CSCE 619-600
Networks and Distributed Processing
Spring 2017

Renewal Process Theory III

Dmitri Loguinov
Texas A&M University

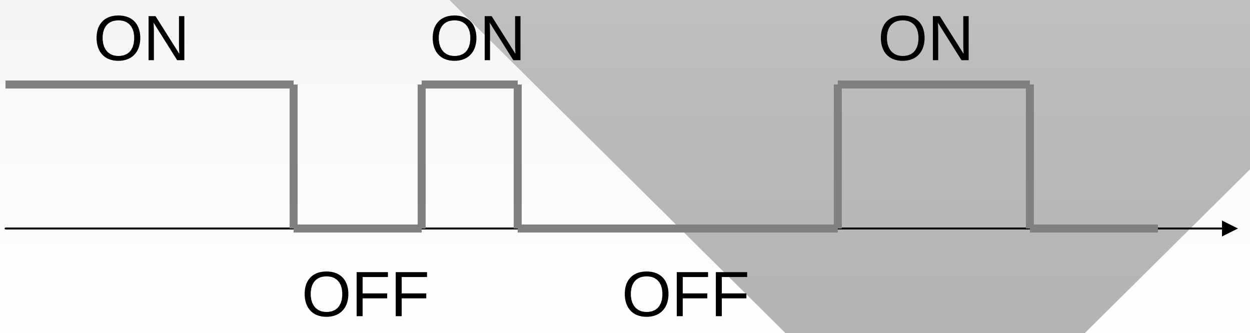
February 14, 2017

Agenda

- Single observation of renewal processes
- Blackwell's theorem
- Poisson processes
- PASTA
- Palm-Khintchine theorem

Single Observation

- Consider an example
  - A machine works for X_j minutes and then breaks down
  - Each repair takes Y_j minutes as soon as the failure is detected; $\{X_j\}$ and $\{Y_j\}$ are iid sequences
  - What is the probability the machine is working at random observation time t?
- This is an alternating (on/off) renewal process
  - Each cycle length is $X_j + Y_j$



Single Observation 2

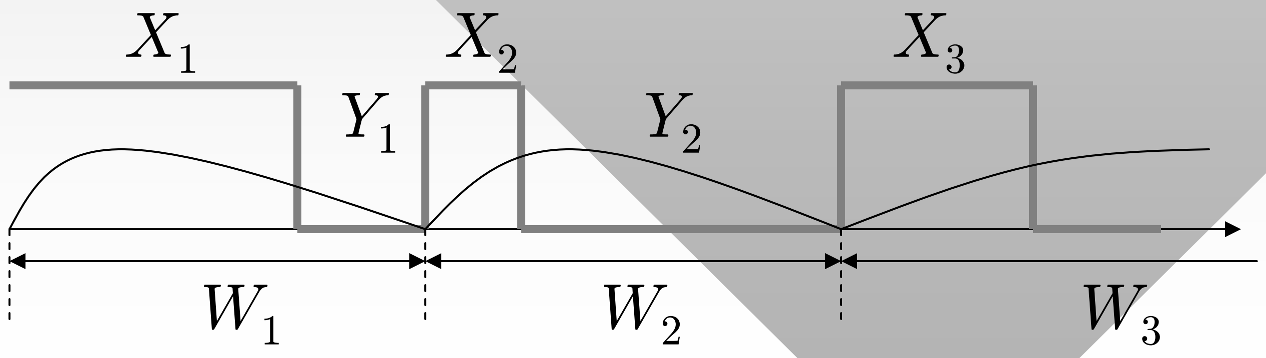
- Define on/off process $I(t)$:

$$I(t) = \begin{cases} 1 & \text{machine is ON at } t \\ 0 & \text{otherwise} \end{cases}$$

- and notice that for uniform $t \in [0, T]$

$$P(I(t) = 1) = \frac{1}{T}\int_0^T I(u)du$$

- Consider a renewal process with cycles $W_j = X_j + Y_j$



Single Observation 3

- If we view each ON duration as the amount of reward for cycle W_j (i.e., $P_j = X_j$), we can apply the renewal-reward theorem:

$$\lim_{t\to\infty} P(I(t)=1) = \lim_{T\to\infty} \frac{1}{T}\int_0^T I(t)dt$$
$$= \lim_{T\to\infty} \frac{C(T)}{T} = \frac{E[P_i]}{E[W_i]}$$

- Therefore: $\lim_{t\to\infty} P(I(t)=1) = \frac{E[X_i]}{E[X_i]+E[Y_i]}$
- Probability to see the object in state j is the ratio of the expected time in state j to the length of each cycle

Single Observation 4

- Knowing both $E[X_j]$ and $E[Y_j]$ allows us to predict
  - Fraction of time the machine is working
  - Probability to hit the ON interval at some large t
- However, what if the observer doesn't know these parameters, but still needs to estimate $P(I(t)=1)$?
- The problem becomes that of sampling $I(t)$
  - Recording $I(t_k)$ in a sequence to points t_1, t_2, …
  - Inferring $P(I(t)=1)$ from the samples
- We next develop the necessary tools for this question
  - We also understand the behavior of the simplest class of renewal processes (i.e., Poisson)

Poisson Process

- Poisson processes are perhaps the most important class of renewal processes
- As before, assume $M(t)$ is the number of arrivals in $[0,\, t]$ and X_j is the j-th inter-renewal delay
- Definition: process $M(t)$ *Poisson* if the following two properties are satisfied:
  1) The number of arrivals in disjoint intervals are independent random variables
  2) The PMF of $M(t)$ is given by

$$P(M(t) = j) = \frac{e^{-\lambda t}(\lambda t)^j}{j!}$$

Poisson Process 2

- First condition (independent increments) says that $M(t)$ and $[M(t{+}h) - M(t)]$ are independent of each other for all t and h
  - The past has no effect on the future
- Second condition says that the number of renewals in $[0,\, t]$ is Poisson-distributed with parameter λt (hence λ is the renewal rate per time unit)
- Using generating functions, one can show that
  - $[M(t{+}h) - M(t)]$ is also a Poisson random variable with parameter λh
  - Thus, the expected number of arrivals in any interval of size h is $m(t{+}h) - m(t) = m(h) = \lambda h$

Poisson Process 3

- Additional observations:
  - For general renewal processes, $M(t+h) - M(t)$ may have a different distribution from $M(h)$
  - Further, even the corresponding expectations are generally not equal: $m(t+h) - m(t) \neq m(h)$
- Example: $X_i \sim$ Pareto(3) $(E[X_i] = 0.5)$ and $h = 0.5$
  - From simulations, $m(h) = 1.29$
  - Restarting the process and re-sampling random variable $M(20+h) - M(20)$ we obtain its mean 0.9945
  - Notice that the latter value is close to the expected result $h\mu = h/E[X_j] = 1$, while the former one is not
- Blackwell's theorem: $\lim_{t\to\infty} [m(t+h) - m(t)] = \frac{h}{E[X_j]}$

Poisson Process 4

- Finally, it can be shown that the distribution of X_j in Poisson processes is always exponential

$$\begin{aligned} P(X_j > h) &= P(M(t+h) - M(h) = 0) \\ &= P(M(h) = 0) = e^{-\lambda h} \end{aligned}$$

- The reverse is also true and can be used to define the Poisson process
- Definition: Poisson process $\Leftrightarrow$ 1) renewal with exponential X_j or 2) independent increments *and* Poisson $M(t)$

Poisson Process 4

- At any time t:
  - Delay to next renewal (i.e., residual $R(t)$) is exponential with the same rate λ
  - Number of arrivals in $[t, t+h]$ is Poisson with parameter λh (i.e., independent of t)
- Poisson processes have been traditionally used to model all types of arrivals
  - Customers, phone calls, packets, etc.
  - Traditional queuing theory in the 1960s
- Important application of Poisson processes
  - Sampling general stochastic processes
  - Superposition of renewal processes

Practice

- Trains arrive to a station as a Poisson(λ_T) process
  - Henry arrives at some random time t
- What is the distribution of his wait time W_T to next train?
- Suppose Henry needs $Y \sim \exp(\mu)$ time units to buy the ticket
  - If train arrives before ticket is bought, Henry boards it
- Probability that Henry rides without a ticket?
- Suppose buses arrive to station as Poisson(λ_B)
  - Henry will take either the train or the bus, whichever comes first; define the delay to first bus as W_B
- Let Z be the delay before Henry leaves the station
  - Probability he rides a bus?
  - Probability he rides without a ticket (same ticket can be used on train or bus)?
  - Define Z using Y, W_T, W_B

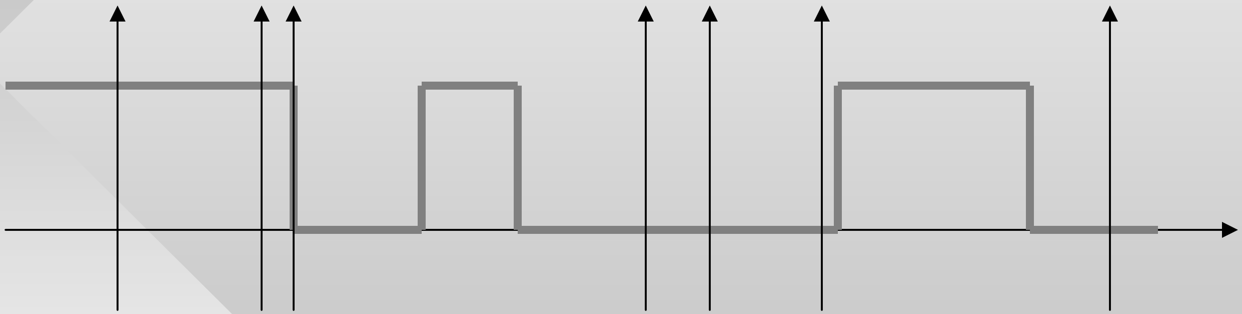
PASTA

- PASTA (Poisson Arrivals See Time Averages)
  - Principle for sampling processes without restarting them
- Definition: indicator variable $$1_A = \begin{cases} 1 & A \text{ occurred} \\ 0 & otherwise \end{cases}$$
- Assume that a discrete process $X(t)$ is sampled by a sequence of Poisson arrivals at points $\{t_i\}$
- Theorem: the fraction of time these arrivals observe $X(t)$ in a given state k converges to the actual probability that $X(t)$ is in that state

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_{\{X(t_i)=k\}} = \lim_{T\to\infty} \frac{1}{T} \int_0^T 1_{\{X(t)=k\}} dt$$

PASTA 2

- Example: machine is ON when working and OFF when being repaired



- Each sample at time t_i produces a value $I(t_i)$, where:

$$I(t) = \begin{cases} 1 & \text{machine is ON at } t \\ 0 & \text{otherwise} \end{cases}$$

PASTA 3

- Then PASTA says:
  - The fraction of inspections that observe the machine in the ON state is the fraction of time the machine is actually ON

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} I(t_i) = \lim_{T\to\infty} \frac{1}{T}\int_0^T I(t)dt = \frac{E[X_i]}{E[X_i]+E[Y_i]}$$

- Note that constant delay between observations does not work in general
  - Example: machine spends 1 hour in each state, inspections every 2 hours will not produce a correct result
- Even in adversarial scenarios (e.g., employees), PASTA cannot be fooled

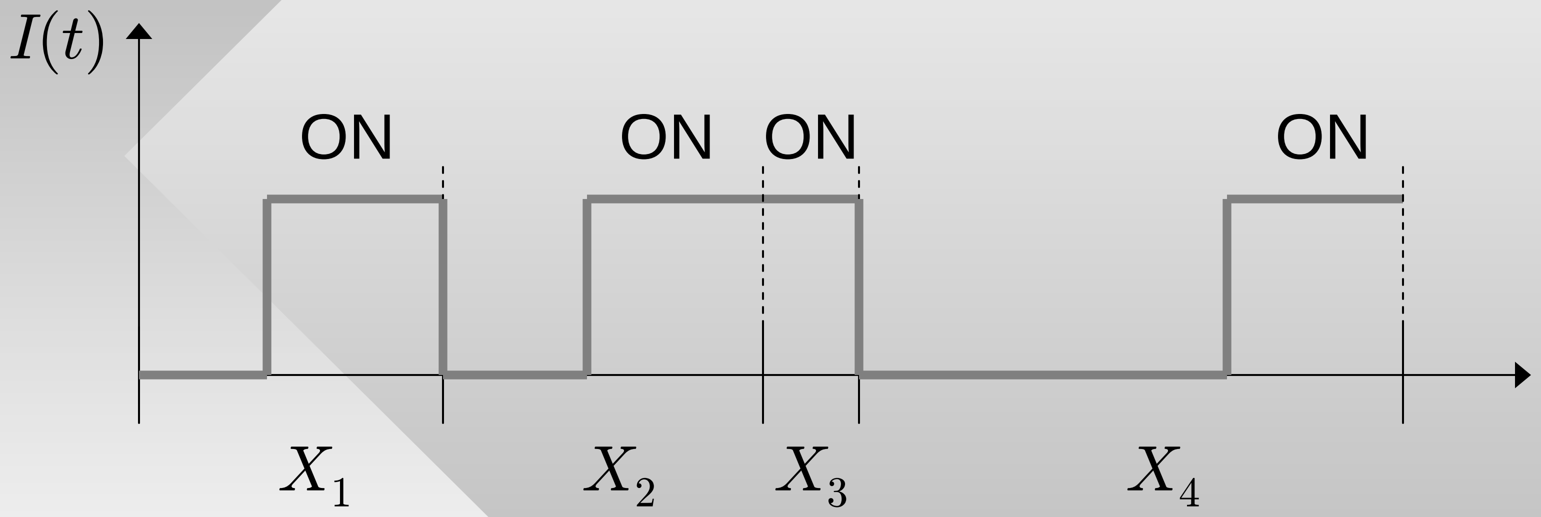
PASTA 4

- Example:
  - Trains arrive to a train station as a renewal process with inter-train delays X_1, X_2 , ... with some CDF $F(x)$
  - Customers arrive to the station as a Poisson(λ) process
  - Find out the fraction of passengers whose delay to the next train is less than some fixed y
- This problem can also be solved without directly applying the PASTA principle (page 76)
- Notice that the sampled process can be in two states:
  - The delay to the next train is larger than y (OFF)
  - The delay is smaller than y (ON)

$$I(t) = \begin{cases} 1 & R(t) \le y \\ 0 & \text{otherwise} \end{cases}$$

PASTA 5

- Recall the picture from the last lecture



- Thus, the fraction of customers that observe the process in the ON state is equal to the probability that the process is ON, i.e.,

$$P(R(t) \le y) = \frac{1}{E[X_i]} \int_0^y (1 - F(x))dx$$

ASTA

- ASTA (Arrivals See Time Averages) is slightly more general
  - It only requires that the delay between samples be a random variable that is independent of the process being sampled
- Assume that $Y_i = t_i - t_{i-1}$
  - If $\{Y_i\}$ are iid, continuous in $[0,c]$ for some $c > 0$ and independent of $X(t)$, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} I(t_i) = \lim_{T\to\infty} \frac{1}{T} \int_0^T I(t)dt$$

  - where

$$I(t) = \begin{cases} 1 & X(t) = k \\ 0 & \text{otherwise} \end{cases}$$

Superposition

- Consider a system of n renewal processes
  - Each renewal point is an arrival of a user into the system
- Example:
  - P2P system

$$Z_i(t) = \begin{cases} 1 & \text{user } i \text{ logged in at } t \\ 0 & \text{otherwise} \end{cases}$$

- Let X_{i1}, X_{i2}, ... be inter-renewal delays of user i, whose distribution is $F_i(x)$, and define $\mu_i > 0$ to be the corresponding average arrival rate
- Suppose $M_i(t)$ is the number of arrivals in $[0, t]$ from user i and $M(t)$ is that from all users:

$$M(t) = \sum_{i=1}^{n} M_i(t)$$

Superposition 2

- Palm-Khintchine Theorem: as $n \to \infty$, superposition process $M(t)$ converges to Poisson(μ) if

$$\mu = \lim_{n \to \infty} \sum_{i=1}^{n} \mu_i < \infty$$

- This explains why many large systems often assume Poisson arrivals even if individual users have non-exponential inter-arrival delays X_{ij}
- Important note: rate μ_i must tend to 0 as $n \to \infty$
  - In other words, users participate at lower rates as n increases (e.g., join rate $\sim 1/n$)