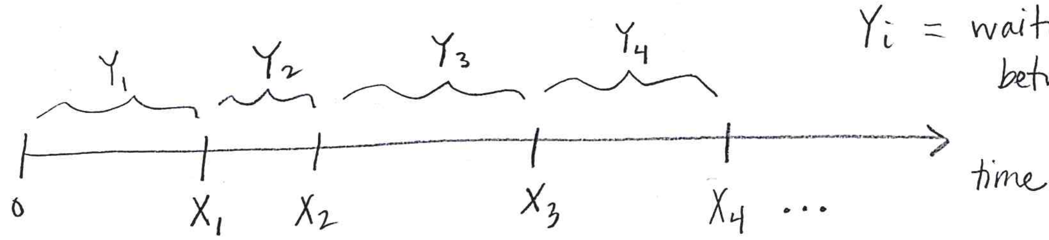
- Type of counting process, connection with the exponential distribution.
- Simple & widely used stochastic process for modeling the times at which arrivals enter a system.

Examples

1. Customers arriving at a store
2. File requests on a server
3. Traffic accidents at an intersection

X_i = time of the i^{th} arrival

Y_i = waiting times between $i-1$ & i^{th} arrivals



$$X_i = \sum_{k=1}^{i} Y_k$$

- If $s<t$, then $N(s) \leq N(t)$

- For $s<t$, $N(t) - N(s)$ equals the number of events that occur in the interval $(s,t]$.

stopped

Independent Increments:

A counting process has independent increments if the numbers of events that occur in disjoint time intervals are independent.

Example: $N(10)$ - the number of events that occur by time 10.

$N(10)$ must be independent of $\underbrace{N(15) - N(10)}_{\text{\# of events that occur between times 10 \& 15}}$.

i.e. $N(10) - N(0)$

→ This assumption might be reasonable for Example 1. in beginning of this section.

→ Might be unreasonable for other counting process examples, e.g. Ex. 2. if the soccer player had a hot streak at end of the game.

Stationary Increments:

A counting process has stationary increments if the distribution of the number of events that occur in any interval of time depends only on the length of the time interval.

i.e. # of events in the interval $(s, s+t)$ has the same distribution $\forall s$.

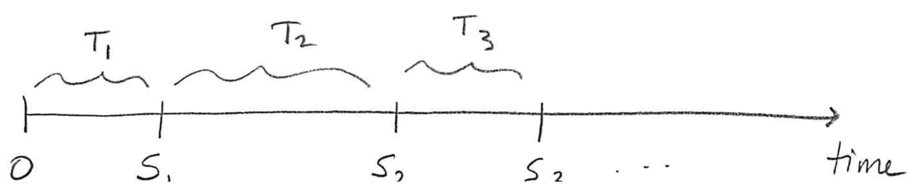
→ This assumption is not reasonable for Example 1 since the store most likely has "rush hours" (e.g. 12-1, 5-6)

extra busy

Poisson Process

def 1: Let $T_1, T_2, \ldots$ be independent exponential(λ) RVs and let $S_n = T_1 + \cdots + T_n$, for $n \geq 1$ and $T_0 = 0$.

Define $N(t) = \max\{n : S_n \leq t\}$. Then $\{N(t) : t \geq 0\}$ is a Poisson Process.



T_i's are times between arrivals

S_n is the arrival time of n^{th} event

$N(t)$ is the # of arrivals by time t

def 2 : The counting process $\{N(t) : t \geq 0\}$ is a Poisson process with rate $\lambda > 0$ if

(i) $N(0) = 0$

(ii) $\{N(t) : t \geq 0\}$ has independent increments

i.e. for any $0 \leq t_0 < t_1 < \cdots < t_n < \infty$

$N(t_0)$, $N(t_1) - N(t_0)$, ..., $N(t_n) - N(t_{n-1})$

are all independent RVs

(iii) $N(s+t) - N(s) \sim \text{Poisson}(\lambda t)$

i.e. the number of events in any interval of length t follows a Poisson distribution with rate λt

mean length of time interval

Further Details :

- $P(N(t+h) - N(t) = 1) = \lambda h + \underbrace{o(h)}_{\text{"little o" of } h}$

- $P(N(t+h) - N(t) \geq 2) = o(h)$

$f(\cdot)$ is $o(h)$ if : $\lim_{h \to 0} \frac{f(h)}{h} = 0$

e.g. $f(x) = x^2$ is $o(h)$ since

$\lim_{h \to 0} \frac{h^2}{h} = 0$

<u>Remark</u> : Since the distribution of $N(s+t)-N(s)$ is the same for all s, it follows that the Poisson process has <u>stationary increments</u>.

Poisson Process Properties

Let $\{N(t) : t \geq 0\}$ be a Poisson process with rate $\lambda > 0$.

$$E[N(t)] = \lambda t$$

$$Var[N(t)] = \lambda t$$

$$\begin{cases} N(t) \sim \text{Poisson}(\lambda t) & \text{\& more generally,} \\ N(s+t) - N(s) \sim \text{Poisson}(\lambda t) \end{cases}$$

i.e. $P(N(t)=n) = \dfrac{e^{-\lambda t} (\lambda t)^n}{n!}$.

Poisson RV w/ rate λt

& likewise $P(N(s+t) - N(s) = n) = \dfrac{e^{-\lambda t}(\lambda t)^n}{n!}$

for all $s \geq 0$

Example: Suppose customers arrive at a store according to a Poisson process with rate $\lambda = 4$ per hour. Given that the store opens at 9am, what is the probability that exactly 1 customer has arrived by 9:30 am and a total of 5 have arrived by 11:30 am?

- Compute $P(N(0.5) = 1)$

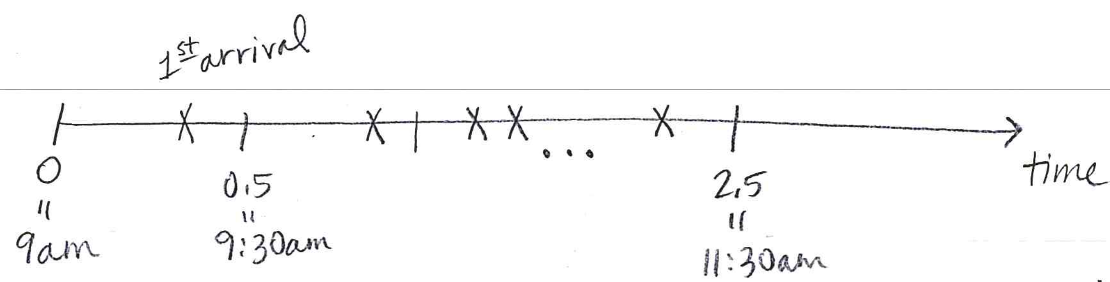
  $\frac{1}{2}$ hour, since time is in units of hours (given by rate parameter λ)

- Compute $P(N(2.5) = 5)$

- Compute the intersection of these events

  "and"

  $$P(N(0.5) = 1, N(2.5) = 5).$$



Yes, indep!

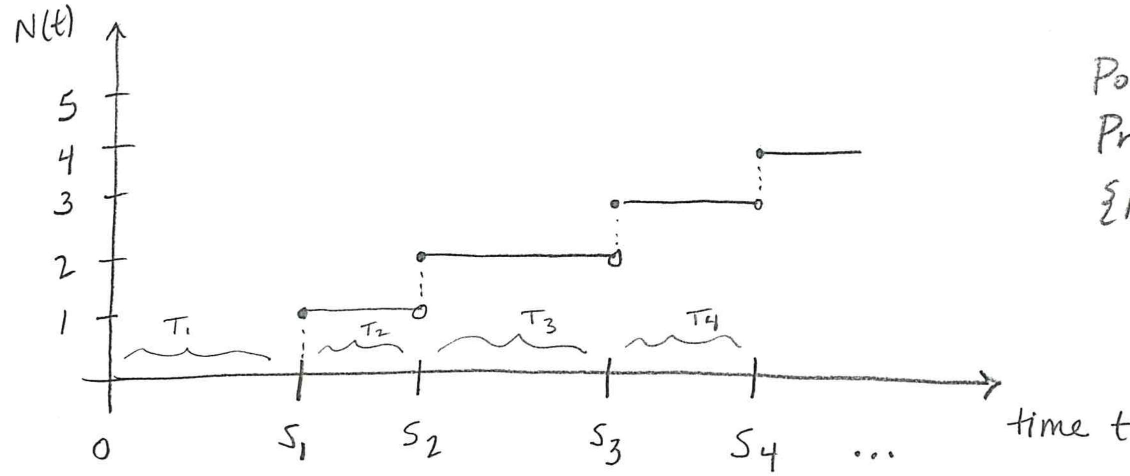
$$P(N(0.5)=1, N(2.5)=5) = P(N(0.5)=1, \underbrace{N(2.5)-N(0.5)=4}_{\text{length of interval} = 2})$$

not indep!

$$= \frac{e^{-4(0.5)}(4 \cdot 0.5)^1}{1!} \times \frac{e^{-4(2)}(4 \cdot 2)^4}{4!}$$

$$\approx 0.0155$$

Interarrival & Waiting Time Distribution



Poisson Process $\{N(t) : t \geq 0\}$

S_i = waiting time for the i^{th} event, $i \geq 1$

A sample path of the process $\{N(t) : t \geq 0\}$ "jumps" by 1 at each S_i

Q. What is the probability distribution of the S_n's?

$$S_n = T_1 + T_2 + \cdots + T_n$$

where T_1 = time of the first event $(= S_1)$

$T_2 = S_2 - S_1$ = interarrival time b/t 1^{st} & 2^{nd} event

$\vdots$

$T_j = S_j - S_{j-1}$ = interarrival time b/t $j-1^{st}$ & j^{th} event

Q. What is the distribution of the T_j's?

$\Rightarrow$ $\boxed{\text{i.i.d. exponential RVs w/parameter } \lambda}$

Recall: Sum of n i.i.d. exponential (λ) RVs is a gamma RV with parameters n & λ

$$\Rightarrow \boxed{S_n \sim \text{gamma}(n, \lambda)}$$

Relation between Poisson & Gamma RVs

Let $X \sim \text{Poisson}(\lambda)$ random variable

PDF of X: $$P(X=n) = \frac{e^{-\lambda}\lambda^n}{n!} \quad \text{for } n = 0,1,2,\ldots \quad \lambda > 0$$

Let $Y \sim \text{Gamma}(\overset{\alpha}{n+1}, \overset{\beta}{1})$ RV

PDF of Y: $$f_Y(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\beta y}, \quad y > 0$$

For $\alpha = n+1$, $\beta = 1$

$$\Rightarrow f_Y(y) = \frac{1^{n+1}}{\Gamma(n+1)} y^{(n+1)-1} e^{-1y}$$

$$= \frac{1}{n!} y^n e^{-y}$$

Evaluate for $y = \lambda$

$$\Rightarrow f_Y(\lambda) = \underbrace{\frac{e^{-\lambda}\lambda^n}{n!}}_{}, \quad \lambda > 0$$

(*) same as PDF of X above!

As a function of λ, (*) is the PDF of Y
As a function of n, (*) " " X

Fact: $P(X \le n) = P(Y \ge \lambda)$

stopped

Simulating a Poisson Process

* Easier approach!

1. Simulate i.i.d. exponential RVs $T_1, T_2, \ldots$ and keep track of $S_n = T_1 + \cdots + T_n$

2. $S_1, S_2, \ldots, S_n$ are the event times of the events in the interval $[0,t]$ where

$$\boxed{S_n \le t \quad \text{but} \quad S_{n+1} > t}$$

Alternative approach for simulation

1. Simulate the number of events in $[0,t]$ (i.e. a Poisson RV with mean = rate = λt)

2. Given $N(t) = n$, generate $S_1, \ldots, S_n$ via conditional distribution of the S_i's

uses order statistics...

Generalizations of the Poisson Process

- Multiple Poisson Processes
- Nonhomogeneous Poisson Processes
- Others (e.g. Spatial PP, compound PP, Marked PP, Mixed PP → next week!)

Multiple Poisson Processes

Let $\{N(t) : t \geq 0\}$ be a Poisson Process with rate λ. Suppose each event, independently of all other events, is classified as type i, $i = 1, \ldots, K$, with probability p_i ($\& \sum_{i=1}^{K} p_i = 1$).

If $N_i(t)$ = # of events of type i occurring in $[0, t]$,

Then $\{N_i(t) : t \geq 0\}$ is a Poisson Process with rate λp_i for each i, and these are all <u>independent</u>.

<u>Example</u> : Customers arrive at a store according to a Poisson Process with rate λ. Suppose that each arrival is a female with probability $\frac{1}{2}$ & a male with prob. $\frac{1}{2}$.
(i.e. type 1) (type 2)

$N_1(t)$ = # of females arriving at store by time t

$N_2(t)$ = # of males " " "

$$N(t) = N_1(t) + N_2(t)$$

$\{N_1(t) : t \geq 0\}$ & $\{N_2(t) : t \geq 0\}$ are independent Poisson processes with rates $\lambda(\frac{1}{2})$.

since $\lambda p_1 = \lambda p_2 = \lambda(\frac{1}{2})$