Stochastic Processes

Kevin Chang

February 22, 2022

1 Laplace Transform

- $\mathcal{L}\{f\}(s) = \int_0^\infty f(t)e^{-st}dt$
- Property
  - $tf(t) \leftrightarrow -F'(s)$
  - $\frac{f(t)}{t} \leftrightarrow \int_s^\infty F(\sigma)d\sigma$
  - $f'(t) \leftrightarrow sF(s) - f(0^-)$
  - $\int_0^t f(\tau)d\tau \leftrightarrow \frac{F(s)}{s}$
  - $e^{at}f(t) \leftrightarrow F(s-a)$
  - $f(t-a)u(t-a) \leftrightarrow e^{-at}F(s)$

2 Moment Generating Function

- Moment Generating Function: $\mathbb{E}[e^{tX}]$
  - Property:
    * $\mathbb{E}[e^{tX}] = \int_{-\infty}^\infty e^{tx}f_X(x)dx$
    * $\mathbb{E}[e^{tX}] = \sum_{k=0}^\infty E[X^k]\frac{t^k}{k!}$
      · $e^{tx} = \sum_{k=0}^\infty \frac{(tx)^k}{k!}$
      · $E[e^{tX}] = E[\sum_{k=0}^\infty \frac{(tX)^k}{k!}] = \sum_{k=0}^\infty E[X^k]\frac{t^k}{k!}$
    * $\frac{d\mathbb{E}[e^{tX}]}{dt} = \mathbb{E}[X]$
    * $\mathbb{E}[e^{t(aX+b)}] = e^tb\mathbb{E}[e^{taX}]$
    * Not all random variables have Moment generating function
- Characteristic Function: $\mathbb{E}[e^{itX}]$
  - Property:
    * All random variables have Moment generating function
- Joint Moment Generating Function: $G(x,y) = \mathbb{E}[e^{xX}e^{yY}]$
- Property:
  - (Joint) moment generating funciton uniquely determines the (joint) CDF
- Example
  - Trapped miner's random walk
    * Miner has probability of $\frac{1}{3}$ to waste 3 hours in vain, $\frac{1}{3}$ to waste 5 hours in vain, and $\frac{1}{3}$ to spend 2 hours to go out of the mine.
    * X is the random variables of the hours to go out of the mine
    * Y_i is the random variables of the hours for the i-th action.
    * $\mathbb{E}[e^{tX}] = \mathbb{E}[e^{tX}|Y_1 = 2] + \mathbb{E}[e^{tX}|Y_1 = 3] + \mathbb{E}[e^{tX}|Y_1 = 5]$
      $= \mathbb{E}[e^{2t}] + \mathbb{E}[e^{t(X+3)}] + \mathbb{E}[e^{t(X+5)}]$
    * Find expectation and variance by joint moment generating function

3 Expectation

- N i.i.d. events, when N is a random variable
  - Suppose N is a integer random variable
  - Suppose $X_1, \dots, X_i, \dots, X_N$ are i.i.d random variables with mean μ and variance σ^2
  - $Y = \sum_{i=1}^{N} X_i$
  - $\mathbb{E}[Y] = \mathbb{E}[N]\mu$
    * $\mathbb{E}[Y] = \sum_{n=1}^{\infty} \mathbb{E}[\sum_{i=1}^{N} X_i | N = n] P[N = n]$
      $= \mu \times \sum_{n=1}^{\infty} nP[N = n] = \mathbb{E}[N]\mu$
  - $\mathbb{E}[Y^2] = \mathbb{E}[N]\mathbb{E}[X^2] + \mathbb{E}[N^2]\mu^2 - \mathbb{E}[N]\mu^2$
    * $\mathbb{E}[Y^2] = \sum_{n=1}^{\infty} \mathbb{E}[(\sum_{i=1}^{N} X_i)^2 | N = n] P[N = n] = \sum_{n=1}^{\infty} (n\mathbb{E}[X_i^2] + n(n-1)\mu^2) P[N = n]$
      $= \mathbb{E}[N]\mathbb{E}[X^2] + \mathbb{E}[N^2]\mu^2 - \mathbb{E}[N]\mu^2$
  - $Var(Y) = \mathbb{E}[N]\sigma^2 + Var(N)\mu^2$
- Expectation by $P[X > x]$
  - $\mathbb{E}[X] = \sum_x P[X > x]$, when X is a non-negative discrete random variable
    * $\mathbb{E}[X] = \sum_{x=0}^{\infty} xP[X = x] = \sum_{x=0}^{\infty} \sum_{y=0}^{x-1} P[X = x] = \sum_{y=0}^{\infty} \sum_{x=y+1}^{\infty} P[X = x] = \sum_{y=0}^{\infty} P[X > y]$
  - $\mathbb{E}[X] = \int_0^{\infty} P[X > x]dx$, when X is a non-negative continuous random variable
    * $\mathbb{E}[X] = \int_0^{\infty} xf_X(x)dx = \int_0^{\infty} \int_0^{x} f_X(x)dydx = \int_0^{\infty} \int_y^{\infty} f_X(x)dxdy = \int_0^{\infty} P[X > y]dy$

4 Inequality

- Markov Inequality

  Definition:
  - Suppose $X \geq 0$, then $P[X \geq \epsilon] \leq \frac{\mathbb{E}[X]}{\epsilon}$

  Proof:
  1. $\mathbb{E}[X] = \int_0^{\infty} xf_X(x) \geq \int_{\epsilon}^{\infty} xf_X(x) \geq \epsilon \int_{\epsilon}^{\infty} f_X(x) = \epsilon P[X \geq \epsilon]$
  2. $X(\omega) \geq \epsilon \mathbb{1}_{X(\omega) \geq \epsilon}, \forall \omega \in S$
     - Calculate expectation on both side.
     - $\mathbb{E}[X] \geq \epsilon P[X \geq \epsilon]$

  Property:
  - The equality happens when $P[X = k] = 0, \forall k \notin \{0, \epsilon\}$.
- Chebyshev Inequality

  Definition:
  - Suppose $m = \mathbb{E}[X], \sigma^2 = Var(X)$, then $P[|X - m| \geq \epsilon] \leq \frac{\sigma^2}{\epsilon^2}$

  Proof:
  - $P[|X - m| \geq \epsilon] = P[(X - m)^2 \geq \epsilon^2]$
  - $P[(X - m)^2 \geq \epsilon^2] \leq \frac{\mathbb{E}[(X-m)^2]}{\epsilon^2}$ (by Markov Inequality)

  Property:
  - The equality happens when $P[X = k] = 0, \forall k \notin \{m - \epsilon, m, m + \epsilon\}$.
  - Might be tighter than Markov Inequality since it requires m, σ
- Chernoff Inequality

  Definition:

- Suppose $X_1, \ldots, X_n$ are independent identically distributed Bernoulli random variable with probability p and $X = \sum_{i=1}^{n} X_i$
- $P[X \geq \epsilon] \leq \frac{(pe^t+1-p)^n}{e^{t\epsilon}} \leq \frac{e^{np(e^t-1)}}{e^{t\epsilon}}$
  - $P[X \geq \epsilon] = P[e^{tX} \geq e^{t\epsilon}] \leq \frac{E[e^{tX}]}{e^{t\epsilon}} = \frac{(E[e^{tX_i}])^n}{e^{t\epsilon}} = \frac{(pe^t+1-p)^n}{e^{t\epsilon}} \leq \frac{e^{np(e^t-1)}}{e^{t\epsilon}}$
- $P[X \geq np(1+\epsilon)] \leq (\frac{e^\epsilon}{(1+\epsilon)^{1+\epsilon}})^{np} \leq \begin{cases} e^{\frac{-\epsilon^2 np}{3}} & \text{if } 0 \leq \epsilon \leq 1 \\ e^{\frac{-\epsilon^2 np}{(2+\epsilon)}} & \text{if } \epsilon \geq 1 \end{cases}$
  - Substitude ϵ with $np(1+\epsilon)$
  - Substitude t with $\log(1+\epsilon)$
  - the last inequality is without proof
- $P[X \leq \epsilon] \leq \frac{(pe^t+1-p)^n}{e^{t\epsilon}} \leq \frac{e^{np(e^t-1)}}{e^{t\epsilon}}$
  - $P[X \leq \epsilon] = P[e^{-tX} \geq e^{-t\epsilon}] \leq \frac{E[e^{-tX}]}{e^{-t\epsilon}} = \frac{(E[e^{-tX_i}])^n}{e^{-t\epsilon}} = \frac{(pe^{-t}+1-p)^n}{e^{-t\epsilon}} \leq \frac{e^{np(e^{-t}-1)}}{e^{-t\epsilon}}$
- $P[X \leq np(1-\epsilon)] \leq (\frac{e^{-\epsilon}}{(1-\epsilon)^{1-\epsilon}})^{np} \leq e^{\frac{-\epsilon^2 np}{2}}$
  - Substitude ϵ with $np(1-\epsilon)$
  - Substitude t with $-\log(1-\epsilon)$
  - the last inequality is without proof

- Chernoff/ Hoeffding Lemma

  Definition:
  - Suppose $X_1, \ldots, X_n$ are independent distributed random variable and $a_i \leq X_i \leq b_i$
  - Suppose $X = \sum_{i=1}^{n} X_i$ and $\mu = \mathbb{E}[X]$
  - $P[|X-\mu| \geq \epsilon] \leq 2e^{\frac{-2\epsilon^2}{\sum_{i=1}^{n}(b_i-a_i)^2}}$ without proof

- Application:
  - Balls in Bins

    Definition: Throw n balls into n bins, find bounds for the maximum number of balls in all bins
    - $P[$ maximum number of balls in all bins $\geq \epsilon]$
$= P[\cup_{i=1}^{n}$ number of balls in i-th bin $\geq \epsilon]$
$\leq n \times P[$ number of balls in one bin $\geq \epsilon]$
    - By Markov inequality:
      - $P[$ number of balls in one bin $\geq \epsilon] \leq \frac{1}{\epsilon} \to$ useless
    - By Chebyshev inequality:
      - $P[$ number of balls in one bin $\geq \epsilon] \leq \frac{(1-\frac{1}{n})}{\epsilon^2}$
      - $P[$ maximum number of balls in all bins $\geq n^{\frac{1}{2}+\epsilon}] \leq \frac{(1-\frac{1}{n})}{n^{2\epsilon}}$
      - when $n \to \infty$, the maximum number of balls should less than $n^{\frac{1}{2}+\epsilon}$
    - By Chernoff inequality:
      - $P[$ number of balls in one bin $\geq 2\log n] \leq \frac{e^{np(e^t-1)}}{n^{2t}}$
      - $P[$ maximum number of balls in all bins $\geq 2\log n] \leq \frac{e^{np(e^t-1)}}{n^{2t-1}}$
      - when t is a constant ≥ 0.5 and $n \to \infty$, the maximum number of balls should less than $2\log n$

5 Law of Large Numbers

- $\{X_i\}_{i=1}^{\infty}$ is a sequence of pairwise uncorrelated random variable with $\mathbb{E}[X_i] = m$, $Var(X_i) = \sigma_i^2$.
- $M_n = \frac{1}{n}\sum_{i=1}^{n} X_i$
- $M_n \to m$ almost surely, in mean square and in probability.

6 Memoryless

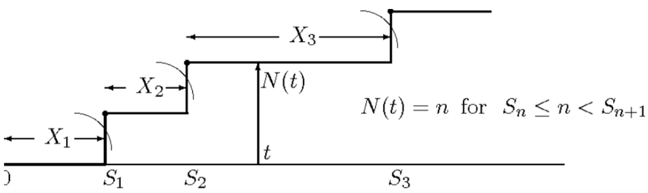
- Definition: $P[X > x_1 + x_2 | X > x_1] = P[X > x_2]$
- Property:
  - Exponential random variable is the only continuous memoryless random variable
  - Bernoulli random variable is the only discrete memoryless random variable

7 Famous Random Variable

- Poisson:
  $P[X = k] = \frac{\lambda^k}{k!} \exp(-\lambda)$
  $\mathbb{E}[X] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} \exp(-\lambda) = \sum_{k=0}^{\infty} \lambda \frac{\lambda^{k-1}}{(k-1)!} \exp(-\lambda) = \lambda$
  Interpretation:
  - Cut total time into infinite period in Binomial random variable, $n \to \infty, p \to \frac{\lambda}{n}$
  - $\to P[X = k] = \lim_{n\to\infty} \binom{n}{k} (\frac{\lambda}{n})^k (\frac{n-\lambda}{n})^{n-k} = \frac{\lambda^k}{k!}(1 - \frac{\lambda}{n})^n = \frac{\lambda^k}{k!} \exp(-\lambda)$
- Erlang:
  $f_X(x) = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!}, \forall x \in \mathbb{R}$
  $\mathbb{E}[X] = \frac{n}{\lambda}$
  Interpretation:
  - Suppose $X_1, X_2, ..., X_n$ are i.i.d exponential random variable with λ.
  - $X = \sum_{i=1}^{n} X_i$
  - Proof by induction:
    Suppose $n = 2$, $f_X(x) = \int_0^x \lambda e^{-\lambda t} \lambda e^{-\lambda(x-t)} dt = \lambda^2 x e^{-\lambda x}$

8 Stochastic Processes

- Stochastic Process: a collection of random variable
  Arrival Process: a sequence of arriving event in continuous time



  - X_i: the time between the i-th event and the $i-1$-th event
  - S_i: the time from start to i-th event
  - $N(t)$: the number of the arrived event at time t
  - X and S Relation:
    * $X_1 = S_1, X_i = S_i - S_{i-1}$
  - N and S Relation:
    * $N(t) < n \leftrightarrow S_{n+1} > t$
    * $N(t) \geq n \leftrightarrow S_n \leq t$
    * $N(t) = n \leftrightarrow S_n \leq t < S_{n+1}$
    * $N(t) = \max\{n : S_n \leq t\}$

– Renewal Process: an arrival process with i.i.d X_i

Delayed Renewal Process: the process becomes a renewal process after several arrivals

X_i Property

* if X_i is dependent on the interval states, then X_i might be dependent on $X_{i-1} \to$ not renewal process

S_i Property

* $P[\lim_{n\to\infty} S_n = \infty] = 1$
  Proof: $\lim_{n\to\infty} P[S_n = \infty] = \lim_{n\to\infty} P[\sum_{i=1}^n X_n = n \times \mathbb{E}[X_i]] = 1$
  Interpretation: infinite events do not take finite time

$N(t)$ Property

* for any $t, P[N(t) < \infty] = 1$
  Proof: $P[\lim_{n\to\infty} S_n = \infty] = 1 \to$ for any $t, P[\lim_{n\to\infty} S_{n+1} > t] = 1$
  Interpretation: infinite events do not take finite time
* $P[\lim_{t\to\infty} N(t) \to \infty] = 1$
  Proof: if $P[\lim_{t\to\infty} N(t) = k] > 0 \to P[X_{k+1} = \infty] > 0$
  Interpretation: finite events do not take infinite time
* $P[\lim_{t\to\infty} \frac{N(t)}{t} = \frac{1}{\mathbb{E}[X_i]}] = 1$
  Proof: $P[\lim_{t\to\infty} \frac{N(t)}{S_{N(t)+1}} \le \lim_{t\to\infty} \frac{N(t)}{t}] = 1$ and $P[\lim_{t\to\infty} \frac{N(t)}{S_{N(t)+1}} = \frac{1}{\mathbb{E}[X_i]}] = 1$
  $P[\lim_{t\to\infty} \frac{N(t)}{t} \le \lim_{t\to\infty} \frac{N(t)}{S_{N(t)}}] = 1$ and $P[\lim_{t\to\infty} \frac{N(t)}{S_{N(t)}} = \frac{1}{\mathbb{E}[X_i]}] = 1$

Inspection Paradox

* $\mathbb{E}[X_{N(t)+1}] \ge \mathbb{E}[X_i]$: inspection paradox
  Interpretation: when selecting t with equal probability, we tend to choose X_i with longer period
* $P[\lim_{t\to\infty} \frac{1}{t}\int_0^t (S_{N(t)+1} - s)ds = \frac{\mathbb{E}[X_i^2]}{2\mathbb{E}[X_i]}] = 1$
  Proof:
  $P[\lim_{t\to\infty} \frac{1}{t}\sum_{i=i}^{N(t)} \frac{\mathbb{E}[X_i^2]}{2} \le \lim_{t\to\infty} \frac{1}{t}\int_0^t (S_{N(t)+1} - s)ds] = 1$ and $P[\lim_{t\to\infty} \frac{1}{t}\sum_{i=i}^{N(t)} \frac{\mathbb{E}[X_i^2]}{2} = \frac{\mathbb{E}[X_i^2]}{2\mathbb{E}[X_i]}] = 1$
  $P[\lim_{t\to\infty} \frac{1}{t}\int_0^t (S_{N(t)+1} - s)ds \le \lim_{t\to\infty} \frac{1}{t}\sum_{i=i}^{N(t)+1} \frac{\mathbb{E}[X_i^2]}{2}] = 1$ and $P[\lim_{t\to\infty} \frac{1}{t}\sum_{i=i}^{N(t)+1} \frac{\mathbb{E}[X_i^2]}{2} = \frac{\mathbb{E}[X_i^2]}{2\mathbb{E}[X_i]}] = 1$
* $P[\lim_{t\to\infty} \frac{1}{t}\int_0^t (s - S_{N(t)})ds = \frac{\mathbb{E}[X_i^2]}{2\mathbb{E}[X_i]}] = 1$
  Proof: similar to above
* $P[\lim_{t\to\infty} \frac{1}{t}\int_0^t X_{N(t)}ds = \frac{\mathbb{E}[X_i^2]}{\mathbb{E}[X_i]}] = 1$
  Proof: $P[\lim_{t\to\infty} \frac{1}{t}\int_0^t X_{N(t)}ds = \lim_{t\to\infty} \frac{1}{t}\int_0^t (S_{N(t)+1} - S_{N(t)})ds] = 1$
* $\mathbb{E}[X_{N(t)+1}] = \frac{\mathbb{E}[X_i^2]}{\mathbb{E}[X_i]}$
  Proof: $P[\lim_{t\to\infty} \frac{1}{t}\int_0^t X_{N(t)}ds = \frac{\mathbb{E}[X_i^2]}{\mathbb{E}[X_i]}] = P[\mathbb{E}[X_{N(t)+1}] = \frac{\mathbb{E}[X_i^2]}{\mathbb{E}[X_i]}] = 1$

Central Limit Theorem

* $\mu = \mathbb{E}[X_i]$
* $\sigma = \sqrt{Var(X_i)}$
* $Z \sim$ Normal(0,1)
* $\lim_{t\to\infty} P[N(t) \le \frac{t}{\mu} + k\frac{\sigma\sqrt{t}}{\sqrt{\mu^3}}] = P[Z \le k]$
  Proof:
  1. Suppose $n(t) = \frac{t}{\mu} + k\frac{\sigma\sqrt{t}}{\sqrt{\mu^3}}$
  2. $P[N(t) \ge n(t)] = P[S_{n(t)} \le t] = P[\frac{S_{n(t)} - n\mu}{\sigma\sqrt{n}} \le \frac{t - n\mu}{\sigma\sqrt{n}}]$.
  3. When $t \to \infty$, $\frac{t-n\mu}{\sigma\sqrt{n}} \to k$
  4. By law of large number, $\lim_{t\to\infty} P[\frac{S_{n(t)} - n\mu}{\sigma\sqrt{n}} \le k] = P[Z \le k]$

  Interpretation:
  · $\frac{t}{\mu}$ is approximately the mean of $N(t)$
  · $k\frac{\sigma\sqrt{t}}{\sqrt{\mu^3}}$ is $k\sigma\sqrt{n}$ after dividing by μ, the ratio between t and $N(t)$ and changing n with $\frac{t}{\mu}$

Wald's Identity

* Stopping Times: a random variable τ s.t. $\{\tau = n\}$ is independent of $\{X_i\}_{i=n+1}^{\infty}$
* Stopping Condition: a condition to stop if we can consider $\tau = \min\{n : \text{ condition}(n) = \top\}$
* Example: $N(t) + 1$ is a stopping times and can be consider $N(t) + 1 = \min\{n : S_n > t\}$
* $\mathbb{E}[\sum_{i=1}^{\tau} X_i] = \mathbb{E}[\tau]\mathbb{E}[X_i]$ if $\mathbb{E}[X_i] < \infty$ and $\mathbb{E}[N] < \infty$
  Proof:
  1. $\mathbb{E}[\sum_{i=1}^{\tau} X_i] = \sum_{i=1}^{\infty} \mathbb{E}[X_i \times \mathbb{1}_{i\le\tau}]$ (by Fubin's Theorem without proof)
     (if $\mathbb{E}[X_i] < \infty$ and $\mathbb{E}[N] < \infty$)
  2. $\sum_{i=1}^{\infty} \mathbb{E}[X_i \times \mathbb{1}_{i\le\tau}] = \mathbb{E}[X_i] \sum_{i=1}^{\infty} \mathbb{E}[\mathbb{1}_{i\le\tau}]$ (by $P[\tau \ge i] = 1 - P[\tau < i]$ is independent of X_i)
  3. $\mathbb{E}[X_i] \sum_{i=1}^{\infty} \mathbb{E}[\mathbb{1}_{i\le\tau}] = \mathbb{E}[\tau]\mathbb{E}[X_i]$
* $\lim_{t\to\infty} \frac{\mathbb{E}[N(t)]}{t} = \frac{1}{\mathbb{E}[X_i]}$
  Proof:
  · Suppose $\mu = \mathbb{E}[X_i]$
  · $\frac{\mathbb{E}[N(t)]}{t} = \frac{\mathbb{E}[S_{N(t)+1}]}{t\times\mu} - \frac{1}{t}$ (by considering $N(t) + 1$ as the stopping time)
  · $\lim_{t\to\infty} \frac{\mathbb{E}[N(t)]}{t} \ge \frac{1}{\mu}$ (by $\mathbb{E}[S_{N(t)+1}] > t$)
  · Suppose $\hat{X}_n = \min\{X_n, T\}$, where T is a constant
  · $\frac{\mathbb{E}[N(t)]}{t} \le \frac{\mathbb{E}[\hat{N}(t)]}{t} = \frac{\mathbb{E}[S_{\hat{N}(t)+1}]}{t\times\hat{\mu}} - \frac{1}{t} \le \frac{t+T}{t\times\hat{\mu}} - \frac{1}{t}$
  · $\lim_{n=\sqrt{t},t\to\infty} \frac{\mathbb{E}[N(t)]}{t} \le \frac{1}{\mu}$

Blackwell's Theorem

* $\mathbb{E}[N(t)] = F_{X_i}(t) + \int_0^t \mathbb{E}[N(t-x)]f_{X_i}(t)dt$
  Proof: $\mathbb{E}[N(t)] = \int_0^t \mathbb{E}[N(t)|X_1 = x]f_{X_1}(x)dx$
  $= \int_0^t \mathbb{E}[N(t-x)+1]f_{X_1}(x)dx = F_{X_i}(t) + \int_0^t \mathbb{E}[N(t-x)]f_{X_i}(t)dt$
* $\mathcal{L}\{\mathbb{E}[N(t)]\}(s) = \frac{\mathcal{L}\{f_{X_i}\}(s)}{s(1-\mathcal{L}\{f_{X_i}\}(s))}$
  Proof: Laplace transform both sides
* Lattice/ Non-Lattice: $N(t)$ is lattice iff X_i only takes on values that are $nd, n \in \mathbb{N}, d \in \mathbb{R}$
* For a non-lattice process: $\lim_{t\to\infty} \mathbb{E}[N(t+\delta) - N(t)] = \frac{\delta}{\mathbb{E}[X_i]}$, for any δ
  Proof: Without Proof
  Interpretation: $\mathbb{E}[N(t)]$ will converge to be linear
* For a lattice process and period d: $\lim_{n\to\infty} \mathbb{E}[\#\text{ events at } t = nd] = \frac{d}{\mathbb{E}[X_i]}$
  Proof: Without Proof
  Interpretation: $\mathbb{E}[N(t)]$ will converge to be stairs with width d and height $\frac{d}{\mathbb{E}[X_i]}$

– Renewal-Reward Process:

Definition

* A renewal process $N(t)$ and $\{R_i\}_{i=1}^{\infty}$ such that (X_i, R_i) are i.i.d.
  ($X_i, R_j, i \ne j$ are independent, but X_i, R_i might be dependent)

Property

* $P[\lim_{t\to\infty} \frac{1}{t} \sum_{i=1}^{N(t)} R_i = \frac{\mathbb{E}[R_i]}{\mathbb{E}[X_i]}] = 1$
  Proof: $P[\lim_{t\to\infty} \frac{1}{t} \sum_{i=1}^{N(t)} R_i = \lim_{t\to\infty} \sum_{i=1}^{N(t)} \frac{R_i}{N(t)} \times \lim_{t\to\infty} \frac{N(t)}{t}] = 1$

– Poisson Process: a renewal process with $X_i \sim \text{Exponential}(\lambda)$

S_i Property

* S_i is an Erlang random variable
  Erlang is the sum of the Exponential random variables
* Joint Distribution $f_{S_1,\ldots,S_n}(s_1, \ldots, s_n) = \lambda^n e^{-\lambda s_n}$
  Prove by induction.
  Induce by $f_{S_1,\ldots,S_n}(s_1, \ldots, s_n) = f_{S_1,\ldots,S_{n-1}}(s_1, \ldots, s_{n-1}) \times f_{S_n|S_1,\ldots,S_{n-1}}(s_n, s_1, \ldots, s_{n-1})$

$N(t)$ Property

* $N(t) \sim \text{Poisson}(\lambda t), P[N(t) = n] = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$
  Prove by $P[N(t) = n] = P[S_n \le t \text{ and } S_{n+1} > t]$

* Conditioned on $N(t) = n$, the set of arrival times $\{s_1, \dots, s_n\}$ have the same distribution with a set of n sorted i.i.d. Uniform$(0, t)$ random variables
  Prove by $f_{S_1,\dots,S_n|N(t)}(s_1, \dots, s_n, n) = \frac{f_{S_1,\dots,S_n}(s_1,\dots,s_n)P[X_{n+1}>t-s_n]}{P[N(t)=n]} = \frac{n!}{t^n}$

Property

* Z is the interval from t to the first arrival $\to Z$ is exponential random variable with same λ and independent of $N(t)$ and the arrival time before t
  Proof:
  $P[Z > z] = \sum_{n=0}^{\infty} \int_0^{\infty} \cdots \int_0^{\infty} P[Z > z|N(t) = n, S_1 = s_1, \dots, S_n = s_n]ds_1 \dots ds_n$
  $= \sum_{n=0}^{\infty} \int_0^{\infty} \cdots \int_0^{\infty} P[X_{n+1} > z + t - s_n|N(t) = n, S_1 = s_1, \dots, S_n = s_n]ds_1 \dots ds_n$
  $= \sum_{n=0}^{\infty} \int_0^{\infty} \cdots \int_0^{\infty} P[X_{n+1} > z + t - s_n|X_{n+1} > t - s_n]ds_1 \dots ds_n = e^{-\lambda z}$
* Stationary Increments: $N(t_1 + t_2) - N(t_1)$ and $N(t_2)$ share the same distribution
  Without Proof
* Independent Increments: $\forall 0 < t_1 < t_2 < \dots, t_k, N(t_1), N(t_2) - N(t_1), \dots$ are independent
  Without Proof
* Any arrival process with stationary and independent increments must be a Poisson process
  Without Proof

Exercise

* $\mathbb{E}[S_i|N(t) = n] = \frac{t \times i}{n+1}$
  - $\mathbb{E}[S_i|N(t) = n] = i \times \mathbb{E}[X_1|N(t) = n] = i \int_0^t \int_0^{s_n} \cdots \int_0^{s_2} s_1 \times \frac{n!}{t^n} ds_1 \dots ds_{n-1}ds_n = \frac{t \times i}{n+1}$
* $\mathbb{E}[\sum_{i=0}^{N(t)} S_i] = \frac{\lambda t^2}{2}$
  - $\mathbb{E}[\sum_{i=0}^{N(t)} S_i] = \sum_{n=0}^{\infty} \mathbb{E}[\sum_{i=0}^{n} S_i|N(t) = n]P[N(t) = n]$
    $= \sum_{n=0}^{\infty} \frac{nt}{2} P[N(t) = n] = \frac{\lambda t^2}{2}$

2D Poisson Process

* Definition:
  - For any region R: number of points in R is a Poisson random variable
  - number of points in the non-overlapping region is independent

Combining Poisson Process

* $N^1(t), N^2(t)$ are two independent Poisson process with λ_1, λ_2
* X_i is the first arrival of X_i^1, X_i^2
* Property
  - X_i is independent of $\{X_i^1 < X_i^2\}$ and $\{X_i^1 > X_i^2\}$
    Proof: $P[X_1^1 < X_1^2] = \frac{\lambda_1}{\lambda_1+\lambda_2}$
    $P[X_1 > x] = P[X_1^1 > x, X_1^2 > x] = e^{-(\lambda_1+\lambda_2)x}$
    $P[X_1 > x, X_1^1 < X_1^2] = P[X_1 > x]P[X_1^1 < X_1^2]$
  - X_i is a Poisson Process with $\lambda = \lambda_1 + \lambda_2$

Splitting Poisson Process

* $N^1(t), N^2(t)$ are two independent Poisson process with λ_1, λ_2
* $N(t)$ is a random process with $\lambda = \lambda_1 + \lambda_2$
  - $N^{1*}(t)$ is the process of the first event
    when $N(t)$ arrives consider it as first event with probability $\frac{\lambda_1}{\lambda_1+\lambda_2}$
  - $N^{2*}(t)$ is the process of the second event
    when $N(t)$ arrives consider it as second event with probability $\frac{\lambda_2}{\lambda_1+\lambda_2}$
* $N^i(t)$ and $N^{i*}(t)$ share the same distribution
* Proof:
  - $B_n(k)$ is a Binomial random variable with $p = \frac{\lambda_1}{\lambda_1+\lambda_2}$
  - $P[N^{1*}(t) = m, N^{2*}(t) = n] = P[N(t) = m + n, B_{m+n}(m)] = P[N^1(t) = m, N^2(t) = n]$

Compound Poisson Process

* $N(t)$ is a Poisson Process
* A_n is a sequence of cost

- $A(t) = \sum_{n=0}^{N(t)} A_n$ is the summation of cost over Poisson Process

Non-Homogeneous Poisson Process

- $N(t) - N(s) \sim \text{Poisson}(\int_s^t \lambda(x)dx$

Queuing Theory

- Definition: $Arrival\_Process/Service\_Process/number\_of\_services$
  - M: memoryless (Poisson) process
  - D: deterministic process
  - G: general renewal process
- T: the random variable of the processing time for each customer
- $Y(t)$: number of cutomers in the service
  - $Y(t) \sim \text{Poisson}(\lambda \int_0^t P[T > x]dx)$
  - Proof:
    Consider $Y(t)$ is a splitting Poisson Process. Since the distribution for the arrival given $N(t)$ is universal, the probability the arrival is still in service: $\frac{1}{t}\int_0^t P[T > t-x]dx = \frac{1}{t}\int_0^t P[T > x]dx$

9 Markov Chain

- Definition
  - Model with states and transition probability matrix
  - States: $\{X_i\}_{i=0}^{\infty}$
  - Transition Probability Matrix: $[P]_{ij} = P[X_{n+1} = j|X_n = i]$
- Terminology
  - p^n: distribution at step n
  - $T_i = \min\{n \geq 1 : X_n = i\}$
  - $f_{ij} = P[T_j < \infty|X_0 = i]$: the probability of starting at i and ever reaching j
  - $\mu_{ij} = \mathbb{E}[T_j|X_0 = i]$
  - $i \to j$ iff $f_{ij} > 0$: j is reachable from i with probability greater than 0
  - $N_i(n)$: number of visits to i by time n
  - Irreducible: $i \leftrightarrow j, \forall$ states i, j
  - aperiodic: period of $X_n = i$ is 1, $\forall$ states i
- Property
  - Updating distribution
    - $p^n = p^0 P$
  - Markovian: transition probability depend only on current state
    - $P[X_{n+1} = j|X_n = i, \ldots, X_0 = x_0] = [P]_{ij}$
  - Stationary Distribution: p

    Property from renewal process
    - consider $X_n = j$ as a event $\to$ Markov Chain becomes a delayed renewal process
    - If $i \leftrightarrow j$ and the model starts from i, then following holds
    - $P[\lim_{n\to\infty} \frac{N_j(n)}{n} = \frac{1}{\mu_{jj}}] = 1$
    - $\lim_{n\to\infty} \frac{\mathbb{E}[N_j(n)]}{n} = \frac{1}{\mu_{jj}}$
    - if the period of $X_n = j$ is $d \to \lim_{n\to\infty} p_j^{nd} = \frac{d}{\mu_{jj}}$

    Theorem of an irreducible, aperiodic Markov Chain
    - Either
      - All states have $\mu_{ii} = \infty$

- All states have $\mu_{ii} < \infty$ and $p_i = \frac{1}{\mu_{ii}}$ is the unique stationary distribution

* Proof
  - From if the period of $X_n = j$ is $d \to \lim_{n\to\infty} p_j^{nd} = \frac{d}{\mu_{jj}}$

Theorem of an irreducible, aperiodic Markov Chain

* All states have $\mu_{ii} < \infty$ and $p_i = \frac{1}{\mu_{ii}}$ is the unique stationary distribution

p can be calculated as the eigenvector corresponds to eigenvalue 1 of P^T

- Hidden Markov Chain
  - Definition: output is a function of the state
  - Interpretation: if the model is not markovian, then reformulate the model as a hidden markov chain by complicating the states and rendering the output as a function of the state