Generalized Semi-Markov Processes (GSMP)

Summary

- Some Definitions
- Markov and Semi-Markov Processes
- Automaton Review
- The Poisson Process
- Properties of the Poisson Process
  - Interarrival times
  - Memoryless property and the residual lifetime paradox
  - Superposition of Poisson processes
Random Process

- Let \((\Omega, F, P)\) be a probability space. A stochastic (or random) process \(\{X(t)\}\) is a collection of random variables defined on \((\Omega, F, P)\), indexed by \(t \in T\) (where \(t\) is usually time). \(X(t)\) is the state of the process.

- Continuous Time and Discrete Time stochastic processes
  - If the set \(T\) is finite or countable then \(\{X(t)\}\) is called \textit{discrete-time process}. In this case \(t \in \{0, 1, 2, \ldots\}\) and we may referred to a \textit{stochastic sequence}. We may also use the notation \(\{X_k\}, k=0,1,2,\ldots\)
  - Otherwise, the process is called \textit{continuous-time process}.

- Continuous State and Discrete State stochastic processes
  - If \(\{X(t)\}\) is defined over a countable set, then the process is \textit{discrete-state}, also referred to as \textit{chain}.
  - Otherwise, the process is \textit{continuous-state}.

Classification of Random Processes

- Joint cdf of the random variables \(X(t_0),\ldots,X(t_n)\)
  \[
  F_X \left( x_0, \ldots, x_n; t_0, \ldots, t_n \right) = \Pr \left\{ X(t_0) \leq x_0, \ldots, X(t_n) \leq x_n \right\}
  \]

- Independent Process
  - Let \(X_1,\ldots,X_n\) be a sequence of independent random variables, then
  \[
  F_X \left( x_0, \ldots, x_n; t_0, \ldots, t_n \right) = F_{X_0} \left( x_0; t_0 \right) \times \ldots \times F_{X_n} \left( x_n; t_n \right)
  \]

- Stationary Process (strict sense stationarity)
  - The sequence \(\{X_n\}\) is stationary if and only if for any \(\tau \in R\)
  \[
  F_X \left( x_0, \ldots, x_n; t_0 + \tau, \ldots, t_n + \tau \right) = F_X \left( x_0, \ldots, x_n; t_0, \ldots, t_n \right)
  \]
Classification of Random Processes

- **Wide-sense Stationarity**
  - Let $C$ be a constant and $g(\tau)$ a function of $\tau$ but not of $t$, then a process is wide-sense stationary if and only if
  \[
  E\{X(t)\} = C \quad \text{and} \quad E\{X(t) \cdot X(t + \tau)\} = g(\tau)
  \]

- **Markov Process**
  - The future of a process does not depend on its past, only on its present.
  \[
  \Pr\{X(t_{k+1}) \leq x_{k+1} \mid X(t_k) = x_k, \ldots, X(t_0) = x_0\} = \Pr\{X(t_{k+1}) \leq x_{k+1} \mid X(t_k) = x_k\}
  \]
  - Also referred to as the Markov property

---

Markov and Semi-Markov Property

- **The Markov Property** requires that the process has no memory of the past. This **memoryless** property has two aspects:
  - All past state information is irrelevant in determining the future (no state memory).
  - How long the process has been in the current state is also irrelevant (no state age memory).
  - The later implies that the lifetimes between subsequent events (inter-event time) should also have the memoryless property (i.e., exponentially distributed).

- **Semi-Markov Processes**
  - For this class of processes the requirement that the state age is irrelevant is relaxed, therefore, the inter-event time is no longer required to be exponentially distributed.
Example

Consider the process

\[ X_{k+1} = X_k - X_{k-1} \]

with \( \Pr\{X_0=0\} = \Pr\{X_0=1\} = 0.5 \) and \( \Pr\{X_1=0\} = \Pr\{X_1=1\} = 0.5 \)

- Is this a Markov process? NO
- Is it possible to make the process Markov?
- Define \( Y_k = X_{k-1} \) and form the vector \( Z_k = [X_k, Y_k]^T \) then we can write

\[
Z_{k+1} = \begin{bmatrix} X_{k+1} \\ Y_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} X_k \\ Y_k \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} Z_k
\]

Renewal Process

- A renewal process is a chain \( \{N(t)\} \) with state space \( \{0,1,2,...\} \) whose purpose is to count state transitions. The time intervals between state transitions are assumed iid from an arbitrary distribution. Therefore, for any \( 0 \leq t_1 \leq ... \leq t_k \leq ... \)

\[ N(0) = 0 \leq N(t_1) \leq N(t_2) \leq ... \leq N(t_k) \leq ... \]
Deterministic Automaton

\[ G = (X, E, f, \Gamma, x_0, X_m) \]

- \( X \) is the state space
- \( E \) is the event set associated with \( G \)
- \( f: X \times E \rightarrow X \) is the transition function
  - \( f(x,e) = y \) means that when \( G \) is at state \( x \) and event \( e \) occurs then \( G \) transitions to \( y \)
- \( \Gamma \) is the feasible event set.
  - For every state \( x \), \( \Gamma(x) \) is the set of all events that are active
- \( x_0 \) is the initial state
- \( X_m \) is the set of marked states and is a subset of \( X \)

Example: Automaton Graphical Representation

[Diagram of an automaton with states I, D, and B, and transitions labeled a, b, c, r, f, and an arrow pointing from I to B and B to D]
A non-deterministic automaton is similar to the deterministic except:
- $f_{nd}$ may probabilistically transition to a set of states.
- $x_0$ the initial state may also be a set of states.

Example

FSM for Stop and Wait Protocol

Transmitter
**FSM for Stop and Wait Protocol**

**Receiver**

- Pkt 0 received
  - Send Ack 0

- Pkt 1 received
  - Wait Pkt 1
  - Pkt 1 received
  - Send Ack 1

- Wait Pkt 0

**Timed Automata Models**

- So far, the models we studied did not consider the time when an event occurred.
- For dynamic systems, the sample path of the system is specified by the sequence of pairs \( \{(e_i, t_i)\} \) where \( e_i \) is the \( i \)-th event and \( t_i \) is the time that the event has occurred.
- Timed automata are similar to the deterministic automata with the addition of a **clock structure**

\[
G = (X, E, f, \Gamma, x_0, V)
\]

- \( V \) is the clock structure
- \( X_m \) (the set of marked states) is omitted for simplicity
Simple Example

Assume a system with a single event, i.e., $E = \{a\}$ and $\Gamma(x) = \{a\}$ for all $x \in X$. We are interested in the sequence $\{(e_i, t_i)\}$ where $e_i$ is the $i$-th event and $t_i$ is the time when $e_i$ occurs.

- $v_k = t_k - t_{k-1}$ is the event lifetime
- $z_k = t - t_{k-1}$ is the age of event
- $y_k = t_k - t$ is the residual lifetime

Generalized Semi-Markov Processes (GSMP)

- A GSMP is a stochastic process $\{X(t)\}$ with state space $X$ generated by a stochastic timed automaton $\left( X, E, \Gamma, f^f, p_0, G \right)$
  - $X$ is the countable state space
  - $E$ is the countable event set
  - $\Gamma(x)$ is the feasible event set at state $x$.
  - $f(x, e)$: is state transition function.
  - $p_0$ is the probability mass function of the initial state
  - $G$ is a vector with the cdfs of all events.

The semi-Markov property of GSMP is due to the fact that at the state transition instant, the next state is determined by the current state and the event that just occurred.
The Poisson Counting Process

Let the process \( \{N(t)\} \) which counts the number of events that have occurred in the interval \([0,t]\). For any \(0 \leq t_1 \leq \ldots \leq t_k \leq \ldots \) \(N(0) = 0 \leq N(t_1) \leq N(t_2) \leq \ldots \leq N(t_k) \leq \ldots\)

- Process with independent increments: The random variables \(N(t_1), N(t_1,t_2), \ldots, N(t_{k-1},t_k), \ldots\) are mutually independent.
- Process with stationary independent increments: The random variable \(N(t_{k-1}, t_k)\) does not depend on \(t_{k-1}, t_k\) but only on \(t_k - t_{k-1}\)

\[N(t_{k-1}, t_k) = N(t_k) - N(t_{k-1})\]

The Poisson Counting Process

Assumptions:
- At most one event can occur at any time instant (no two or more events can occur at the same time)
- A process with stationary independent increments

\[
\Pr \{N(t_{k-1}, t_k) = n\} = \Pr \{N(t_k - t_{k-1}) = n\}
\]

Given that a process satisfies the above assumptions, find

\[
P_n(t) \equiv \Pr \{N(t) = n\}, n = 0, 1, 2, \ldots
\]
The Poisson Process

- **Step 1**: Determine
  \[ P_0(t) = \Pr\{N(t) = 0\} \]
  Starting from
  \[
  \Pr\{N(t+s) = 0\} = \Pr\{N(t) = 0\} \text{ and } N(t, t+s) = 0
  \]
  \[= \Pr\{N(t) = 0\} \Pr\{N(s) = 0\} \]
  \[\Rightarrow P_0(t+s) = P_0(t)P_0(s)\]

- **Lemma**: Let \( g(t) \) be a differentiable function for all \( t \geq 0 \) such that \( g(0) = 1 \) and \( g(t) \leq 1 \) for all \( t > 0 \). Then for any \( t, s \geq 0 \)
  \[ g(t+s) = g(t)g(s) \iff g(t) = e^{-\lambda t} \text{ for some } \lambda > 0 \]

The Poisson Process

- Therefore
  \[ P_0(t) = \Pr\{N(t) = 0\} = e^{-\lambda t} \]

- **Step 2**: Determine \( P_0(\Delta t) \) for a small \( \Delta t \).
  \[ \Pr\{N(\Delta t) = 0\} = e^{-\lambda \Delta t} = 1 - \lambda \Delta t + \frac{(\lambda \Delta t)^2}{2!} - \frac{(\lambda \Delta t)^3}{3!} + \ldots \]
  \[= 1 - \lambda \Delta t + o(\Delta t). \]

- **Step 3**: Determine \( P_n(\Delta t) \) for a small \( \Delta t \).
  For \( n = 2, 3, \ldots \) since by assumption no two events can occur at the same time
  \[ P_n(\Delta t) = \Pr\{N(\Delta t) = n\} = o(\Delta t) \]
  As a result, for \( n = 1 \)
  \[ P_1(\Delta t) = \Pr\{N(\Delta t) = 1\} = \lambda \Delta t + o(\Delta t) \]
Step 4: Determine $P_n(t + \Delta t)$ for any $n$

$$P_n(t + \Delta t) \equiv \Pr \{ N(t + \Delta t) = n \} = \sum_{k=0}^{n} P_{n-k} (t) P_k (\Delta t)$$

$$= P_n (t) P_0 (\Delta t) + P_{n-1} (t) P_1 (\Delta t) + o(\Delta t).$$

$$= [1 - \lambda \Delta t + o(\Delta t)] P_n (t) + [\lambda \Delta t + o(\Delta t)] P_{n-1} (t) + o(\Delta t)$$

Moving terms between sides,

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

Taking the limit as $\Delta t \to 0$

$$\frac{dP_n (t)}{dt} = -\lambda P_n (t) + \lambda P_{n-1} (t)$$
The Poisson Process

Step 5: Solve the differential equation to obtain

\[ P_n(t) = \Pr \{ N(t) = n \} = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \quad t \geq 0, \quad n = 0, 1, 2, \ldots \]

This expression is known as the Poisson distribution and it fully characterizes the stochastic process \( \{N(t)\} \) in \([0,t)\) under the assumptions that

- No two events can occur at exactly the same time, and
- Independent stationary increments

You should verify that

\[ E[N(t)] = \lambda t \quad \text{and} \quad \text{var}[N(t)] = \lambda t \]

Parameter \( \lambda \) has the interpretation of the “rate” that events arrive.

Properties of the Poisson Process: Interevent Times

Let \( t_{k-1} \) be the time when the \( k-1 \) event has occurred and let \( V_k \) denote the (random variable) interevent time between the \( k \)-th and \( k-1 \) events.

What is the cdf of \( V_k \), \( G_k(t) \)?

\[
G_k(t) = \Pr \{ V_k \leq t \} = 1 - \Pr \{ V_k > t \} = 1 - \Pr \{ 0 \text{ arrivals in the interval } [t_{k-1}, t_{k-1} + t) \} = 1 - e^{-\lambda t} \]

\[ \Rightarrow G(t) = 1 - e^{-\lambda t} \]

Exponential Distribution

Stationary independent increments
Properties of the Poisson Process:
Exponential Interevent Times

- The process \( \{V_k\}_{k=1,2,...} \) that corresponds to the
  interevent times of a Poisson process is an iid stochastic
  sequence with cdf
  \[
  G(t) = \Pr\{V_k \leq t\} = 1 - e^{-\lambda t}
  \]
- The corresponding pdf is
  \[
  g(t) = \lambda e^{-\lambda t}, \quad t \geq 0
  \]
- One can easily show that
  \[
  E[V_k] = \frac{1}{\lambda} \quad \text{and} \quad \text{var}[V_k] = \frac{1}{\lambda^2}
  \]

Properties of the Poisson Process:
Memoryless Property

- Let \( t \) be the time when previous event has occurred and let \( V \) denote
  the time until the next event.
- Assuming that we have been at the current state for \( z \) time units, let \( Y \)
  be the remaining time until the next event.
- What is the cdf of \( Y \)?
  \[
  F_Y(t) = \Pr\{Y \leq t\} = \Pr\{V - z < t \mid V > z\}
  \]
  \[
  = \frac{\Pr\{V > z \text{ and } V < z + t\}}{\Pr\{V > z\}} = \frac{1 - \Pr\{V < z + t\}}{1 - \Pr\{V < z\}}
  \]
  \[
  = \frac{G(t + z) - G(z)}{1 - G(z)} = \frac{1 - e^{-\lambda(t+z)}}{1 - 1 + e^{-\lambda z}}
  \]
  \[
  \Rightarrow F_Y(t) = 1 + e^{-\lambda t} = G(t)
  \]

Memoryless! It does not matter that we have already spent \( z \) time units at the
current state.
Memoryless Property

This is a unique property of the exponential distribution. If a process has the memoryless property, then it must be exponential, i.e.,

\[ \Pr \{ V \leq z + t \mid V > z \} = \Pr \{ V \leq t \} \iff \Pr \{ V \leq t \} = 1 - e^{-\lambda t} \]

Superposition of Poisson Processes

Consider a DES with \( m \) events each modeled as a Poisson Process with rate \( \lambda_i, i=1,...,m \). What is the resulting process?

- Suppose at time \( t_k \) we observe event 1. Let \( Y_1 \) be the time until the next event 1. Its cdf is \( G_1(t) = 1 - \exp(-\lambda_1 t) \).
- Let \( Y_2,\ldots,Y_m \) denote the residual time until the next occurrence of the corresponding event.
- Their cdfs are:

\[ G_i(t) = 1 - e^{-\lambda_i t} \]

Let \( Y^* \) be the time until the next event (any type).

\[ Y^* = \min \{ Y_i \} \]

Therefore, we need to find

\[ G_{Y^*}(t) = \Pr \{ Y^* \leq t \} \]
Superposition of Poisson Processes

$$G_{Y^*}(t) = \Pr\{Y^* \leq t\} = \Pr\{\min\{Y_i\} \leq t\}$$

$$= 1 - \Pr\{\min\{Y_i\} > t\}$$

$$= 1 - \Pr\{Y_i > t, \ldots, Y_m > t\}$$

$$= 1 - \prod_{i=1}^{m} \Pr\{Y_i > t\} = 1 - \prod_{i=1}^{m} e^{-\lambda_i t}$$

$$\Rightarrow G_{Y^*}(t) = 1 - e^{-\Lambda t} \quad \text{where} \quad \Lambda = \sum_{i=1}^{m} \lambda_i$$

- The superposition of $m$ Poisson processes is also a Poisson process with rate equal to the sum of the rates of the individual processes.

---

Superposition of Poisson Processes

- Suppose that at time $t_k$ an event has occurred. What is the probability that the next event to occur is event $j$?
- Without loss of generality, let $j = 1$ and define $Y' = \min\{Y_i; i=2,\ldots,m\}$. 

$$\Pr\{\text{next event is } j = 1\} = \Pr\{Y_i \leq Y'\} =$$

$$\int_{0}^{y} \int_{0}^{y'} \lambda_i e^{-\lambda_i y} \Lambda' e^{-\Lambda' y'} dy' dy'$$

$$= \int_{0}^{y} \Lambda' (1 - e^{-\lambda_i y}) e^{-\Lambda' y'} dy'$$

$$= \frac{\lambda_1}{\Lambda}$$

where \(\Lambda = \sum_{i=1}^{m} \lambda_i\)
Residual Lifetime Paradox

- Suppose that buses pass by the bus station according to a Poisson process with rate $\lambda$. A passenger arrives at the bus station at some random point.
- How long does the passenger have to wait?
- **Solution 1:**
  - $E[V] = 1/\lambda$. Therefore, since the passenger will (on average) arrive in the middle of the interval, he has to wait for $E[Y] = E[V]/2 = 1/(2\lambda)$.
  - But using the memoryless property, the time until the next bus is exponentially distributed with rate $\lambda$, therefore $E[Y] = 1/\lambda$ not $1/(2\lambda)$!
- **Solution 2:**
  - Using the memoryless property, the time until the next bus is exponentially distributed with rate $\lambda$, therefore $E[Y] = 1/\lambda$.
  - But note that $E[Z] = 1/\lambda$ therefore $E[V] = E[Z] + E[Y] = 2/\lambda$ not $1/\lambda$!