Stochastic Processes

Kevin Chang

April 13, 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:
    * $\mathbb{E}[e^{tX}] = \int_{-\infty}^\infty e^{tx}f_X(x)dx$
    * $\mathbb{E}[e^{tX}] = \sum_{k=0}^\infty E[X^k]\frac{t^k}{k!}$
      · $e^{tx} = \sum_{k=0}^\infty \frac{(tx)^k}{k!}$
      · $E[e^{tX}] = E[\sum_{k=0}^\infty \frac{(tX)^k}{k!}] = \sum_{k=0}^\infty E[X^k]\frac{t^k}{k!}$
    * $\frac{d\mathbb{E}[e^{tX}]}{dt} = \mathbb{E}[X]$
    * $\mathbb{E}[e^{t(aX+b)}] = e^t b\mathbb{E}[e^{taX}]$
    * 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 $X_1, \ldots, X_i, \ldots, X_N$ 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
    * $\mathbb{E}[X] = \int_0^{\infty} x f_X(x)dx = \int_0^{\infty} \int_0^{x} f_X(x)dydx = \int_0^{\infty} \int_y^{\infty} f_X(x)dxdy = \int_0^{\infty} P[X > y]dy$

4 Inequality

- Markov Inequality

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

  Proof:
  1. $\mathbb{E}[X] = \int_0^{\infty} x f_X(x) \geq \int_{\epsilon}^{\infty} x f_X(x) \geq \epsilon \int_{\epsilon}^{\infty} f_X(x) = \epsilon P[X \geq \epsilon]$
  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

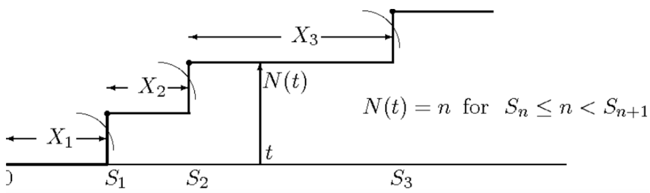
- Definition: $P[X > x_1 + x_2 | X > x_1] = P[X > x_2]$
- Property:
  - Exponential random variable is the only continuous memoryless random variable
  - Bernoulli random variable is the only discrete memoryless random variable

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, $n \to \infty, p \to \frac{\lambda}{n}$
  - $\to P[X = k] = \lim_{n\to\infty} \binom{n}{k} (\frac{\lambda}{n})^k (\frac{n-\lambda}{n})^{n-k} = \frac{\lambda^k}{k!}(1 - \frac{\lambda}{n})^n = \frac{\lambda^k}{k!} \exp(-\lambda)$
- Gaussian: $N(m, \sigma^2)$
  - $f_X[x] = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-m)^2}{2\sigma^2}}, \forall x \in \mathbb{R}$
  - $\mathbb{E}[e^{cX}] = e^{cm + \frac{c^2\sigma^2}{2}}$
- 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 $X_{i-1} \rightarrow$ not renewal process

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

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

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

  Central Limit Theorem
    * $\mu = \mathbb{E}[X_i]$
    * $\sigma = \sqrt{Var(X_i)}$
    * $Z \sim$ Normal(0,1)
    * $\lim_{t\rightarrow\infty} P[N(t) \leq \frac{t}{\mu} + k\frac{\sigma\sqrt{t}}{\sqrt{\mu}^3}] = P[Z \leq k]$
      Proof:
        1. Suppose $n(t) = \frac{t}{\mu} + k\frac{\sigma\sqrt{t}}{\sqrt{\mu}^3}$
        2. $P[N(t) \geq n(t)] = P[S_{n(t)} \leq t] = P[\frac{S_{n(t)} - n\mu}{\sigma\sqrt{n}} \leq \frac{t - n\mu}{\sigma\sqrt{n}}]$.

3. When $t \to \infty$, $\frac{t-n\mu}{\sigma\sqrt{n}} \to k$
4. By law of large number, $\lim_{t\to\infty} P[\frac{S_{n(t)}-n\mu}{\sigma\sqrt{n}} \leq k] = P[Z \leq k]$

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. $\mathbb{E}[\sum_{i=1}^{\tau} X_i] = \sum_{i=1}^{\infty} \mathbb{E}[X_i \times \mathbb{1}_{i\leq\tau}]$ (by Fubin's Theorem without proof)
     (if $\mathbb{E}[X_i] < \infty$ and $\mathbb{E}[N] < \infty$)
  2. $\sum_{i=1}^{\infty} \mathbb{E}[X_i \times \mathbb{1}_{i\leq\tau}] = \mathbb{E}[X_i]\sum_{i=1}^{\infty} \mathbb{E}[\mathbb{1}_{i\leq\tau}]$ (by $P[\tau \geq i] = 1 - P[\tau < i]$ is independent of X_i)
  3. $\mathbb{E}[X_i]\sum_{i=1}^{\infty} \mathbb{E}[\mathbb{1}_{i\leq\tau}] = \mathbb{E}[\tau]\mathbb{E}[X_i]$
* $\lim_{t\to\infty} \frac{\mathbb{E}[N(t)]}{t} = \frac{1}{\mathbb{E}[X_i]}$

  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)

  · $\lim_{t\to\infty} \frac{\mathbb{E}[N(t)]}{t} \geq \frac{1}{\mu}$ (by $\mathbb{E}[S_{N(t)+1}] > t$)

  · Suppose $\hat{X}_n = \min\{X_n, T\}$, where T is a constant

  · $\frac{\mathbb{E}[N(t)]}{t} \leq \frac{\mathbb{E}[\hat{N}(t)]}{t} = \frac{\mathbb{E}[S_{\hat{N}(t)+1}]}{t\times\hat{\mu}} - \frac{1}{t} \leq \frac{t+T}{t\times\hat{\mu}} - \frac{1}{t}$

  · $\lim_{n=\sqrt{t}, t\to\infty} \frac{\mathbb{E}[N(t)]}{t} \leq \frac{1}{\mu}$

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

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

Exercise

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

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 $\lambda = \lambda_1 + \lambda_2$
  · $\min(X_1, X_2)$ is an exponential random variable 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
        * $N(t) - N(s) \sim$ 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
            · $Y(t) \sim$ 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: $\{X_n\}_{n=0}^{\infty}$
    - Transition Probability Matrix: $[P]_{ij} = P[X_{n+1} = j|X_n = i]$
- Terminology
    - $p^n = [P[X_n = 0], P[X_n = 1], \dots]^T$: distribution at step n
    - $T_i = \min\{n \geq 1 : X_n = i\}$: a random variable of the minimum time step to go to state 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
    - Consider a given distribution as an event $\tau : [P[X_n = 0|\tau], P[X_n = 1|\tau], \dots]^T$
    - Updating distribution
        * $p^n = p^0P^n$
    - Markovian: transition probability depend only on current state
        * $P[X_{n+1} = j|X_n = i, \dots, X_0 = x_0] = [P]_{ij}$
    - Transient and Recurrant of state i
        * Transient: if $f_{ii} < 1$
        * Null Recurrant: if $f_{ii} = 1$ and $\mu_{ii} = \infty$
        * Positive Recurrant: if $f_{ii} = 1$ and $\mu_{ii} < \infty$

    - Markov Chain with transient or null recurrant state $\rightarrow$ no limiting distribution exists
- Stationary Distribution: p s.t. if $p^n = p \rightarrow p^{n+1} = p$
  Property from renewal process
    - consider $X_n = j$ as a event $\rightarrow$ Markov Chain becomes a delayed renewal process
    - If $i \leftrightarrow j$ and the model starts from i, then following holds
    - $P[\lim_{n\rightarrow\infty} \frac{N_j(n)}{n} = \frac{1}{\mu_{jj}}] = 1$
    - $\lim_{n\rightarrow\infty} \frac{\mathbb{E}[N_j(n)]}{n} = \frac{1}{\mu_{jj}}$
    - if the period of $X_n = j$ is $d \rightarrow \lim_{n\rightarrow\infty} p_j^{nd} = \frac{d}{\mu_{jj}}$

  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 $d \rightarrow \lim_{n\rightarrow\infty} p_j^{nd} = \frac{d}{\mu_{jj}}$
        Proof: $\lim_{n\rightarrow\infty} p_j^{nd} = \lim_{n\rightarrow\infty} \mathbb{E}[\#$ events at $nd]$

  Theorem of an finite irreducible, aperiodic Markov Chain
    - All states have $\mu_{ii} < \infty$ and $p_i = \frac{1}{\mu_{ii}}$ is the unique stationary distribution

  Property
    - p can be calculated as the eigenvector corresponds to eigenvalue 1 of P^T
    - p satisfy $p_i \sum_{j\neq i} R_{ij} = \sum_{j\neq i} p_j R_{ji}$: sum of out-distribution equals sum of in-distribution
- Detailed Balance
  Definition:
    - Given a distribution π
    - $\pi_i P_{ij} = \pi_j P_{ji}, \forall i, j$

  Property:
    - distribution π satisfying Detailed Balance is the stationary distribution p
    - symmetric transition probability matrix $\rightarrow$ uniform stationary distribution
- Reversible
  Definition: A Markov Chain with stationary distribution p is reversible if it satisfies detailed balance
  Interpretation
    - Transitions forward and backward in the stationary distribution have the same probability
    - $P[X_{n+1} = j | X_n = i] = P_{ij}$
    - $P[X_{n-1} = j | X_n = i] = \frac{P[X_{n-1}=j, X_n=i]}{P[X_n=i]} = \frac{p_j P_{ji}}{p_i} = P_{ij}$
- Metropolis Update Rule
  Definition
    - Given a Markov Chain and distribution p', find P' such that p' is the stationary distribution

  Procedure
    - For each pair (i, j), $P'_{ij} = P_{ij} \times \min\{1, \frac{p'_j P_{ji}}{p'_i P_{ij}}\}$
    - construct self loop to satisfy $\sum_j P'_{ij} = 1$

  Proof
    - To satisfy detailed balance, for each pair (i, j), we should set $p'_i P'_{ij} = \min\{p'_i P_{ij}, p'_j P_{ji}\}$
- Distance between Probability Measure
  Definition:
    - Total Variation Distance between P_1 and P_2 is: $d_{TV}(P_1, P_2) = \frac{1}{2} \sum_\omega |P_1[\omega] - P_2[\omega]|$

  Interpretation:
    - consider the distributions as events τ_1, τ_2

    * $P_i[\omega] = P[\omega|\tau_i]$
    * $d_{TV}(P_1, P_2) = \frac{1}{2}\sum_\omega |P[\omega|\tau_1] - P[\omega|\tau_2]| = \sum_\omega |P[\omega \wedge \tau_1] - P[\omega \wedge \tau_2]|$
  - Mixing Time
    Definition
    * Mixing time τ is the least t such that for all initial state p^0, $d_{TV}(p, p^0P^t) \leq \frac{1}{2e}$

    Interpretation
    * the factor $\frac{1}{2e}$ is set such that $d_{TV}(p, p^0P^t) \leq \epsilon$ if $t \geq \tau \times \log(\frac{1}{\epsilon})$
      Without proof
  - Example
    Random Walk on Graph
    * Definition: move from vertex i to vertex j with probability $P_{ij} = \begin{cases} 0 & \text{if } (i,j) \notin E \\ \frac{1}{\text{degree}(i)} & \text{if } (i,j) \in E \end{cases}$
    * Distribution π, $\pi_i = \frac{\text{degree}(x)}{2|E|}$ satisfies detailed balance
    * If we want stationary distribution to be uniform $\rightarrow P'_{ij} = \begin{cases} \frac{1}{\text{degree}(i)} & \text{if degree}(i) \geq \text{degree}(j) \\ \frac{1}{\text{degree}(j)} & \text{if degree}(i) < \text{degree}(j) \end{cases}$

    Random graph coloring
    * Given a graph with V vertices, maximum degree Δ and q colors, to color each vertex one color such that adjacent vertex do not share the same color
    * Assume $q > 4\Delta$
    * Markov Chain Transition:
      · Pick random vertex and random color, if the color is changeable then change
    * Property
      · Aperiodic: there exist self loops
      · Symmetric: symmetric transition
      · Irreducible
    * Mixing time is $O(V \log V)$
      Proof:
      · Assume X is a event s.t. Markov Chain starts with any valid coloring and Y is a event s.t. Markov Chain starts with uniform distribution
      · Apply same transition on both X and Y
      · D_n is a random variable for the number of vertices in different colors in X and Y at time n
      · Good moves: number of vertices in different colors decrease $\geq D_n \times (q - 2\Delta) \geq (2\Delta + 1)D_n$
        (vertices with different colors $\times$ color that is different with any adjacent color in X and Y)
      · Bad moves: number of vertices in different colors increase $\leq (D_n\Delta) \times 2$
        (vertices adjacent to different colors vertices $\times$ color of the differen colors vertices)
      · $\mathbb{E}[D_{n+1} - D_n] \leq V(1 - \frac{1}{qV})^n$
      · $\mathbb{E}[D_n] \leq V(1 - \frac{1}{qV})^n$
      · $P[D_n \geq 1] \leq V(1 - \frac{1}{qV})^n$

- 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

10 Continuous Markov Chain

- Interpretation
  - v_i: coefficient of exponential distribution, where time in state i before next step is $\sim$ Exponential(v_i)
- Definition

- Model with states and transition rate matrix
- States: $X(t), \forall 0 \le t < \infty$
- Transition Probability Matrix R

- $P_{ij}(t)$
  - Definition: $P_{ij}(t) = P[X(t) = j | X(0) = i]$
  - Chapman-Kolmogorov Equation
    * Definition: $P(s+t) = P(s) \times P(t)$
    * Proof
      · $P_{ij}(s+t) = P[X(s+t) = j | X(0) = i]$
      $= \sum_k P[X(s+t) = j | X(s) = k, X(0) = i] P[X(s) = k | X(0) = i]$
      $= \sum_k P[X(s+t) = j | X(s) = k] P[X(s) = k | X(0) = i] = \sum_k P_{kj}(t) P_{ik}(s)$
  - Kolmogorov's Differential Equation
    * Forward: $\frac{dP(t)}{dt} = P(t)R$
      Interpretation:
      · Change of distribution at t equals the distribution at $t \times R$
      Proof:
      · $\frac{dP(t)}{dt} = \lim_{\delta \to 0} \frac{P(t+\delta) - P(t)}{\delta} = P(t) \lim_{\delta \to 0} \frac{P(\delta) - P(0)}{\delta} = P(t)R$
    * Backward: $\frac{dP(t)}{dt} = RP(t)$
      Interpretation:
      · Change of distribution at t equals the distribution at $t = 0 \times P(t)$
      Proof:
      · $\frac{dP(t)}{dt} = \lim_{\delta \to 0} \frac{P(t+\delta) - P(t)}{\delta} = \lim_{\delta \to 0} \frac{P(\delta) - P(0)}{\delta} P(t) = RP(t)$
    * Solution: $P(t) = e^{Rt}$

- R
  - Definition:
    * $R_{ij} = \frac{dP_{ij}(t)}{dt}|_{t=0}$
    * $R_{ij} = \begin{cases} -v_i & \text{if } i = j \\ v_i P_{ij} & \text{if } i \neq j \end{cases}$ (if there is no self-transition)
  - Interpretation
    * πR is the change of distribution of π (by Kolmogorov's Differential Equation)
    * simulation by transition from state i to j when $e^{-R_{ij}t}$ event arrives
      Proof
      · $\frac{dP_{ii}(t)}{dt} = R_{ii} P_{ii}(t) \to P_{ii}(t) = e^{-R_{ii}t}$
      · simulate the transition out of state i by $e^{-R_{ii}t}$ and transition to j state by probability $\frac{R_{ij}}{R_{ii}}$ is the same as transition from state i to j when $e^{-R_{ij}t}$ event arrives
      Property
      · Continuous Markov Chain with same R are of the same functionality
  - Property:
    * $\sum_j R_{ij} = 0$: sum of element is a row of R is 0

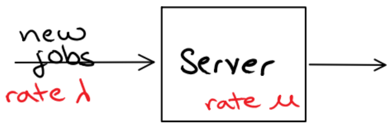
- Property
  - Self Transition:
    * Since R defines the Markov Chain, we can modify v_i to conduct self transition without changing R
  - Uniformization:
    * Since R defines the Markov Chain, we can modify v_i such that v_i are the same for all states without changing R

- Stationary Distribution: p s.t. $pR = 0, pe^{Rt} = p$
  Interpretation:
  * $\frac{dpP(t)}{dt} = p\frac{dP(t)}{dt} = pRP(t) = 0$
  * p is the eigenvector of eigenvalue 0 of R, then p is the eigenvector of eigenvalue 1 of e^{Rt} $\rightarrow$ the distribution would not change, if start with p

  Property
  * $\pi_i \sum_{j\neq i} R_{ij} = \sum_{j\neq i} \pi_j R_{ji}$: sum of out-distribution equals sum of in-distribution

  Trick:
  1. cluster states such that every state in the cluster share the same R_{ij} to use property 1
  2. assume distribution is independent of the cluster and check $pR = 0$ after the calculation
- Poisson process is a special case of Continuous Markov Chain
  * $v_i = \lambda, \forall i$
  * i-th state transition to $i+1$-th state
- Exploding process: only if $v_i \rightarrow \infty$
  * exploding process: traverse infinite states in finite time

• Example
- Queue



  * Stationary Distribution $\pi : \pi_i = (1 - \frac{\lambda}{\mu})(\frac{\lambda}{\mu})^i$
  * For queue with feedback: find the stationary increment frequency λ and process frequency μ then stationary distribution is $\pi : \pi_i = (1 - \frac{\lambda}{\mu})(\frac{\lambda}{\mu})^i$

11 Martingales

• Definition
- Discrete
  General Discrete Martingales
  * $\{Z_i\}_{i=0}^{\infty}$ such that
    1. $\mathbb{E}[|Z_n|] < \infty$
    2. $\mathbb{E}[Z_{n+1}|Z_0, \ldots, Z_n] = Z_n$
       · sub-martingales: $\mathbb{E}[Z_{n+1}|Z_0, \ldots, Z_n] \geq Z_n$
       · super-martingales: $\mathbb{E}[Z_{n+1}|Z_0, \ldots, Z_n] \leq Z_n$

  Discrete Martingales with respect to X_i
  * $\{Z_i\}_{i=0}^{\infty}$ such that
    1. $\mathbb{E}[|Z_n|] < \infty$
    2. $\mathbb{E}[Z_{n+1}|X_0, \ldots, X_n] = Z_n$
       · sub-martingales: $\mathbb{E}[Z_{n+1}|X_0, \ldots, X_n] \geq Z_n$
       · super-martingales: $\mathbb{E}[Z_{n+1}|X_0, \ldots, X_n] \leq Z_n$
  * $\mathbb{E}[Z_{n+1}|X_0, \ldots, X_n] = Z_n$ implies $\mathbb{E}[Z_{n+1}|Z_0, \ldots, Z_n] = Z_n$
    · Z_n is a function of $X_0, \ldots, X_n$
    · $\mathbb{E}[Z_{n+1}|Z_0, \ldots, Z_n] = \mathbb{E}[\mathbb{E}[Z_{n+1}|X_0, \ldots, X_n, Z_0, \ldots, Z_n]|Z_0, \ldots, Z_n]$
      $= \mathbb{E}[\mathbb{E}[Z_{n+1}|X_0, \ldots, X_n]|Z_0, \ldots, Z_n] = \mathbb{E}[Z_n|Z_0, \ldots, Z_n] = Z_n$
- Continuous Martingales with respect to $N(t)$
  * $Y(t)$ such that

1. $\mathbb{E}[|Y(t)|] < \infty$
2. $\mathbb{E}[Y(t)|\{N(s)|0 \le s \le \tau\}] = Y(\tau), \forall \tau \le t$
   · sub-martingales: $\mathbb{E}[Y(t)|\{N(s)|0 \le s \le \tau\}] \ge Y(\tau), \forall \tau \le t$
   · super-martingales: $\mathbb{E}[Y(t)|\{N(s)|0 \le s \le \tau\}] \le Y(\tau), \forall \tau \le t$

- Property
  - $\mathbb{E}[Z_n] = \mathbb{E}[Z_1]$
    Proof: $\mathbb{E}[Z_{n+1} - Z_n] = \mathbb{E}[\mathbb{E}[Z_{n+1} - Z_n|Z_0, \dots, Z_n]] = 0$
  - $\mathbb{E}[Z_n|\{Z_i|i \in S\}] = Z_{\max_{i\in S} i}$, where $\forall i \in S, i < n$
    Proof: $\mathbb{E}[Z_n|Z_i] = \mathbb{E}[\mathbb{E}[Z_n|Z_0, \dots, Z_{n-1}]|Z_i] = \mathbb{E}[Z_{n-1}|Z_i]$
  - Azuma's Inequality
    * $\mu = \mathbb{E}[Z_0]$
    * $-a_i \le Z_i - Z_{i-1} \le b_i$
    * $P[|Z_n - \mu| \ge \delta] \le 2e^{-\frac{2\delta^2}{\sum_{i=1}^{n}(b_i+a_i)^2}}$
  - Kolmogorov's sub-martingales inequality
    * $P[\sup_{n\ge1} Z_n \ge a] \le \frac{\mathbb{E}[Z_1]}{a}$
  - Martingales Stopping Theorem
    * 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 $\tau = \min\{n : \text{ condition}(n) = \top\}$
    * $\mathbb{E}[Z_\tau] = \mathbb{E}[Z_0]$ if the either of the following holds
      1. $P[\tau \le k] = 1$
      2. $P[\max_{i\le\tau} |Z_\tau| \le k] = 1$
      3. $\mathbb{E}[\tau] < k$ and $\mathbb{E}[|Z_{n+1} - Z_n||Z_0, \dots, Z_n] < k$

- Application for generating Martingales
  - Sum of iid. random variables
    * $\{X_i\}_{i=1}^{\infty}$ are iid. random variables
    * $Z_n = \sum_{i=1}^{n} X_i - n\mathbb{E}[X_i]$ is a martingales.
    * Proof: $\mathbb{E}[Z_{n+1}|Z_0, \dots, Z_n] = \mathbb{E}[Z_n + X_{n+1} - \mathbb{E}[X_i]|Z_0, \dots, Z_n] = Z_n$
  - Squre of sum of iid. random variables
    * $\{X_i\}_{i=1}^{\infty}$ are iid. random variables and $\mathbb{E}[X_i] = 0$
    * $Z_n = (\sum_{i=1}^{n} X_i)^2 - n\mathbb{E}[X_i^2]$ is a martingales.
    * Proof: $\mathbb{E}[Z_{n+1}|Z_0, \dots, Z_n] = \mathbb{E}[Z_n + X_{n+1}^2 + 2X_{n+1}(\sum_{i=1}^{n} X_i) - \mathbb{E}[X_i^2]|Z_0, \dots, Z_n] = Z_n$
  - Product of iid. random variables
    * $\{X_i\}_{i=1}^{\infty}$ are iid. random variables
    * $Z_n = \frac{\prod_{i=1}^{n} X_i}{\mathbb{E}[X_i]^n}$ is a martingales.
    * Proof: $\mathbb{E}[Z_{n+1}|Z_0, \dots, Z_n] = \mathbb{E}[Z_n(\frac{X_{n+1}}{\mathbb{E}[X_i]})|Z_0, \dots, Z_n] = Z_n$
  - Poisson Process
    * $N(t)$ is a poisson process
    * $Y(t) = N(t) - \lambda t$ is a martingales.
    * Proof: $\mathbb{E}[Y(t)|\{N(s)|0 \le s \le \tau\}] = \mathbb{E}[Y(\tau) + Y(t) - Y(\tau)|\{N(s)|0 \le s \le \tau\}]$
      $= Y(\tau) + \mathbb{E}[N(t) - N(\tau) + \lambda(t - \tau)|\{N(s)|0 \le s \le \tau\}] = Y(\tau)$
  - Doob-type Martingales
    * $X, \{Y_i\}_{i=1}^{\infty}$ are random variables
    * $Z_n = \mathbb{E}[X|Y_1, Y_2, \dots, Y_n]$ is a martingales
    * Proof: $\mathbb{E}[Z_{n+1}|Y_1, \dots, Y_n] = \mathbb{E}[\mathbb{E}[X|Y_1, Y_2, \dots, Y_n, Y_{n+1}]|Y_1, Y_2, \dots, Y_n]$
      $= \mathbb{E}[X|Y_1, Y_2, \dots, Y_n] = Z_n$

- Example

- Symmetric Random Walk
  * $p = 0.5$
  * $\tau = \min\{i | \sum_{i=0}^{n} X_i \in \{-a, b\}\}$
  * $Z_n = \sum_{i=0}^{n} X_i$, by second rule of Martingales Stopping Theorem: $\mathbb{E}[Z_\tau] = 0$
    $\to P[Z_\tau \text{ at } a] = \frac{b}{a+b}, P[Z_\tau \text{ at } b] = \frac{a}{a+b}$
  * $Z_n = (\sum_{i=0}^{n} X_i)^2 - n$, by third rule of Martingales Stopping Theorem: $\mathbb{E}[Z_\tau] = 0$
    $\to \mathbb{E}[\tau] = ab$
- Unbiased Random Walk
  * $\tau = \min\{i | \sum_{i=0}^{n} X_i \in \{-a, b\}\}$
  * $Z_n = (\frac{1-p}{p})^{\sum_{i=0}^{n} X_i}$, by second rule of Martingales Stopping Theorem: $\mathbb{E}[Z_\tau] = 0$
    $P[Z_\tau \text{ at } a] = \frac{(\frac{1-p}{p})^b - 1}{(\frac{1-p}{p})^b - (\frac{1-p}{p})^{-a}}, P[Z_\tau \text{ at } b] = \frac{1-(\frac{1-p}{p})^{-a}}{(\frac{1-p}{p})^b - (\frac{1-p}{p})^{-a}}$
  * $Z_n = \sum_{i=0}^{n} X_i - n\mathbb{E}[X_0]$, by third rule of Martingales Stopping Theorem: $\mathbb{E}[Z_\tau] = 0$
    $\to \mathbb{E}[\tau] = \frac{\mathbb{E}[\sum_{i=0}^{\tau} X_i]}{\mathbb{E}[X_0]}$

12 Random Walk

- Definition
  - $X_i = \begin{cases} 1 & \text{with probability } p \\ -1 & \text{with probability } 1-p \end{cases}$
  - $S_n = \sum_{i=0}^{n} X_i$
- The monkey at the cliff
  - $P_k = P[\exists n \text{ such that } S_n = k] = \begin{cases} 1 & \text{if } p \geq \frac{1}{2} \\ (\frac{p}{1-p})^k & \text{if } p < \frac{1}{2} \end{cases}$ where $k \in \mathbb{N}$

    Proof
    * $P_k = P_1^k$ by memoryless property
    * $P_1 = p + q \times P_2 \to P_1 = 1$ or $\frac{p}{1-p}$
    * if $p \geq 0.5 \to P_1 = 1$ $(P_1 \leq 1)$
    * if $p < 0.5 \to P_1 = \frac{p}{1-p}$
      Since $P_1 \leq \frac{p}{1-p}$ by induction on n to ∞ for $P_1(n) = P[S_n = k]$
  - $\mathbb{E}_k = \mathbb{E}[\min\{n : S_n = k\}] = \begin{cases} \infty & \text{if } p \leq \frac{1}{2} \\ \frac{k}{2p-1} & \text{if } p > \frac{1}{2} \end{cases}$ where $k \in \mathbb{N}$

    Proof
    * $\mathbb{E}_k = \mathbb{E}_1 \times k$ by memoryless property
    * $\mathbb{E}_1 = 1 + 0 \times p + \mathbb{E}_2 \times (1-p)$
    * if $p < 0.5 \to P_1 = \frac{p}{1-p} \to \mathbb{E}_1 = \infty$
    * if $p = 0.5 \to \mathbb{E}_1 = 1 + \mathbb{E}_1$ (no solution) $\to \mathbb{E}_1 = \infty$
    * if $p > 0.5 \to \mathbb{E}_1 = \frac{1}{2p-1}$
  - $P_0 = P[\exists n \text{ such that } S_n = 0] = 1 - |2p-1|$ where $k \in \mathbb{N}$
    * $P_0 = p \times P_{-1} + (1-p) \times P_1$
  - $\mathbb{E}_0 = \mathbb{E}[\min\{n : S_n = 0\}] = \infty$
    * if $p \neq \frac{1}{2} \to P_0 \neq 1 \to \mathbb{E}_0 = \infty$
    * if $p = \frac{1}{2} \to \mathbb{E}_0 = 1 + \frac{1}{2}\mathbb{E}_{-1} + \frac{1}{2}\mathbb{E}_1 = \infty$
- The Gambler's Ruin
  - Definition: $\tau = \min\{i | S_n \in \{-a, b\}\}$
  - $A_k = P[S_\tau = b | X_0 = k]$
    * $A_k = pA_{k+1} + (1-p)A_{k-1}$

- $A_0 = \begin{cases} \frac{a}{a+b} & \text{if } p = \frac{1}{2} \\ \frac{(\frac{1-p}{p})^a - 1}{(\frac{1-p}{p})^{a+b} - 1} & \text{if } p \neq \frac{1}{2} \end{cases}$
  Solved by previous recursive equation
- $E_k = \mathbb{E}[\tau | X_0 = k]$
  * $E_k = 1 + pE_{k+1} + (1-p)E_{k-1}$
- $E_0 = \begin{cases} ab & \text{if } p = \frac{1}{2} \\ \frac{a}{1-2p} - \frac{a+b}{1-2p} \times \frac{(\frac{1-p}{p})^a - 1}{(\frac{1-p}{p})^{a+b} - 1} & \text{if } p \neq \frac{1}{2} \end{cases}$
  Solved by previous recursive equation

- Observation
  - $S_n = O(n)$
    Upperbound: $\lim_{n\to\infty} P[S_n \leq k\sqrt{n}] = \int_{-\infty}^{k} \frac{1}{2\pi} e^{-\frac{x^2}{2}} dx$
    Lowerbound: $P[|S_n| \geq k\sqrt{n}] \leq 2e^{-\frac{k^2}{2}}$

13 Brownian Motion

- Standard Brownian Motion
  - Interpretation: generalize discrete time and space of random walk to be in continuous time and space
    * $S_t = \delta_x (\sum_{i=0}^{\frac{t}{\delta_t}} X_i)$
    * let $\delta_x = \sqrt{\delta_t}$ and $\delta_x \to 0$
    * $\mathbb{E}[S_t] = 0$
    * $Var(S_t) = t$
      · $Var(S_t) = \delta_x^2 \frac{t}{\delta_t} = t$
  - Definition:
    * $X(0) = 0$
    * $X(t) \sim N(0, \sigma^2 = t)$
    * $X(t)$ has independent, stationary increment
      · independent: $X(t_{i_2}) - X(t_{i_1})$ and $X(t_{i_1}) - X(t_{i_0})$ are independent
      · stationary: $X(s+t) - X(t) = X(s)$
  - Property
    * Distribution self-similarity
      · $X(t) \sim N(0, t)$
      · $\sqrt{k} X(\frac{t}{k}) \sim N(0, t)$
    * Nowhere Differentiable
      · With probability 1, $X(t)$ is nowhere differentiable
      · $\lim_{\delta_t \to 0} \frac{X(t+\delta_t) - X(t)}{\delta_t} = \lim_{\delta_t \to 0} \frac{N(0, \delta_t}{\delta_t} = \lim_{\delta_t \to 0} N(0, \frac{1}{\delta_t})$
    * Unbounded Variation
      · Length of distance $\to \infty$ in finite time t
      · $\lim_{n\to\infty} \sum_{j=1}^{n} |X(\frac{jt}{n}) - X(\frac{(j-1)t}{n})| = \infty$
        Proof: $\lim_{n\to\infty} \sum_{j=1}^{n} |X(\frac{jt}{n}) - X(\frac{(j-1)t}{n})| = lim_{n\to\infty} \sum_{j=1}^{n} |X(\frac{t}{n})| = n \times \sqrt{\frac{2}{\pi} \frac{t}{n}} = \infty$
    * Hitting Time
      The Gambler's Ruin
      · $\tau = \min\{t \geq 0 : X(t) \in \{-A, B\}\}$
      · $P[X(\tau) = A] = \frac{B}{A+B}, P[X(\tau) = B] = \frac{A}{A+B}$
        Prove by Martingales Stopping Theorem on $X(t)$:
        $\to \mathbb{E}[X(t)] = P[X(\tau) = A]A + P[X(\tau) = B]B = 0$

- $\mathbb{E}[\tau] = AB$
  Prove by Martingales Stopping Theorem on $X(t)^2 - t$:
  $\rightarrow \mathbb{E}[X(t)^2 - t] = P[X(\tau) = A]A^2 + P[X(\tau) = B]B^2 - \mathbb{E}[\tau] = 0$

The monkey at the cliff

- $\tau = \min\{t \geq 0 : X(t) = B\}$
- $P[\tau < \infty] = 1$
  Prove by let $A = -\infty$ in The Gambler's Ruin
- $P[\tau \leq t] = 2P[X(\tau) \geq B]$
  $P[\tau \leq t] = P[\tau \leq t \text{ and } X(t) \geq B] + P[\tau \leq t \text{ and } X(t) < B]$
  $= 2P[\tau \leq t \text{ and } X(t) \geq B] = 2P[X(t) \geq B]$
- $\mathbb{E}[\tau] = \infty$
  Prove by let $A = -\infty$ in The Gambler's Ruin

* Diffusion Equation
  - Forward Diffusion Equation: $\frac{\partial f}{\partial t} = \frac{1}{2}\frac{\partial^2 f}{\partial x^2}$
  - Backward Diffusion Equation: $\frac{\partial f}{\partial t} = \frac{-1}{2}\frac{\partial^2 f}{\partial x^2}$
  - $f(X(t_2) = x|X(t_1) = k)$ satisfies Forward Diffusion equation
  - $f(X(t_2) = k|X(t_1) = x)$ satisfies Backward Diffusion equation
* Martingales
  - $X(t)$ is a martingale
  - $X(t)^2 - t$ is a martingale
  - $e^{cX(t) - \frac{c^2}{2}t}$ is a martingale
* Zeros
  Definition: $P[X(t) = 0, t_0 < t < t_1] = \frac{2}{\pi}\cos^{-1}(\sqrt{\frac{t_0}{t_1}})$
  - Prove by $P[X(t) = 0, t_0 < t < t_1] = \int_{-\infty}^{\infty} f_{X(t_0}(x_1)P[T_{-x} \leq t_1 - t_0]dx_1$

  Property
  - $P[X(t) = 0, 0 < t < t_1] = 1, \forall t_1 > 0$
  - $P[inf\{t > 0 : X(t) = 0\} = 0] = 1$
  - $P[$ there are infinitely many zeros in $[0,t]] = 1$

- Brownian Bridge
  * Definition: the distribution of t_1 given the result of the future $X(t_2)$
  * Property
    $f_{X(t_1)|X(t_2)}(x_1, x_2) = \frac{f_{X(t_1),X(t_2)}(x_1,x_2)}{f_{X(t_2)}(x_2)} \sim N(\frac{t_1}{t_2}x_2, \frac{t_1(t_2-t_1)}{t_2})$
    - let $s = t_2 - t_1$
    - $f_{X(t_1)|X(t_2)}(x_1, x_2) = \frac{f_{X(t_1),X(t_2)}(x_1,x_2)}{f_{X(t_2)}(x_2)}$
      $= \frac{f_{X(t_1),X(s)}(x_1,x_2-x_1)}{f_{X(t_2)}(x_2)}$ (By transformation of 2-D random variables)
      $= \frac{f_{X(t_1)}(x_1)f_{X(s)}(x_2-x_1)}{f_{X(t_2)}(x_2)} = \frac{\frac{1}{\sqrt{2\pi t_1}}e^{\frac{-x_1^2}{2t_1}}\frac{1}{\sqrt{2\pi(t_2-t_1)}}e^{\frac{-(x_2-x_1)^2}{2(t_2-t_1)}}}{\frac{1}{\sqrt{2\pi t_2}}e^{\frac{-x_2^2}{2t_2}}}$
      $= \frac{1}{\sqrt{2\pi\frac{t_1(t_2-t_1)}{t_2}}}e^{\frac{-(x_1-\frac{t_1}{t_2}x_2)^2}{2\frac{t_1(t_2-t_1)}{t_2}}} \rightarrow X(t_1) - \frac{t_1}{t_2}X(t_2) \sim N(0, \frac{t_1(t_2-t_1)}{t_2})$
    - $\mathbb{E}[X(t_1)|X(t_2)] = \frac{t_1}{t_2}X(t_2)$
    - $Var(X(t_1)|X(t_2)) = \frac{t_1(t_2-t_1)}{t_2}$
    - $Y(t_1) = X(t_1) - \frac{t_1}{t_2}X(t_2)$ share the same distribution as $X(t_1)|X(t_2) = 0$

    $Cov(X(t_1), X(t_2)|X(t_3)) = \frac{t_1(t_3-t_2)}{t_3}$
    - $Cov(X(t_1), X(t_2)|X(t_3))$
      $= \mathbb{E}[X(t_1)X(t_2)|X(t_3)] - \mathbb{E}[X(t_1)|X(t_3)] \times \mathbb{E}[X(t_2)|X(t_3)]$
      $= \mathbb{E}[X(t_1)^2 + X(t_1)(X(t_2) - X(t_1))|X(t_3)] - \frac{t_1 t_2}{t_3^2}X(t_3)^2$
      $= \mathbb{E}[X(t_1)(X(t_2) - X(t_1))|X(t_3)] + \frac{t_1(t_1-t_2)}{t_3^2}X(t_3)^2 + \frac{t_1(t_3-t_1)}{t_3}$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi \frac{t_1(t_3-t_1)}{t_3}}} e^{\frac{-(x_1 - \frac{t_1}{t_3}X(3))^2}{2\frac{t_1(t_3-t_1)}{t_3}}} \mathbb{E}[x_1(X(t_2) - x_1)|X(t_3), X(t_1) = x_1]dx_1 + \frac{t_1(t_1-t_2)}{t_3^2}X(t_3)^2 + \frac{t_1(t_3-t_1)}{t_3}$$
$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi \frac{t_1(t_3-t_1)}{t_3}}} e^{\frac{-(x_1 - \frac{t_1}{t_3}X(3))^2}{2\frac{t_1(t_3-t_1)}{t_3}}} x_1(-x_1 + X(3))\frac{t_2-t_1}{t_3-t_1}dx_1 + \frac{t_1(t_1-t_2)}{t_3^2}X(t_3)^2 + \frac{t_1(t_3-t_1)}{t_3}$$
$$= \frac{t_1(t_2-t_1)}{t_3^2}X(t_3)^2 - \frac{t_1(t_2-t_1)}{t_3} + \frac{t_1(t_1-t_2)}{t_3^2}X(t_3)^2 + \frac{t_1(t_3-t_1)}{t_3}$$
$$= \frac{t_1(t_3-t_2)}{t_3}$$

- Brownian Motion with drift
  - Interpretation: generalize discrete time and space of biased random walk to be in continuous time and space
    * $X_i = \begin{cases} 1 & \text{with probability } p \\ -1 & \text{with probability } 1-p \end{cases}$
    * $S_t = \delta_x(\sum_{i=0}^{\frac{t}{\delta_t}} X_i)$
    * let $\delta_x = \sqrt{\delta_t}$, $p = \frac{1+\mu\sqrt{\delta_t}}{2}$, and $\delta_x \to 0$
    * $\mathbb{E}[S_t] = \mu t$
      · $\mathbb{E}[S_t] = \delta_x \frac{t}{\delta_t}(2p-1) = \mu t$
    * $Var(S_t) = t$
      · $Var(S_t) = \delta_x^2 \frac{t}{\delta_t}(1-(2p-1)^2) = t$
  - Definition:
    * $X(t)$ is Standard Brownian Motion
    * $Y(t) = X(t) + \mu t$
  - Property
    * Hitting Time
      The Gambler's Ruin
      · $\tau = \min\{t \geq 0 : Y(t) \in \{-A, B\}\}$
      · $P[Y(t) = A] = \frac{e^{-2\mu B}-1}{e^{-2\mu B}-e^{2\mu A}}, P[Y(t) = B] = \frac{1-e^{2\mu A}}{e^{-2\mu B}-e^{2\mu A}}$
        Prove by Martingales Stopping Theorem on $e^{cX(t)-\frac{c^2}{2}t}$ and $c = -2\mu$:
        $\to \mathbb{E}[e^{cX(t)-\frac{c^2}{2}t}] = \mathbb{E}[e^{-2\mu Y(t)}] = 1$
      · $\mathbb{E}[\tau] = \frac{1}{\mu}(P[Y(t) = B] \times (A+B) - A)$
        Prove by Martingales Stopping Theorem on $X(t)$:
        $\to \mathbb{E}[X(t)] = P[Y(t) = B]\mathbb{E}[B - \mu t|Y(t) = B] + P[Y(t) = A]\mathbb{E}[-A - \mu t|Y(t) = A] = 0$
      The monkey at the cliff
      · $\tau = \min\{t \geq 0 : X(t) = B\}$
      · $P[\tau < \infty] = \begin{cases} e^{2\mu B} & \text{if } \mu < 0 \\ 1 & \text{if } \mu \geq 0 \end{cases}$
        Prove by let $A = -\infty$ in The Gambler's Ruin
- Gaussian Process
  - Definition: A stochastic process $\{X(t) : t \geq 0\}$ such that for every $\{t_i\}_{i=1}^n$, $[X(t_1), X(t_2), \ldots, X(t_n)]$ is a joint Gaussian distribution
    * Defined by
      · $\mathbb{E}[X(t)], \forall t$
      · $Cov(X(s), X(t)), \forall s, t$
  - Property
    * Standard Brownian Motion is a Gaussian Process with $\mathbb{E}[X(t)] = 0$, $Cov(X(s), X(t)) = \min\{s, t\}$
      · $Cov(X(s), X(t)) = \min(s, t)$ (by $X(t) = X(s) + X(t-s)$ if $t > s$)
- Geometric Brownian Motion
  - Definition:

    - $Y(t) = e^{\sigma X(t)}$
  - Property:
    - $\mathbb{E}[Y(t)] = e^{\frac{\sigma^2 t}{2}}$
    - $Var[Y(t)] = e^{\sigma^2 t}$
- Brownian Motion reflected at the origin
  - Definition:
    - $Z(t) = |X(t)|$
  - Property
    - $P[Z(t) \geq x] = \frac{2}{\sqrt{2\pi t}} e^{\frac{-x^2}{2t}}$
      - same distribution as Maximum Brownian Motion
- Maximum Brownian Motion
  - Definition:
    - $Z(t) = \max_{0 \leq s \leq t} X(t)$
  - Property
    - $P[Z(t) \geq x] = \frac{2}{\sqrt{2\pi t}} e^{\frac{-x^2}{2t}}$
      - $P[Z(t) \geq x] = P[T_x \leq t] = \frac{2}{\sqrt{2\pi t}} e^{\frac{-x^2}{2t}}$
      - same distribution as Brownian Motion reflected at the origin
- Tricks
  - Creat $Y_1, Y_2 \sim N(0,1)$ and $Cov(Y_1, Y_2) = \cos\theta$
    - $X_1, X_2 \sim N(0,1)$ and independent
    - $Y_1 = X_1$
    - $Y_2 = \cos\theta \times X_1 + \sin\theta \times X_2$

14 Ornstein-Uhlenbeck Process