EECS 126 Crib Sheet - DIVIT RAWAL

$\Pr[A|B] = \frac{\Pr[A\cap B]}{\Pr[B]}$ $\quad \mathbb{1}^2 = \mathbb{1}$

$\Pr[B] = \sum_{i=1}^{n} \Pr[A_i]\Pr[B|A_i]$

$\Pr[A_i|B] = \frac{\Pr[A_i]\Pr[B|A_i]}{\Pr[B]}$

$\Pr[A|B] = \Pr[A] \Leftrightarrow A \perp\!\!\!\perp B$

$\{A_i\}_{i\in S}$ indep $\Leftrightarrow$ $\Pr[\bigcap_{i\in S} A_i] = \prod_{i\in S}\Pr[A_i]$

${}^nP_k = \frac{n!}{(n-k)!}$ $\quad \binom{n}{k} = \frac{n!}{k!(n-k)!}$

Partition of $n$ obj into $n_i$ groups:

$\binom{n}{n_1,\dots,n_k} = \frac{n!}{n_1!\dots n_k!}$

Stars & Bars: $n$ bins, $r$ balls $\binom{r+n-1}{n-1}$

$\Pr[X\in(a,b)] = \int_a^b f_X(x)\,dx \quad F_X(x) = \int_{-\infty}^{x} f(t)\,dt$

$\text{Exp}(\lambda): f_X(x) = \lambda e^{-\lambda x}, x>0; \quad F_X(x) = 1-e^{-\lambda x} \quad \mathbb{E}[X] = \frac{1}{\lambda} \quad \text{Var}[X] = \frac{1}{\lambda^2}$

$\Pr[X>t] = e^{-\lambda t}$

$N(\mu,\sigma^2): f_X(x) = (2\pi\sigma^2)^{-1/2}\exp(-(x-\mu)^2/(2\sigma^2))$

$aN(\mu_1,\sigma_1^2) + bN(\mu_2,\sigma_2^2) = N(a\mu_1+b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$.

Joint PDFs $\quad f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad X\perp\!\!\!\perp Y \Rightarrow f_{X,Y}(x,y) = f_X(x)f_Y(y)$

Convolution $\quad P_Z(z) = \Pr[X+Y=z] = \sum_x \Pr[X=x, Y=z-x]$

if $\perp\!\!\!\perp \rightsquigarrow = \sum_x \Pr_X[x]\Pr_Y[z-x]$

$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\,dx$

Order Stats $\quad Y = \min_{1\le k\le n} X_k \overset{iid}{\sim}$ CDF $F_X \quad F_Y(y) = 1-(1-F_X(y))^n$

$Y = \max_{1\le k\le n} X_k \overset{iid}{\sim}$ CDF $F_X \quad F_Y(y) = (F_X(y))^n$

Discrete RVs

| | $\Pr[X=k]$ | $\mathbb{E}[X]$ | $\text{Var}[X]$ |
|---|---|---|---|
| Bernoulli | $\begin{cases}1 & \text{w.p. } p\\ 0 & \text{else}\end{cases}$ | $p$ | $p(1-p)$ |
| Binomial | $\binom{n}{k}p^k(1-p)^{n-k}$ | $np$ | $np(1-p)$ |
| Geometric | $(1-p)^{k-1}p$ | $1/p$ | $\frac{1-p}{p^2}$ |
| Poisson | $\frac{e^{-\lambda}\lambda^k}{k!}$ | $\lambda$ | $\lambda$ |
| Uniform | $\begin{cases}((b-a)+1)^{-1} & k\in[a,b]\\ 0 & \text{else}\end{cases}$ | $\frac{b+a}{2}$ | $\frac{(b-a)}{12}$ |

$\text{Pois}(\lambda) + \text{Pois}(\mu) = \text{Pois}(\lambda+\mu)$

Moment Generating Functions

$M_X(s) = \mathbb{E}[\exp(sX)] = \int_{-\infty}^{\infty} e^{sx} f_X(x)\,dx$

$\left.\frac{d^n M_X(s)}{ds^n}\right|_{s=0} = \int x^n f(x)\,dx = \mathbb{E}[X^n]$

Pois: $M(s) = e^{\lambda(e^s-1)}$ $\quad$ Exp: $M(s) = \frac{\lambda}{\lambda-s} \quad s<\lambda$

Gaussian: $M(s) = \exp(s^2/2) \cdot \exp(\mu t + \frac{\sigma^2 t^2}{2})$

Geom: $M(s) = (p\exp(s))/(1-(1-p)e^s)$

Bernoulli: $M(s) = 1-p+pe^s$

Binom: $M(s) = (1-p+pe^s)^n$ $\quad Z = \sum X_i \Rightarrow M_Z(s) = \prod_i M_{X_i}(s)$

Unif.: $M(s) = \begin{cases}\frac{e^{bs}-e^{as}}{s(b-a)} & s\neq 0\\ 1 & s=0\end{cases}$

Expectation, Variance, Covariance

$\mathbb{E}[aX+bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$

$\mathbb{E}[X] = \sum_i x_i \Pr[x_i]$ for $x_i$ values $X$ can take

$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]$

$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ if $X\perp\!\!\!\perp Y$.

$\text{Var}[X] = \mathbb{E}[(X-\mathbb{E}[X])^2]$
$= \mathbb{E}[X^2] - \mathbb{E}[X]^2$
$= \text{Cov}[X,X]$

$\text{Var}[X] = \text{Var}[\mathbb{E}[X|Y]] + \mathbb{E}[\text{Var}[X|Y]]$

$\text{Cov}[X,Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$

$X\perp\!\!\!\perp Y \Rightarrow \text{Var}[XY] = \mathbb{E}[X^2]\mathbb{E}[Y^2] - \mathbb{E}[X]^2\mathbb{E}[Y]^2$

$\mathbb{E}[X] = \sum_y \Pr_Y[y]\,\mathbb{E}[X|Y=y]$

$Y = g(X)$

$\Rightarrow f_Y(y) = f_X(g^{-1}(y))\left|\frac{d}{dy}g^{-1}(y)\right|$

Bayes Rule: $\Pr[A|B] = \frac{\Pr[B|A]\Pr[A]}{\sum_j \Pr[B|A_j]\Pr[A_j]}$.

Inclusion-Exclusion:

$|\bigcup_{i=1}^{n} A_i| = \sum_i |A_i| - \sum_{i<j}|A_i\cap A_j| + \sum_{i<j<k}|A_i\cap A_j\cap A_k| \dots + (-1)^{n-1}\sum_{i_1<\dots<i_n}|A_{i_1}\cap\dots\cap A_{i_n}|$

General $\quad X_i \overset{iid}{\sim} \text{Exp}(\lambda_i) \Rightarrow \min_i X_i \sim \text{Exp}(\sum_i \lambda_i) \quad X_i \overset{iid}{\sim} \text{Geom}(p)$

$\Pr[X_k = \min_i X_i] = \frac{\lambda_k}{\sum \lambda_i} \quad \min X_i \sim \text{Geom}(1-(1-p)^n)$

If $X = \sum_i \mathbb{1}_i$, $\mathbb{E}[X(X-1)] = \sum_{i\neq j}\Pr[\mathbb{1}_i, \mathbb{1}_j]$.

Exp: $\Pr[X>s+t \mid X>s] = \Pr[X>t]$

Geom: $\Pr[X>s+t \mid X>s] = \Pr[X>t]$

$S_n \sim \text{Binom}(n, p_n)$ w/ $p_n n \to \lambda$

$\Rightarrow S_n \to \text{Pois}(\lambda)$

$\mathbb{E}[g(X,Y)] = \sum_x\sum_y g(x,y)\Pr_{X,Y}[x,y]$

$\text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X,Y] \quad \mathbb{E}[\mathbb{E}[X|Y]g(Y)] = \mathbb{E}[Xg(Y)]$

## Concentration Inequalities

Markov ($X \geq 0$, finite $\mathbb{E}$): $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$

Chebyshev (finite $\mathbb{E}$, Var): $\Pr[|X-\mu| \geq c] \leq \frac{\sigma^2}{c^2}$

Chernoff (sum of indep. RVs):

$\Pr[X \geq a] \leq \frac{\mathbb{E}[e^{sX}]}{e^{sa}}$ $s>0$

$\Pr[X \leq a] \leq \frac{M(s)}{e^{sa}}$ $s<0$

CLT: $Z \doteq \frac{S_n - n\mu}{\sigma\sqrt{n}}$, $F_Z(z) \to \Phi(z)$

Union Bound: $\Pr[\bigcup_{i=1}^{n} A_i] \leq \sum_{i=1}^{n} \Pr[A_i]$

WLLN: $\lim_{n\to\infty} \Pr[|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}[X]| \geq \varepsilon] = 0$

## Convergence

a.s. : $\Pr[\lim_{n\to\infty} X_n = X] = 1$

$\Downarrow$

i.p. : $\lim_{n\to\infty} \Pr[|X_n - X| \geq \varepsilon] = 0$

$\Downarrow$

i.d. : $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)\ \forall x$

Borell-Cantelli Lem.

$\sum_{n=1}^{\infty} \Pr[E_n] < \infty \Rightarrow \Pr[\limsup_{n\to\infty} E_n] = 0$

$\limsup_{n\to\infty} E_n \doteq \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} E_k$

$\sum_{n=1}^{\infty} \Pr[A_n] < \infty \Rightarrow \Pr[A_n \text{ i.o.}] = 0$

SLLN: i.i.d. $X$s, $M_n \doteq \frac{1}{n}\sum_{i=1}^{n} X_i$

$\Pr[\lim_{n\to\infty} M_n = \mu] = 1$

## Poisson Processes

$\Pr[N_t = n] \sim \mathrm{Pois}(\lambda t)$

$\forall t,s>0$ $N(t,t+s) \overset{d}{=} N(s)$ For $0<t_1<\dots<t_k$;

$N(t) \sim \mathrm{Pois}(\lambda t)$ $N(t_1), N(t_1,t_2), \dots$ are jointly indep.

Fix $\lambda > 0$ & sample interarrival times $S_1, S_2, \dots \sim \mathrm{Exp}(\lambda)$.

$\forall n \geq 1$ $T_n \doteq \sum_{j=1}^{n} S_j$. Then $N(t) = \max\{n \geq 0 \mid T_n \leq t\}$ = arrivals in time $t$

$T_n = \sum_{k=1}^{n} S_k \sim \mathrm{Gamma}(n,\lambda)$ $f_{T_n}(s) = \frac{\lambda e^{-\lambda s}(\lambda s)^{n-1}}{(n-1)!}$

Memorylessness: $N_{T_i} - N_{T_{i-1}} \sim \mathrm{Pois}(\lambda(t_i - t_{i-1}))$ $\Pr[N(t) \text{ even}] = \frac{1+e^{-2\lambda t}}{2}$

Merging: $PP(\lambda_1) + PP(\lambda_2) = PP(\lambda_1+\lambda_2)$ Splitting: $\Pr[T_a = \min(T_a, T_b)] = \frac{\lambda_a}{\lambda_a+\lambda_b}$

## Descrete Time Markov Chain

Markov Property: $\Pr[X_{n+1} | X_n, \dots, X_1] = \Pr[X_{n+1} | X_n]$

Stationarity: $\pi = \pi P$ Chapman-Kolmogorov: $P^n_{ij} = [P^n]_{ij}$

Periodicity: $d(i) = \gcd_{n\geq 1} P^n_{ii} > 0$ Detailed Balance: $\pi_i P_{ij} = \pi_j P_{ji}$

Hitting Time: $\beta(i) = \begin{cases} 1 + \sum_j P_{ij}\beta(j) & i \notin A \\ 0 & i \in A \end{cases}$ $\underset{\text{transient}}{T} = S \setminus \underset{\text{absorbing}}{A}$

Absorbtion Prob. $b_i = \sum_{j\in T} P_{ij} b_j + \sum_{j \in A} P_{ij}$ $b_j = 1\ \forall j \in A$

$\exists \pi \Rightarrow$ every state positive recurrent transient: return pr < 1

null recurrent: $\mathbb{E}$[steps to return] $= \infty$, else positive recurrent

irreducible: every state reachable from every other in finite steps

finite + irreducible $\Rightarrow \exists \pi$ $P_{ij}$, row $i$, col $j$ = $\Pr[i \to j]$.

non-closed communicating class: inside $\to$ outside possible, else closed

Prob A before B: $\alpha(A) = 1$, $\alpha(B) = 0$, $\alpha(k) = \sum_{m \in \text{neighbors}} P_{km}\alpha(m)$

## Continuous Time Markov Chain

Stationary Dist of Jump Chain: $\mu(i) = \frac{q(i)\pi(i)}{\sum_{j=1} q(j)\pi(j)}$

Generator matrix $Q$. $q_{ij} \geq 0\ \forall i \neq j$

solve $\pi Q = 0$, $\sum_i \pi(i) = 1$

$q_{ii} = -\sum_{i\neq j} q_{ij} \Rightarrow$ row sums $= 0$

Holding Time $T_i \sim \mathrm{Exp}(q_i)$, $q_i = \sum_{i\neq j} q_{ij}$

$P(t) = e^{Qt}$ $e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$

Hitting Times: For $i \neq a$ $-1 = \sum_j q_{ij} h_j$, $h_a = 0$

Uniformization: $\Lambda > \max_i q_i$ $P = I + \frac{1}{\Lambda} Q$. Then $P(t) = \sum_{n=0}^{\infty} e^{-\Lambda t} \frac{(\Lambda t)^n P^n}{n!}$

$\Pr[\text{next} = j | \text{leaving } i] = \frac{q_{ij}}{q_i}$ $\pi P_{\text{unif}} = \pi$

Birth-Death: birth $\lambda_n$ death $\mu_n$ $\pi_n = \pi_0 \prod_{k=1}^{n} \frac{\lambda_{k-1}}{\mu_k}$ $\pi_0 = \left(1 + \sum_{n=1}^{\infty}\prod_{k=1}^{n}\frac{\lambda_{k-1}}{\mu_k}\right)^{-1}$

$\exists \pi$ iff $\sum_{n=0}^{\infty}\prod_{k=1}^{n} \frac{\lambda_{k-1}}{\mu_k} < \infty$

## Estimation

MLE: $\hat{\theta}_{MLE} = \arg\max_\theta L(\theta) = \arg\max_\theta f(X|\theta) = \arg\max_\theta \ell(\theta)$

$\ell(\theta) = \sum_i \log f(x_i | \theta)$

MAP (minimizes P[err]): $\hat{\theta}_{MAP} = \arg\max_\theta f(x|\theta)\pi(\theta) = \arg\max_\theta (\ell(\theta) + \log(\pi(\theta)))$ ← posterior mode

MMSE: posterior mean $\hat{\theta} = \mathbb{E}[\theta | x]$; posterior median optimal for abs. loss

LLSE: $\hat{Y} = aX + b$ $\hat{Y} = \mathbb{E}[Y] + \frac{\mathrm{Cov}[X,Y]}{\mathrm{Var}[X]}(X - \mathbb{E}[X])$ $a = \mathbb{E}[Y] - b\mathbb{E}[X]$

$b = \mathrm{Cov}[X,Y]/\mathrm{Var}[X]$

jointly gaussian $\Rightarrow$ MMSE = LLSE $\hat{\beta} = (X^T X)^{-1} X^T Y$

$\mathbb{E}[(Y-\hat{Y})X] = 0$, $\mathbb{E}[g(Y)(\hat{X} - X)] = 0$

Bayes Rule: $f(\theta|x) = \frac{f(x|\theta)f(\theta)}{f(x)}$

Viterbi Alg $V_t(j)$ = max prob of ending in $j$ at $t$.

$V_0(j) = \pi_j \cdot Q_j(y_0)$

$V_t(j) = Q_j(y_t) \cdot \max_i [V_{t-1}(i) \cdot P_{ij}]$

$B_t(j) = \arg\max_i [V_{t-1}(i) P_{ij}]$

$\hat{x}_n = \arg\max_j V_n(j)$

$\hat{x}_{t-1} = B_t(\hat{x}_t)$

## Hypothesis Testing

Errors:

Type I ($\alpha$): reject $H_0$ but is True

Type II ($\beta$): fail to reject $H_0$ but is False

Power: $1-\beta$ ← want to maximize

$\Lambda^*(x) = \frac{f(x|H_0)}{f(x|H_1)}$ ← LRT

Reject $H_0$ if $\Lambda(x) \gtrless \eta$

Neyman-Pearson Lem: UMP test of level $\alpha$ is LRT.

$$\phi(y) = \begin{cases} 1 & \Lambda > \eta \\ \rho & \Lambda = \eta \\ 0 & \Lambda < \eta \end{cases}$$

$Pr[\text{false alarm}] = Pr[\hat{H} = H_1 | H_0] = \sum_i \alpha_i Pr_0[Y = y_i] = \beta$  $\alpha_i \in [0,1]$

$\phi(y) = \mathbb{1}\{\Lambda(y) > \eta\} + \rho\, \mathbb{1}\{\Lambda(y) = \eta\}$ s.t. $\mathbb{E}_{H_0}[\phi(y)] = \beta$

Set greatest LR to 1, then up others if FA "budget"

## Metropolis-Hastings

Target $\pi(x)$, Proposal dist $q(x'|x)$ to sample from

Alg: Given current state $x$, propose $x' \sim q(x'|x)$

$$\alpha(x,x') = \min\left(1, \frac{\pi(x')\, q(x|x')}{\pi(x)\, q(x'|x)}\right)$$

w.p. $\alpha$ set $x_{t+1} = x'$, else $x_{t+1} = x$.

## Joint Gaussians

Def. Let $X = (X_1, \dots X_n)^T$. Let $Z \in \mathbb{R}^l$ s.t. $Z_i \sim N(0,1)$ iid.

$X_1, \dots, X_n$ are JG if $\exists \mu \in \mathbb{R}^n$ s.t. $X = AZ + \mu$  $A \in \mathbb{R}^{n \times l}$  $\Sigma = AA^T$

Def. $X_1, \dots, X_n$ are JG if any ~~lin comb of~~ $u^T X$ follows a normal distr.

For $\Sigma \succ 0$, $f_X(x) = ((2\pi)^n \det(\Sigma))^{-1/2} \exp(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu))$

Correlation $\rho = \frac{Cov[X,Y]}{\sigma_X \sigma_Y}$

Thm JG are indep. ⟺ uncorrelated

Thm lin combs of JG RVs are JG

Thm MMSE $\mathbb{E}[X|Y]$ = LLSE $\mathbb{L}[X|Y]$

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix}\right)$$

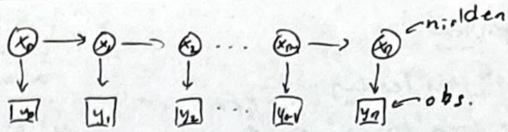
$\mathbb{E}[X|Y=y] = \mu_X + \Sigma_{XY} \Sigma_{YY}^{-1} (y - \mu_Y)$

if $X \sim N(\mu, \Sigma)$, $AX + b \sim N(A\mu + b, A\Sigma A^T)$

$Cov[X|Y] = \Sigma_{XX} - \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX}$

$\hat{X} = \mathbb{E}[X|Y]$

## Hidden Markov Models




MLSE: $MAP[X^n | Y^n = y^n]$

$$x^{n*} = \arg\max_{x \in \mathcal{X}^n} Pr[X^n = x^n | Y^n = y^n]$$

$$= \arg\max_{x^n \in \mathcal{X}^n} \Big[\log \underbrace{\pi_0(x_0)}_{\text{initial state}} \underbrace{Q(x_0, y_0)}_{\text{hidden} \to \text{obs}} + \sum_{m=1}^{n} \log[\underbrace{P(x_{m-1}, x_m)}_{\text{hidden transitions}} Q(x_m, y_m)]\Big]$$

Define $d_0(x_0) = -\log \pi_0(x_0) Q(x_0, y_0)$

$d_m(x_{m-1}, x_m) = -\log[P(x_{m-1}, x_m) Q(x_m, y_m)]$

$$\Rightarrow x^{n*} = \arg\min_{x^n} \Big[d_0(x_0) + \sum_{m=1}^{n} d_m(x_{m-1}, x_m)\Big]$$

## Kalman Filters

State: $X_n = AX_{n-1} + V_n$  $V_n \sim N(0, \Sigma_V)$  LLSE = MMSE

Obs: $Y_n = CX_n + W_n$  $W_n \sim N(0, \Sigma_W)$

At every time $n$, maintain:

estimate: $\hat{X}_{n|n} = \mathbb{E}[X_n | Y_{1:n}]$

err cov: $\Sigma_{n|n} = Cov(X_n - \hat{X}_{n|n})$

use:

Prediction: $\hat{X}_{n|n-1}$

Pred cov: $\Sigma_{n|n-1}$

Prediction step: $\hat{X}_{n|n-1} = A\hat{X}_{n-1|n-1}$

$\Sigma_{n|n-1} = A\Sigma_{n-1|n-1}A^T + \Sigma_V$

Innovation: $\tilde{Y}_n = Y_n - C\hat{X}_{n|n-1}$

Kalman Gain: $K_n = \Sigma_{n|n-1} C^T (C\Sigma_{n|n-1} C^T + \Sigma_W)^{-1}$

If $C = 1$ (scalar case): $K_n = \frac{\sigma^2_{n|n-1}}{\sigma^2_{n|n-1} + \sigma^2_W}$

Update Step:

$\hat{X}_{n|n} = \hat{X}_{n|n-1} + K_n \tilde{Y}_n = (I - K_n C)\hat{X}_{n|n-1} + K_n Y_n$

$\Sigma_{n|n} = (I - K_n C)\Sigma_{n|n-1}$

### Scalar Case

$\sigma^2_{n|n-1} = A^2 \sigma^2_{n-1|n-1} + \sigma_V^2$  $K_n = \frac{\sigma^2_{n|n-1}}{\sigma^2_{n|n-1} + \sigma_W^2}$

↑ pred. var.

$\hat{X}_{n|n} = (1 - K_n)\hat{X}_{n|n-1} + K_n Y_n$  $\sigma^2_{n|n} = (1 - K_n)\sigma^2_{n|n-1}$

## General

- For affine $f(Y)$, $\mathbb{E}[X|Y] = \mathbb{L}[X|Y]$
- $Z - \mathbb{L}[Z|Y] \perp Y$
- $X \perp\!\!\!\perp Y \Rightarrow Cov(AX + Y) = A\,Cov(X)A^T + Cov(Y)$
- $S_n = \frac{a(1 - r^n)}{1 - r}$  $a$ - first term, $r$ - common ratio  $S = \frac{a}{1-r}$ for $|r| < 1$
- $Var[Y|X] = \mathbb{E}[Y^2|X] - (\mathbb{E}[Y|X])^2$

$m$th order stat of $n$ unif $U[0,1)$ is $\frac{m}{n+1}$

# Examples

## Hypothesis Testing

$\alpha = PFA = .05$

$X \sim Geom(p)$, $H_0: p = .5$, $H_1: p = .25$

$\Lambda(x) = \frac{f_1(x)}{f_0(x)} = \frac{1}{2}\left(\frac{3}{2}\right)^{x-1}$ MLR.

N-P needs $k$ s.t. $Pr_{H_0}[\Lambda(x) > k] \le .05$

inc $\Rightarrow \Lambda(x) > k \iff \frac{1}{2}\left(\frac{3}{2}\right)^{x-1} > k$

$\iff x > \frac{\ln(2k)}{\ln(3/2)} + 1 \doteq t(k)$

$Pr_{H_0}[X > m] = \left(\frac{1}{2}\right)^m$

$\Rightarrow Pr_{H_0}[X > \lceil t(k) \rceil] = (.5)^{\lceil t(k) \rceil}$

$\lceil t(k) \rceil = c$ w/ $(.5)^c = .05$

$\Rightarrow t(k) = \frac{\ln(.05)}{\ln(.5)} \approx 4.32$

$\Rightarrow c = 5$

$Pr_{H_0}[X > 5] = (.05)^5 = .03125 < .05$

$\Rightarrow$ need randomization

$Pr[X = 5] = .5(.5)^4 = .03125$

Then solve

$.03125 + p(.03125) = .05$

$\Rightarrow p = .6$

$\Rightarrow \phi^*(x) = \begin{cases} \text{Reject } H_0 \text{ if } x > 5 \\ \text{Reject } H_0 \text{ w.p. } p = .6 \text{ if } x = 5 \\ \text{Accept } H_0 \text{ if } x < 5 \end{cases}$

## JG Sin

$X, Y, Z$ JG w/ cov $\begin{bmatrix} 4 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 4 \end{bmatrix}$ $\mu = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$.

$\mathbb{E}[\sin(X) Y \sin(Z)] = \mathbb{E}[\mathbb{E}[\sin(X) Y \sin(Z) \mid X, Z]]$

$= \mathbb{E}[\sin(X)\sin(Z)\,\mathbb{E}[Y \mid X, Z]]$

$= \mu_Y + \Sigma_{Y,(X,Z)} \Sigma_{(X,Z)}^{-1} \begin{bmatrix} X \\ Z \end{bmatrix} = 2 + [1\ 1]\begin{bmatrix} 1/4 & 0 \\ 0 & 1/4 \end{bmatrix}\begin{bmatrix} X \\ Z \end{bmatrix}$

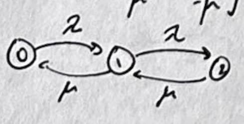
$= 2 + \frac{1}{4}X + \frac{1}{4}Z$ $\quad \mathbb{E}[\sin(X)] = 0 = \mathbb{E}[\sin(Z)]$

$\Rightarrow$ full $= 0$

## ~~Even Times Kalman Filter~~ CTMC $\pi$

2 servers, time until break $\sim Exp(\mu)$, repair time $\sim Exp(\lambda)$

$Q = \begin{bmatrix} -\lambda & \lambda & 0 \\ \mu & -(\lambda+\mu) & \lambda \\ 0 & \mu & -\mu \end{bmatrix}$ CTMC in # operational servers




$\lambda\pi(0) = \mu\pi(1)$

$(\lambda+\mu)\pi(1) = \lambda\pi(0) + \mu\pi(2)$

$\mu\pi(2) = \lambda\pi(1)$

$1 = \pi(0) + \pi(1) + \pi(2)$

$\pi(2) = \frac{\lambda}{\mu}\pi(1) \Rightarrow \mu\pi(1) = \lambda\pi(0)$

$1 = \pi(0) + \left(1 + \frac{\lambda}{\mu}\right)\pi(1)$ $\quad \pi(1) = \frac{\lambda}{\mu}\pi(0)$

$\Rightarrow \pi(0) = \frac{1}{1 + \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu}\right)^2}$ $\quad \pi(1) = \frac{\lambda}{\mu}\left(\frac{1}{1 + \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu}\right)^2}\right)$

$\pi(2) = \left(\frac{\lambda}{\mu}\right)^2\left(\frac{1}{1 + \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu}\right)^2}\right)$

## DTMC A before B

Fair coin, $\mathbb{E}$ flips before HT seq.

$S = \{0, 1, 2\}$ for $S = 0 \Rightarrow$ last was tails or start, $S = 1 \Rightarrow$ last was heads, $S = 2$ is absorbing HT

Prs: $0 \to 0 = .5$, $0 \to 1 = .5$, $1 \to 2 = .5$

$1 \to 1 = .5$, $2 \to 2 = 1$

$\beta(i)$ denotes # of steps to 2 from $i$

$\beta(2) = 0$, first-step eqs give

$\beta(0) = 1 + .5\beta(0) + .5\beta(1)$

$\beta(1) = 1 + .5\beta(1) + .5\beta(2)$

$\Rightarrow \beta(1) = 2 \Rightarrow \beta(0) = 4$

## MMSE

$X \perp\!\!\!\perp Z$, $X, Z$ iid $S \doteq X + Z$

$\mathbb{E}[X \mid S = s] = \mathbb{E}[Z \mid S = s]$

$\mathbb{E}[S \mid S = s] = \mathbb{E}[X + Z \mid S = s] = 2\mathbb{E}[X \mid S = s]$

$\mathbb{E}[S \mid S = s] = s \Rightarrow \mathbb{E}[X \mid S = s] = \frac{s}{2}$

## Poisson Process

$n$ bulbs, toggled $\sim Pois(\lambda)$ $N_i(t) \doteq$ # times $i$ toggled up to $t$

Find $\mathbb{E}[N_1(t) \mid N_1(t) + N_2(t) + N_3(t) = k]$

Let $a_i = \mathbb{E}[N_i(t) \mid \sum_{i=1}^{3} N_i(t) = k]$. By symmetry $a_1 = a_2 = a_3$ & $a_1 + a_2 + a_3 = k \Rightarrow a_1 = \frac{k}{3}$

## Poisson Process Arrival Times

PP $(N_t)_{t \ge 0}$ w/ rate 1. $T_k$ is time of $k$th arrival

(a) $\mathbb{E}[T_3 \mid N_1 = 2] = 1 + \mathbb{E}[T_1] = 2$ ← memoryless

(b) Given $T_3 = s > 0$, find joint of $T_1, T_2$

$f_{T_k} = \frac{s^{k-1}}{(k-1)!} e^{-s} \mathbb{1}_{s \ge 0}$ $\quad$ Via Bayes & memorylessness

$f_{T_1, T_2 \mid T_3}(s_1, s_2 \mid s) = \frac{f_{T_1, T_2, T_3}(s_1, s_2, s)}{f_{T_3}(s)}$

$= \frac{e^{-s_1} e^{-(s_2 - s_1)} e^{-(s - s_2)}}{s^2 e^{-s}/2!} \mathbb{1}_{\{0 \le s_1 \le s_2 \le s\}}$

$= \frac{2}{s^2} \mathbb{1}_{\{0 \le s_1 \le s_2 \le s\}}$

(c) ~~Find~~ $\mathbb{E}[T_2 \mid T_3 = s]$ is max of 2 $U[0, s]$ RVs. Then for $0 \le x \le s$

$F_{T_2 \mid T_3}(x \mid s) = Pr[T_2 \le x \mid T_3 = s] = \left(\frac{x}{s}\right)^2$

$f_{T_2 \mid T_3}(x \mid s) = \frac{2x}{s^2}$

$\mathbb{E}[T_2 \mid T_3 = s] = \int_0^s \frac{2x^2}{s^2} dx = \frac{2s}{3}$

## Another Poisson Process

$(N_t)_{t \ge 0}$ is $PP(\lambda)$. $T_k$ s.t. $k \ge 1$ is $k$th arrival time

Given $0 \le s < t$ $N(s, t) \doteq N(t) - N(s)$

(a) $Pr[N(1) + N(2, 4) + N(3, 5) = 0] = Pr[N(1) = 0]\,Pr[N(2, 5) = 0]$

$= e^{-\lambda} e^{-3\lambda} = e^{-4\lambda}$

(b) $\mathbb{E}[N(1, 3) \mid N(1, 2) = 3] = \mathbb{E}[N(2, 3)] + N(1, 2) = 3 + \lambda$

(c) $\mathbb{E}[T_2 \mid N(2) = 1] \ne \mathbb{E}[T_2 - 2 \mid N(2) = 1] = \frac{1}{\lambda}$ by memoryless & so answer is $2 + \lambda^{-1}$

## Even Times Kalman Filter

Random Process $(X_n)_{n \in \mathbb{N}}$ w/ $\quad V_n \perp\!\!\!\perp W_n$

$X_{n+1} = aX_n + V_n$ $\quad V_n \overset{iid}{\sim} N(0, \sigma_V^2)$

$Y_n = X_n + W_n$ $\quad W_n \overset{iid}{\sim} N(0, \sigma_W^2)$

We only observe $Y_0, Y_2, Y_4, \ldots$

(a) Recurrence reln for $\hat{X}_{2n \mid 2n} \doteq \mathbb{L}(X_{2n} \mid Y_0, Y_2, \ldots, Y_{2n})$ in terms of $\hat{X}_{2n-2 \mid 2n-2}$.

$X_{2n+2} = a^2 X_{2n} + (aV_{2n} + V_{2n+1})$

$\Rightarrow$ updates are $\quad aV_{2n} + V_{2n+1} \sim N(0, (a^2+1)\sigma_V^2)$

$\hat{X}_{2n+2 \mid 2n+2} = a^2 \hat{X}_{2n \mid 2n} + K_{2n+2} \tilde{Y}_{2n+2}$

$\tilde{Y}_{2n+2} = Y_{2n+2} - a^2 \hat{X}_{2n \mid 2n}$ for

$K_{2n+2} = \frac{\sigma^2_{2n+2 \mid 2n}}{\sigma^2_{2n+2 \mid 2n} + \sigma_W^2}$ $\quad \sigma^2_{2n+2 \mid 2n} = (a^2 \sigma^2_{2n \mid 2n})^2 + (a^2+1)\sigma_V^2$

$\sigma^2_{2n+2 \mid 2n+2} = (1 - K_{2n+2})\sigma^2_{2n+2 \mid 2n}$

(b) $\hat{X}_{2n+1 \mid 2n} = a\hat{X}_{2n \mid 2n}$ by linearity of LLSE

## Hidden Markov Models

(a) $Pr[X_0 = x_0, Y_0 = y_0, \ldots, X_n = x_n, Y_n = y_n]$

$= \pi_0(x_0) Q(x_0, y_0) \prod_{i=1}^{n} P(x_{i-1}, x_i) Q(x_i, y_i)$

(b) $Pr[X_0 = x_0 \mid Y_0 = y_0]$ by Bayes Rule

$= \frac{Pr[X_0 = x_0, Y_0 = y_0]}{Pr[Y_0 = y_0]} = \frac{\pi_0(x_0) Q(x_0, y_0)}{\sum_{x \in \mathcal{X}} \pi_0(x) Q(x, y_0)}$

(c) Observe $(y_0, \ldots, y_n)$, find most likely $(x_0, \ldots x_n)$.

$U(x_m, m) = \max_{x_{m+1}, \ldots, x_n \in \mathcal{X}} Pr[X_m = x_m, X_{m+1:n} = x_{m+1:n}, Y_{0:n} = y_{0:n}]$

is largest prob for seq. of states starting @ $x_m$ w/ $(y_0, \ldots, y_n)$

$U(x_m, m) = \max_{x_{m+1} \in \mathcal{X}} P(x_m, x_{m+1}) Q(x_{m+1}, y_{m+1}) \cdot U(x_{m+1}, m+1)$

## Joint Gaussian MMSE

$\begin{bmatrix} X \\ Y \end{bmatrix} \sim N\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\right)$ $\quad W \doteq \begin{cases} 1 & Y > 0 \\ 0 & Y = 0 \\ -1 & Y < 0 \end{cases}$

(a) $\mathbb{E}[WX \mid Y] = W\mathbb{E}[X \mid Y] = W\mathbb{L}[X \mid Y]$

$\Rightarrow \mathbb{E}[X \mid Y] = \mathbb{L}[X \mid Y] = 1 + \frac{Y}{2}$

$\Rightarrow \mathbb{E}[WX \mid Y] = \begin{cases} 1 + \frac{Y}{2} & Y > 0 \\ 0 & Y = 0 \\ -1 - \frac{Y}{2} & Y < 0 \end{cases}$

(b) LLSE $\ne$ MMSE (varies w/ sign)

(c) LLSE $\ne$ MMSE $\Rightarrow$ not JG

# Kalman Filter

Standard but obs have const. & unknown bias $\Omega$ but no other noise. $\forall i \geq 0$

$$X_{i+1} = a X_i \qquad X_0, \Omega \text{ are 0 mean}$$
$$Y_i = X_i + \Omega. \qquad \text{\& indep. w/ var } \sigma_x^2, \sigma_\Omega^2$$

We want to do Kalman prediction

$$\hat{x}_{i|i-1} = \mathbb{L}[X_i | Y_0 \dots Y_{i-1}] \;\&\; \sigma^2_{i|i-1} = \mathbb{E}[(X_i - x_{i|i-1})^2]$$

Define $\hat{\Omega}_i = \mathbb{L}[\Omega | Y_0 \dots Y_i]$

$$\lambda_i^2 = \mathbb{E}[(\Omega - \hat{\Omega}_i)^2]$$
$$\rho_i = \mathbb{E}[(X_{i+1} - x_{i+1|i})(\Omega - \hat{\Omega}_i)]$$

(i) $\hat{x}_{0|-1} = \mathbb{E}[X_0]$, $\hat{\Omega}_{-1} = E[\Omega]$, give $\sigma^2_{0|-1}, \lambda^2_{-1}, \rho_{-1}$

$$\sigma^2_{0|-1} = \sigma_x^2,\ \lambda^2_{-1} = \sigma_\Omega^2,\ \rho_{-1} = \mathrm{Cov}[X_0, \Omega] = 0$$

(ii) Find Kalman update for $\hat{x}_{i+1|i}$. $\hat{x}_{i+1|i} = a\hat{x}_{i|i-1} + K_i \hat{Y}_i$

$$\hat{x}_{i+1|i} = a\hat{x}_{i|i-1} + \frac{\mathrm{Cov}[X_{i+1}, \hat{Y}_i]}{\mathrm{Var}[\hat{Y}_i]}\hat{Y}_i.$$

Note $\hat{Y}_i = Y_i - \mathbb{L}[X_i + \Omega | Y_0 \dots Y_{i-1}] = Y_i - \hat{x}_{i|i-1} - \hat{\Omega}_{i-1}$

$$\mathrm{Cov}[X_{i+1}, \hat{Y}_i] = a\sigma^2_{i|i-1} + a\rho_{i-1} \text{ b/c}$$

$$\mathrm{Cov}[X_i, X_i - \hat{x}_{i|i-1} + \Omega - \hat{\Omega}_{i-1}] = \mathrm{Cov}[X_i, X_i - \hat{x}_{i|i-1}] + \mathrm{Cov}[X_i, \Omega - \hat{\Omega}_{i-1}]$$
$$= \mathrm{Cov}[X_i - \hat{x}_{i|i-1}, X_i - \hat{x}_{i|i-1}] + \mathrm{Cov}[X_i - \hat{x}_{i|i-1}, \Omega - \hat{\Omega}_{i-1}]$$

by orthogonality

Similarly, $\mathrm{Var}[\hat{Y}_i] = \mathrm{Var}[X_i + \Omega - \hat{x}_{i|i-1} - \hat{\Omega}_{i-1}]$

$$= \sigma^2_{i|i-1} + \lambda^2_{i-1} + 2\rho_{i-1}$$

$$\Rightarrow K = \frac{a\sigma^2_{i|i-1} + a\rho_{i-1}}{\sigma^2_{i|i-1} + \lambda^2_{i-1} + 2\rho_{i-1}}$$

(iii) Find $L_i$ in $\hat{\Omega}_i = \hat{\Omega}_{i-1} + L_i \hat{Y}_i$

$$\hat{\Omega}_i = \hat{\Omega}_{i-1} + \frac{\mathrm{Cov}[\Omega, \hat{Y}_i]}{\mathrm{Var}[\hat{Y}_i]}\hat{Y}_i$$

$$\mathrm{Cov}[\Omega, \hat{Y}_i] = \mathrm{Cov}[\Omega, Y_i - \hat{x}_{i|i-1} - \hat{\Omega}_{i-1}]$$
$$= \mathrm{Cov}[\Omega, X_i + \Omega - \hat{x}_{i|i-1} - \hat{\Omega}_{i-1}]$$
$$= \lambda^2_{i-1} + \rho_{i-1}$$

$$\Rightarrow L_i = \frac{\lambda^2_{i-1} + \rho_{i-1}}{\sigma^2_{i|i-1} + \lambda^2_{i-1} + 2\rho_{i-1}}$$

(iv) Kalman Update for $\sigma^2_{i+1|i}$ ($\sigma^2_{i+1|i} = a^2\sigma^2_{i|i-1} - K_i \alpha_i$)

$$\sigma^2_{i+1|i} = \sigma^2_{i+1|i-1} - K_i \mathrm{Cov}[X_{i+1}, \hat{Y}_i] \Rightarrow \sigma^2_{i+1|i} = \sigma^2_{i+1|i-1} - aK_i(\sigma^2_{i|i-1} + \rho_{i-1})$$

find $\alpha_i$

and $\sigma^2_{i+1|i-1} = \mathrm{Var}[X_{i+1} - \hat{x}_{i+1|i-1}] = \mathrm{Var}[aX_i - a\hat{x}_{i|i-1}] = a^2\sigma^2_{i|i-1}$

$$\alpha_i = a(\sigma^2_{i|i-1} + \rho_{i-1}).$$

(v) Find $\beta_i$ in

$$\lambda_i^2 = \lambda^2_{i-1} - L_i\beta_i$$
$$\lambda_i^2 = \lambda^2_{i-1} - L_i \mathrm{Cov}[\Omega, \hat{Y}_i]$$
$$\Rightarrow \lambda_i^2 = \lambda^2_{i-1} - L_i(\lambda^2_{i-1} + \rho_{i-1})$$
$$\Rightarrow \beta_i = (\lambda^2_{i-1} + \rho_{i-1})$$

(vi) Find $\gamma_i$ in

$$\rho_i = a\rho_{i-1} - L_i\alpha_i - K_i\beta_i + K_iL_i\gamma_i$$
$$\rho_i = \mathbb{E}[(X_{i+1} - x_{i+1|i})(\Omega - \hat{\Omega}_i)]$$
$$= \mathbb{E}[(aX_i - a\hat{x}_{i|i-1} - K_i\hat{Y}_i)(\Omega - \hat{\Omega}_{i-1} - L_i\hat{Y}_i)]$$
$$= a\rho_{i-1} - aL_i(\sigma^2_{i|i-1} + \rho_{i-1}) - K_i(\lambda^2_{i-1} + \rho_{i-1})$$
$$+ K_iL_i(\lambda^2_{i-1} + \sigma^2_{i|i-1} + 2\rho_{i-1}).$$