Aula 4 – Autômatos Temporais Estocástico (1)

(Capítulo 6 – Cassandras)

1. Introdução

Como na realidade os sistemas de eventos discretos existem na presença de incertezas, modelos mais refinados são exigidos, incorporando elementos estocásticos.

2. Processo Estocástico

Um processo estocástico, ou randômico, é simplesmente um conjunto de variáveis randômicas indexadas por meio de algum parâmetro. (normalmente o "tempo").

Definição: Um processo estocástico ou randômico $X(\omega, t)$ é uma coleção de variáveis randômicas indexadas por t. As variáveis randômicas são definidas sobre um espaço de probabilidade (Ω, E, P) com $\omega \in \Omega$.

Ω ... conjunto de variáveis randômicas (Espaço Amostral)

E — subconjunto de Ω (Espaço de Eventos de interesse)

P — Probabilidade ($A \in E \rightarrow P[A]$)

t — varia dentro de um conjunto $T \subseteq \mathbb{R}$

a) Processo Estocástico de Estado-Contínuo x Estado-Discreto

Estado-Discreto: $X(t)$ definido sobre um conjunto finito ou contável de estados. Em geral um processo de estado-discreto é denominado por Cadeia (Chain).

Caso contrário: Estado-Contínuo

b) Processo Estocástico Tempo-Contínuo x Tempo-Discreto

Tempo-Discreto: O conjunto T, que especifica os valores permitidos da variável indexadora t, é finito ou contável.

Caso contrário, tempo contínuo.

c) Classes Importantes de Processo Estocástico

Para caracterizar uma variável randômica X, normalmente é especificada uma função distribuição de probabilidade cumulativa ("cumulative distribution function") - cdf

$$F_X(x) = P[X \leq x]$$

Seja o vetor randômico: $X = [X(t_0), X(t_1) \ldots X(t_n)]$

que pode assumir os valores $\mathbf{x} = [x_0, x_1, \ldots x_n]$

$$\Rightarrow F_X(x_0, \ldots x_n; t_0, \ldots t_n) = P[X(t_0) \leq x_0, \ldots, X(t_n) \leq x_n]$$

Algumas características são mais relevantes quando se modela os Sistemas de Eventos Discretos.

c.1) Processos Estacionários

Estacionária é a propriedade de um sistema em manter seu comportamento dinâmico invariante no tempo.

$$F_X(x_0, \ldots x_n; t_0 + \tau, \ldots, t_n + \tau) = F_X(x_0, \ldots x_n; t_0, \ldots t_n)$$

para $\forall \tau \in \mathbb{R}$

c.2) Processos Independentes

Seja a sequencia de variáveis randômicas:

$\{X_0, X_1, \dots X_n\}$

Se diz que trata-se de processos independentes se:

$$F_X(x_0, \dots, x_n; t_0, \dots t_n) = F_{X_0}(x_0; t_0) \dots \dots F_{X_n}(x_n; t_n)$$

Neste caso $\{X_K\}$ representa uma sequencia de variáveis randômicas independentes e distribuidas identicamente (sequencia iid) se também todos tiverem a mesma função distribuição.

c.3) Processo de Markov

Suponha que observemos uma cadeia de t_0 até t_K, onde $t_0 \leq t_1 \leq \dots \leq t_{K-1} \leq t_K$.

Observamos o valor x_K no instante t_K. ("Estado presente")

História Passada: $\{x_0, x_1, \dots, x_{K-1}\}$

Futuro Desconhecido: $\{x_{K+1}, x_{K+2}, \dots\}$

Numa Cadeia de Markov existe a independencia, ou seja, o futuro é completamente independente do passado. Este fato é conhecido como Propriedade Sem memória, depende apenas do presente.

Formalmente, um processo é Markoviano se 4

$$\underline{P}[X(t_{K+1}) \leq x_{K+1} \mid X(t_K) = x_K, X(t_{K-1}) = x_{K-1}, \ldots, X(t_0) = x_0]$$

$$= \underline{P}[X(t_{K+1}) \leq x_{K+1} \mid X(t_K) = x_K]$$

(Espaço de Estado Contínuo)

ou

$$\underline{P}[X(t_{K+1}) = x_{K+1} \mid X(t_K) = x_K, X(t_{K-1}) = x_{K-1}, \ldots X(t_0) = x_0]$$

$$= \underline{P}[X(t_{K+1}) = x_{K+1} \mid X(t_K) = x_K]$$

(Espaço de Estado Discreto)

No caso de Cadeias de Markov em tempo-discreto, as transições de estado estão limitadas aos instantes de tempo 0, 1, 2, .. K, ..

Assim a propriedade sem memória tem dois aspectos:

(M1) Todas informações dos estados passados são irrelevantes

(M2) Quanto tempo o processo está no estado presente também é irrelevante, ou seja, sua idade é irrelevante.

Num Sistema de Evento ~~Estado~~ Discreto, as transições de estado são causadas por ocorrência de eventos. Dessa forma, as variáveis randômicas com propriedade sem memória ~~são causadas por~~ correspondem aos tempos entre eventos.

Há diversos sistemas reais que essa propriedade é válida, ou pode ser justificada em larga escala.

3. Estrutura de Relógio Estocástico (Clock)

Nos Autômatos Temporais foi vista a estrutura do tempo como: $V = \{v_i : i \in E\}$ onde

$v_i = \{v_{i1}, v_{i2}, \ldots\}$ corresponde a sequencia de clock para o evento $i \in E$.

O objetivo agora é extender este modelo para incorporar as características estocásticas dos Sistemas de Eventos Discretos.

Desta forma, teremos que:

$$\{V_{i,k}\} = \{\underset{\substack{\uparrow \\ \text{randômico}}}{V_{i,1}}, V_{i2}, V_{i3}, \ldots\}$$

Definição: Autômato Temporal Estocástico

$$(\mathcal{E}, X, \Gamma, p, p_0, G)$$

$\mathcal{E}$ -- Conjunto de Eventos contáveis

X -- Espaço de Estados contáveis

$\Gamma(x)$ -- função de eventos permitidos

$p(x'; \underset{\substack{\uparrow \\ \text{atual}}}{x}, e')$ -- probabilidade de transição de estado

$p_0(x)$ -- probabilidade de estar no estado inicial X_0

G -- Estrutura do Clock Estocástico

O Processo de Contagem de Poisson

Considere um Sistema de Evento Discreto com um único evento. Adotemos $\{N(t)\}$ como sendo o processo que conta o número de ocorrências deste evento no intervalo $(0,t]$.

$$N(0) \leq N(t_1) \leq \cdots \leq N(t_K),$$

$$0 \leq t_1 \leq \cdots \leq t_K.$$

$$N(t_{K-1}, t_K) = N(t_K) - N(t_{K-1}), \quad K = 1, 2, \ldots$$

conta o número de ocorrências do evento no intervalo (t_{K-1}, t_K).

A probabilidade de n eventos ocorrerem em $(0,t]$ será dada por: $P_n(t) = P[N(t) = n]$, $n = 0, 1, \ldots$

Hipóteses:

(1) Apenas 1 evento pode ocorrer num instante de tempo

(2) As variáveis aleatórias $N(t)$, $N(t,t_1)$, $N(t_1,t_2)$, ..., $N(t_{K-1}, t_K)$, ... são mutuamente independentes para $0 \leq t \leq t_1 \ldots \leq t_K$ e qualquer $K = 1, 2, \ldots$

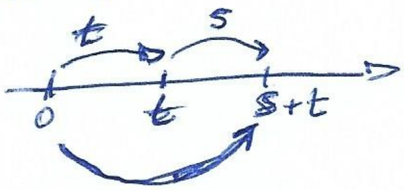
Os eventos futuros em algum intervalo futuro (t_{K-1}, t_K) não são afetados pelo que aconteceu no passado, antes de t_{K-1}. [Processo com Incrementos Independentes]

(3) $P[N(t_{K-1}, t_K) = n]$, $n = 0, 1, \ldots$, pode depender do comprimento do intervalo $(t_K - t_{K-1})$, mas não do instante t_{K-1}, para todo intervalo. [Processo com Incrementos Independentes Estacionários]

Objetivo: Calcularmos $P_n(t) = P[N(t)=n]$

Passo 1: Determinar $P_0(t) = P[N(t)=0]$

$$P_0(t+s) = P_0(t) \cdot P_0(s)$$



$$\frac{d}{ds} P_0(t+s) = P_0(t) \cdot \underbrace{\frac{dP_0(s)}{ds}}_{P_0'(s)}$$

Fazemos $s=0 \Rightarrow \dfrac{dP_0(t)}{ds} = P_0(t) \cdot \underbrace{P_0'(0)}_{cte}$

$$\Rightarrow \frac{dP_0(t)}{ds} = cte \cdot P_0(t) \Rightarrow P_0(t) = e^{cte \cdot t}$$

Para satisfazer $P_0(t) \leq 1$ para todo $t \geq 0 \Rightarrow cte < 0$
$(cte = -\lambda)$

$$\Rightarrow \boxed{P_0(t) = e^{-\lambda t}}$$

Passo 2: Determinar $P[N(\Delta t)=0]$ para um pequeno intervalo Δt.

$$P[N(\Delta t)=0] = e^{-\lambda \Delta t} = 1 - \lambda \Delta t + \lambda^2 \frac{(\Delta t)^2}{2!} - \lambda^3 \frac{(\Delta t)^3}{3!} \cdots$$

[Série de Taylor]

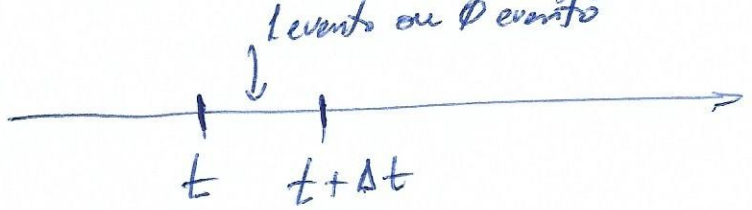
Quando $\Delta t \to 0 \Rightarrow \boxed{P[N(\Delta t)=0] \simeq 1 - \lambda \Delta t}$

Passo 3: Determinar $P[N(\Delta t)=n]$, $n=1,2,\ldots$ para pequeno Δt.

Como Δt muito pequeno, vamos assumir baseado em (1) que no intervalo Δt acontece apenas um evento (no máximo)

$$\Rightarrow P[N(\Delta t)=1] = 1 - (1 - \lambda \Delta t) \Rightarrow \boxed{P[N(\Delta t)=1] = \lambda \Delta t}$$

Passo 4: Derivar a expressão $P[N(t+\Delta t)=n]$ em termos de $P[N(t)=n]$. (Δt muito pequeno.)



$$P_n(t+\Delta t) = P_n(t)\,\underbrace{(1-\lambda\Delta t)}_{\text{zero eventos}} + P_{n-1}(t)\cdot\underbrace{\lambda\Delta t}_{\text{1 evento}}$$

Passo 5: Fazer $\lim \Delta t \to 0$

$$\frac{P_n(t+\Delta t) - P_n(t)}{\Delta t} = -\lambda \cdot P_n(t) + \lambda P_{n-1}(t)$$

$\Downarrow \lim \Delta t \to 0$

$$\boxed{\frac{dP_n(t)}{dt} = -\lambda \cdot P_n(t) + \lambda \cdot P_{n-1}(t)}$$

Passo 6: Resolver a equação diferencial

Para $n=1 \Rightarrow \dfrac{dP_1(t)}{dt} = -\lambda \cdot P_1(t) + \lambda \cdot P_0(t)$

$$\frac{dP_1(t)}{dt} = -\lambda \cdot P_1(t) + \lambda \cdot e^{-\lambda t} \Rightarrow \boxed{P_1(t) = (\lambda t)\cdot e^{-\lambda t}}$$

para $n=2,3,\dots \Rightarrow \boxed{P_n(t) = \dfrac{(\lambda t)^n}{n!}\cdot e^{-\lambda t}}$; $t \geq 0$

$n = 0,1,2,3\dots$

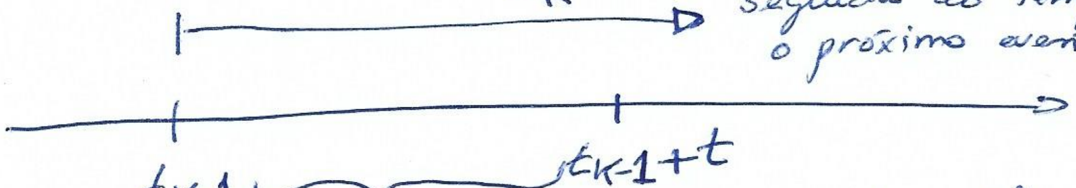
Esta expressão é conhecida como <u>Distribuição de Poisson</u> (com hipóteses (1), (2) e (3))

$$\begin{cases} E[N(t)] = \lambda t & \text{— Média (*)} \\ Var[N(t)] = \lambda t & \text{— Variância} \end{cases}$$

Propriedades do Processo de Poisson

a) Tempo entre Eventos distribuído Exponencialmente

v_K → variável randômica de tempo seguida ao tempo t_{K-1} até o próximo evento.



t_{K-1} $\underbrace{\qquad}$ → $N(t_{K-1}, t_{K-1}+t) = 0$

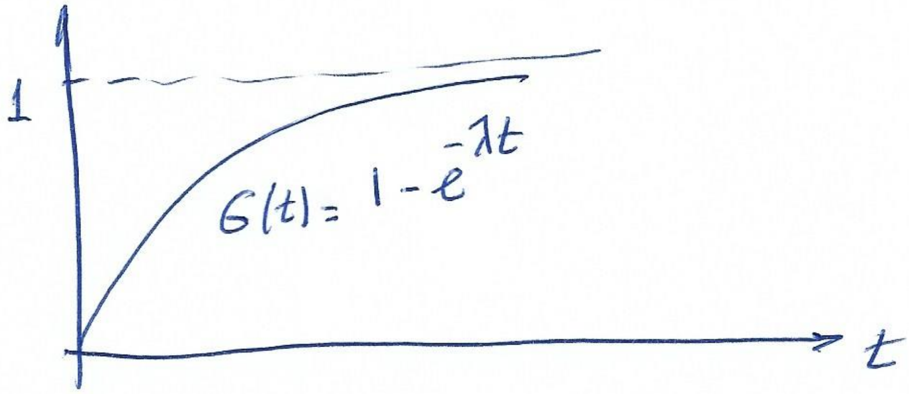
já ocorreram $(K-1)$ events

Seja $G_K(t) = P[v_K \le t] = 1 - P[v_K > t]$

$$\Downarrow \; = P(N(t_{K-1}, t_{K-1}+t) = 0) = P_0(t)$$

$\Rightarrow \underline{G_K(t) = 1 - e^{-\lambda t}}$ probabilidade de ocorrer um evento até t.

(função distribuição de probabilidade cumulativa)

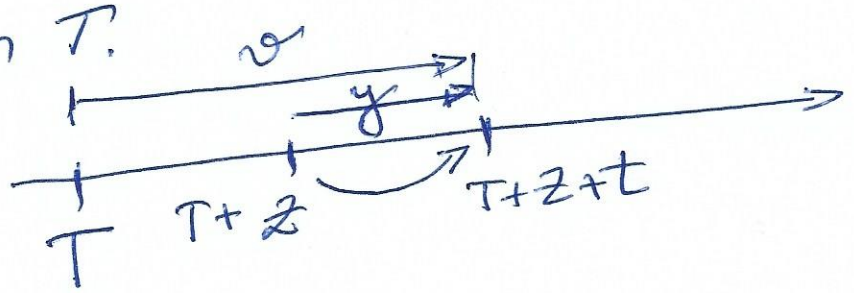


(*) $\Rightarrow \boxed{\lambda = \dfrac{E[N(t)]}{t}}$ Intuitivamente podemos <u>interpretar</u> o <u>parâmetro λ</u> como sendo a <u>taxa média</u> em que os eventos ocorrem por unidade de tempo.

No caso de eventos de falha, λ passa a representar <u>taxa de falha</u>.

A Propriedade "Sem Memória"

Vamos supor que no instante T ocorreu o último evento observado,
Quer-se calcular a probabilidade do próximo evento, a partir T, sabendo-se que ele não ocorreu até o tempo z depois da ocorrência em T.



$$P[\vartheta \leq z+t \mid \vartheta > z]$$

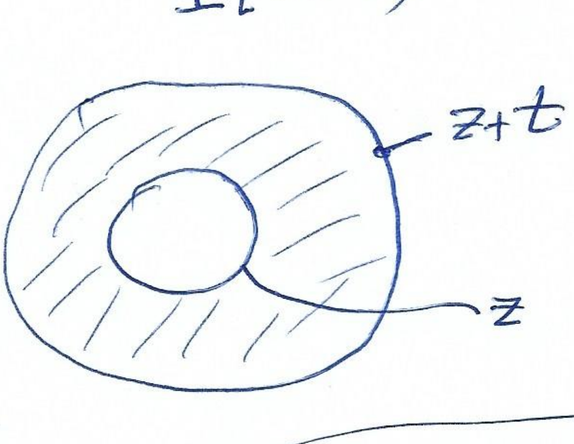
Da teoria de probabilidade: $P(A \cap B) = P(A|B) \cdot P(B)$

$$\Rightarrow P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Aplicando-se tem:

$$P[\vartheta \leq z+t \mid \vartheta > z] = \frac{P[\vartheta \leq z+t \text{ e } \vartheta > z]}{P[\vartheta > z]}$$

$$= \frac{P[z < \vartheta \leq z+t]}{P[\vartheta > z]}$$



$$= \frac{(\cancel{1} - e^{-\lambda(z+t)}) - (\cancel{1} - e^{-\lambda z})}{e^{-\lambda z}}$$

$$= \frac{e^{-\lambda z} - e^{-\lambda z} \cdot e^{-\lambda t}}{e^{-\lambda z}} \Rightarrow \boxed{P[\vartheta \leq z+t \mid \vartheta > z] = 1 - e^{-\lambda t}}$$

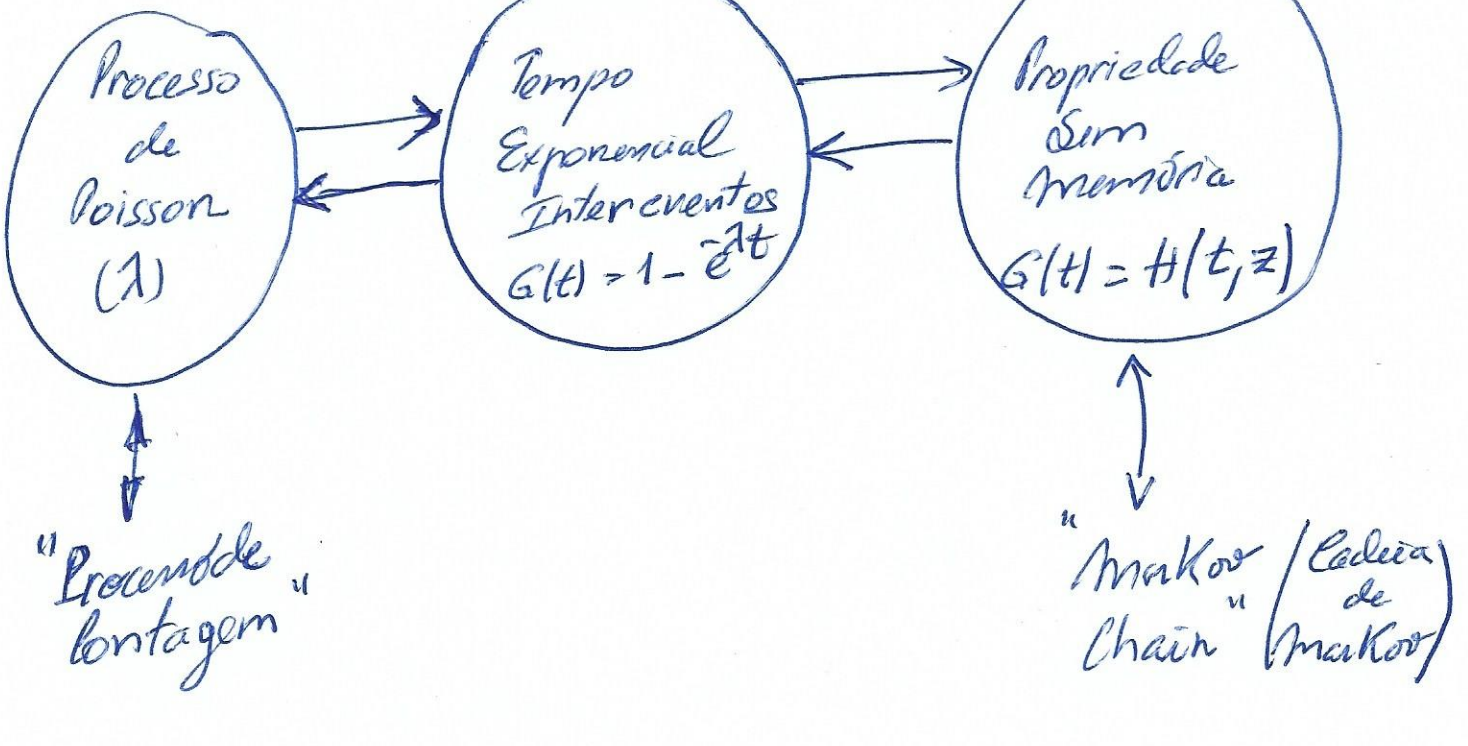
Ou seja, a probabilidade cumulativa da ocorrência do próximo evento até t independe da idade (z) do ~~último~~ próximo evento. Isso é o que se denomina de propriedade "sem memória" da distribuição exponencial.

$\Downarrow$

Definição: Cada tempo de vida residual (clock), num Processo de Poisson, é caracterizado pela mesma distribuição exponencial do tempo de vida original:

$$H(t,z) = P[Y \le t \mid Y > z] = \underline{G(t) = 1 - e^{-\lambda t}}$$

independentemente da idade do evento (z).



Superposição de Processos de Poisson

Nós havíamos considerado no Sistema de Evento Discreto com apenas um único evento.

Vamos considerar agora este Sistema de Evento Discreto com <u>m</u> eventos, $m > 1$.

Assume-se que cada evento é modelado por um processo de Poisson com parâmetros λ_i, $i = 1, \ldots, m$.

Assume-se também que estes <u>m</u> processos de Poisson são <u>mutuamente</u> <u>independentes</u>.

Pergunta: o processo resultante continua sendo de Poisson?

$$P[Y^* \leq t] = 1 - P[Y^* > t]$$

$$P[Y^* \leq t] = 1 - \left[\prod_{i=1}^{m} P[Y_i > t]\right]$$

$$P[Y^* \leq t] = 1 - \prod_{i=1}^{m} e^{-\lambda_i t}$$

$$\boxed{P[Y^* \leq t] = 1 - e^{-\Lambda t}}$$

onde $\boxed{\Lambda = \sum_{i=1}^{m} \lambda_i}$

Ou seja, a superposição de <u>m</u> Processos de Poisson Independentes, com parâmetros λ_i; $i = 1, \ldots m$; é também um Processo de Poisson com parâmetro

$$\underline{\Lambda = \sum_{i=1}^{m} \lambda_i}$$