

Lecture 8: Random Processes, Part I

Yi, Yung (이웅)

EE210: Probability and Introductory Random Processes
KAIST EE

MONTH DAY, 2021

- A lot of applications in engineering systems
- Basics on Random Process
- Bernoulli Process
- Poisson Process
- Use of Bernoulli and Poisson Processes
 - Coding of both processes
 - Merge and Split
- Markov Chain

MONTH DAY, 2021 1 / 1

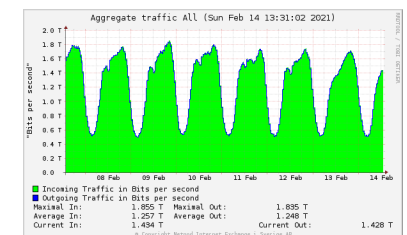
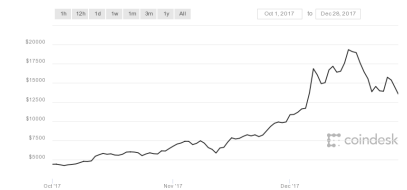
MONTH DAY, 2021 2 / 1

Roadmap

Things that evolve in time

- A lot of applications in engineering systems
- Basics on Random Process
- Bernoulli Process
- Poisson Process
- Use of Bernoulli and Poisson Processes
 - Coding of both processes
 - Merge and Split
- Markov Chain

- Many probabilistic experiments that **evolve in time**
 - Sequence of daily prices of a stock
 - Sequence of scores in football
 - Sequence of failure times of a machine
 - Sequence of traffic loads in Internet
- Random process is a mathematical model for it.



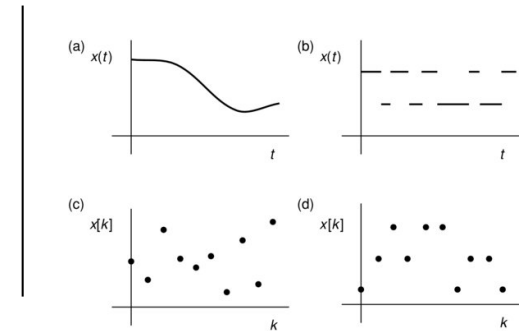
MONTH DAY, 2021 3 / 1

MONTH DAY, 2021 4 / 1

- A random process is a **sequence** of random variables indexed by **time**.
- Time: discrete or continuous
- Notation
 - $(X_t)_{t \in \mathcal{T}}$ or $(X(t))_{t \in \mathcal{T}}$, where $\mathcal{T} = \mathbb{R}$ (continuous) or $\mathcal{T} = \{0, 1, 2, \dots\}$ (discrete)
 - For the discrete case, we also often use $(X_n)_{n \in \mathbb{Z}_+}$.
 - We will use all of them, unless confusion arises.
- For a fixed time t , X_t is a random variable.
- The values that X_t can take: discrete or continuous

- Types of time and value

- (a) continuous time, continuous value
- (b) continuous time, discrete value
- (c) discrete time, continuous value
- (d) discrete time, discrete value



- A lot of applications in engineering systems
- Basics on Random Process
- **Bernoulli Process**
- Poisson Process
- Use of Bernoulli and Poisson Processes
 - Coding of both processes
 - Merge and Split
- Markov Chain

- At each minute, we toss a coin with probability of head $0 < p < 1$.
 - Sequence of lottery wins/loses
 - Customers (each second) to a bank
 - Clicks (at each time slot) to server
- A sequence of **independent Bernoulli trials** X_1, X_2, \dots ,
 - We call index 1, 2, ... **time slots** (or simply slots)

0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | →

- Discrete time, discrete value
- One of the simplest random processes
- A type of “arrival” process

- **Question.** We've already studied a sequence of Bernoulli rvs X_1, X_2, \dots, X_n . What's the difference?
- **Physical difference:** infinite sequence of X_1, X_2, \dots .
 - Sample space? set of all outcomes?
 - an outcome: an infinite sequence of sample values x_1, x_2, \dots , e.g., $(0, 1, 1, 0, 0, 1, \dots)$
- **Semantic difference:** Understand i in X_i as time. Also, interesting questions from the random process point of view.
 - Dependence: How X_1, X_2, \dots are related to each other as a time series
 - Long-term behavior: What is the fraction of times that a machine is idle?
 - Other interesting questions, depending on the target random process
- Next: Key questions and answers about Bernoulli process

(Q1) # of arrivals in n slots?

- $S_n = X_1 + X_2 + \dots + X_n$
- $S_n \sim \text{Bin}(n, p)$
- $\mathbb{E}(S_n) = np, \text{var}(S_n) = np(1-p)$

(Q2) # of slots T_1 until the first arrival?

- $T_1 \sim \text{Geom}(p)$
- $\mathbb{E}(T_1) = 1/p, \text{var}(T_1) = \frac{1-p}{p^2}$

- T_1 is geometric? **Memoryless**
- Conditioned no first arrival at some time, what's the distribution of the remaining time until the first arrival?
- Still, geometric.
- But, more than that, as we will see. Independence across time slots leads to many useful properties, allowing the quick solution of many problems.

(Q3) $U = X_1 + X_2 \perp\!\!\!\perp V = X_5 + X_6$?

- Yes
- Because X_i s are independent

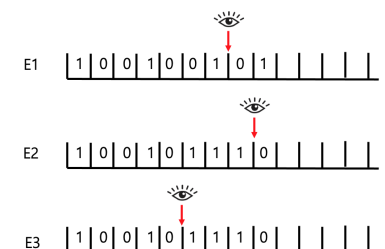
(Q4) The process $(X_n)_{n=6}^\infty$?

- $(X_1, \dots, X_5) \perp\!\!\!\perp (X_n)_{n=6}^\infty$
- **Fresh-start** after a deterministic time n
- If you watch the on-going Bernoulli process(p) from some time n , you still see the same Bernoulli process(p).

(Q5) The process $(X_N, X_{N+1}, X_{N+2}, \dots)$? Fresh-start even after **random** N ?

- Examples of N

- E1. Time of 3rd arrival
- E2. First time when 3 consecutive arrivals have been observed
- E3. Time just before 3 consecutive arrivals



- Difference of N from n
 - The time when I watch the on-going Bernoulli process is **random**.

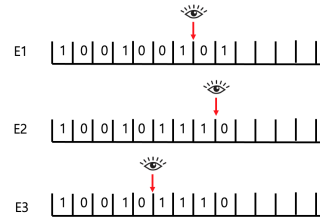
(Q5) The process $(X_N, X_{N+1}, X_{N+2}, \dots)$? Fresh-start even after random N ?

- Examples of N

E1. Time of 3rd arrival

E2. First time when 3 consecutive arrivals have been observed

E3. Time just before 3 consecutive arrivals



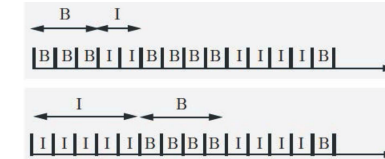
E1. When I watch the process, N has been already determined. **Yes**

E2. Same as E1. **Yes**

E3. Need the future knowledge. '111' does not become random. **No**

- The question of $N = n$? can be answered just from the knowledge about X_1, X_2, \dots, X_n ? Then, Yes! (see pp. 301 for more formal description)

- Regard an arrival as business of a server
- First busy period B_1 : starts with the first busy slot and ends just before the first subsequent idle slot

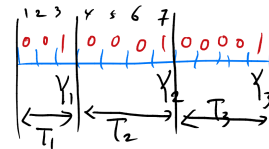


- (Q6) Distribution of B_1 ?
- N : time of the first busy slot. Fresh-start after N .
- B_1 is geometric with parameter $(1 - p)$
- Question: What about the second busy period B_2 ? B_3, B_4 ?

- Time of the first arrival $Y_1 \sim \text{geom}(p)$

(Q7) Time of the k -th arrival Y_k ?

- $T_k = Y_k - Y_{k-1}$: k -th inter-arrival ($k \geq 2, T_1 = Y_1$)
- $Y_k = T_1 + T_2 + \dots + T_k$.



- After each T_k , the fresh-start occurs.

- $\{T_i\}$ are i.i.d. and $\sim \text{geom}(p)$

$$\mathbb{E}[Y_k] = \frac{k}{p}, \text{var}[Y_k] = \frac{k(1-p)}{p^2}$$

$$Y_k = T_1 + T_2 + \dots + T_k.$$

- $\{T_i\}$ are i.i.d. and $\sim \text{geom}(p)$

$$\begin{aligned} \mathbb{P}(Y_k = t) &= \mathbb{P}(X_k = 1 \text{ and } k-1 \text{ arrivals during the first } t-1 \text{ slots}) \\ &= \mathbb{P}(X_k = 1) \cdot \mathbb{P}(k-1 \text{ arrivals during the first } t-1 \text{ slots}) \\ &= p \times \binom{t-1}{k-1} p^{k-1} (1-p)^{t-k} = \binom{t-1}{k-1} p^k (1-p)^{t-k}, \quad t = k, k+1, \dots \end{aligned}$$

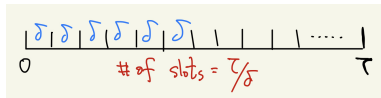
- Y_k is called **Pascal rv** with parameter (k, p) .
- $\text{Pascal}(1, p) = \text{Geometric}(p)$

- A lot of applications in engineering systems
- Basics on Random Process
- Bernoulli Process
- Poisson Process
- Use of Bernoulli and Poisson Processes
 - Coding of both processes
 - Merge and Split
- Markov Chain

- Very useful to both continuous and discrete random processes that are “twins” and share the key properties.
 - Independence between what happens in a different time region
 - Memoryless and fresh-start property
- Choose one of discrete or continuous versions for our modeling convenience.
- **Question.** How do we design the continuous analog of Bernoulli process?
 - Key idea: Making it as a **limiting system** of a sequence of Bernoulli processes
- Need a “modeling sense” to make this possible. It’s a good practice for engineers!

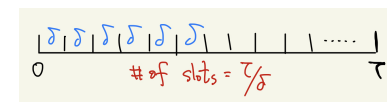
Key Design Idea to Develop a Continuous Twin (1)

- Continuous twin
 - Key point: Understand the number of arrivals over a given interval $[0, \tau]$.
 - Assume that it has some arrival rate λ (# of arrivals/unit time).
 - We know how to handle Bernoulli process with discrete time slots.
- Divide $[0, \tau]$ into slots whose length = δ . Then, $n = \#$ of slots = $\frac{\tau}{\delta}$.



- What's the limit as $\delta \rightarrow 0$ (equivalently, $n \rightarrow \infty$)

Key Design Idea to Develop a Continuous Twin (2)

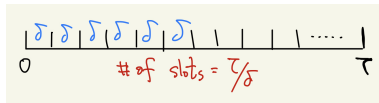


- Now, our design idea: during one time slot of length δ ,

$$\begin{aligned} \mathbb{P}(1 \text{ arrival}) &\propto \text{arrival rate and slot length} \\ \mathbb{P}(\geq 2 \text{ arrivals}) &\propto \text{something, but very small} \\ &\quad \text{for small slot length} \\ \mathbb{P}(0 \text{ arrival}) &= 1 - \mathbb{P}(1 \text{ arrival or } \geq 2 \text{ arrivals}) \end{aligned}$$

$$\begin{aligned} \mathbb{P}(1 \text{ arrival}) &= \lambda\delta + o(\delta) \\ \mathbb{P}(\geq 2 \text{ arrivals}) &= 0o(\delta) \\ \mathbb{P}(0 \text{ arrival}) &= 1 - \lambda\delta + o(\delta) \end{aligned}$$

- $o(\delta)$: some function that goes to zero faster than δ goes to zero.
 - Thus, for very small δ , $o(\delta)$ becomes negligible.
 - Example: $o(\delta) = \delta^\alpha$, where any $\alpha > 1$



- Our interest: Prob. of k arrivals over $[0, \tau]$
- Given "small" δ , # of arrivals $\sim \text{Binomial}(n, p)$, where $n = \tau/\delta$ and $p = \lambda\delta$
- As $\delta \rightarrow \infty$, $np = \tau/\delta \times \lambda\delta = \lambda\tau$.
- # of arrivals over $[0, \tau]$, $\sim \text{Poisson}(\lambda\tau)$
- This is a continuous twin process of Bernoulli process, which we call **Poisson process**.

- N_s : number of arrivals over the interval $[0, s]$.
- **(Independence)** If $s < t$, the number $N_t - N_s$ of arrivals over $[s, t]$ is independent of the times of arrivals during $[0, s]$.
- Thus, N_s can be a random variable over **any** interval of length s .
- **(Small interval probability)** The probabilities $\mathbb{P}(k, s)$ satisfy:

$$\mathbb{P}(0, s) = 1 - \lambda s + o(s)$$

$$\mathbb{P}(1, s) = \lambda s + o_1(s)$$

$$\mathbb{P}(k, s) = o_k(s) \quad \text{for } k = 2, 3, \dots,$$

where

$$\lim_{s \rightarrow 0} \frac{o(s)}{s} = 0, \quad \lim_{s \rightarrow 0} \frac{o_k(s)}{s} = 0$$

- **(Q1)** Number of arrivals of any interval with length $\tau \sim \text{Poisson}(\lambda\tau)$, i.e.,

$$\mathbb{P}(k, \tau) = e^{-\lambda\tau} \frac{(\lambda\tau)^k}{k!}, \quad k = 0, 1, 2, \dots$$

- $\mathbb{E}[N_\tau] = \lambda\tau$ and $\text{var}[N_\tau] = \lambda\tau$

- **(Q2)** Time of first arrival T

$$F_T(t) = \mathbb{P}(T \leq t) = 1 - \mathbb{P}(T > t) = 1 - P(0, t) = 1 - e^{-\lambda t}$$

$$f_T(t) = \frac{dF_T(t)}{dt} = \lambda e^{-\lambda t}, \quad t \geq 0.$$

- $T \sim \text{expo}(\lambda)$. Thus $\mathbb{E}[T] = 1/\lambda$ and $\text{var}[T] = 1/\lambda^2$
- Continuous twin of geometric rv in Bernoulli process
- Memoryless

- Receive emails according to a Poisson process at rate $\lambda = 5$ messages per hour
- Mean and variance of mails received during a day
- $5 * 24 = 120$
- $\mathbb{P}[\text{one new message in the next hour}]$
- $\mathbb{P}(1, 1) = \frac{(5 \cdot 1)^1 e^{-5 \cdot 1}}{1!} = 5e^{-5}$
- $\mathbb{P}[\text{exactly two msgs during each of the next three hours}]$
- $\left(\frac{5^2 e^{-5}}{2!} \right)^3$

- **Remind.** Similar property for Bernoulli processes, but here no time slots.
- **Fresh-start at deterministic time:** Start watching at time t , then you see the Poisson process, independent of what has happened in the past.
- **Fresh-start at random time:** Similarly holds. For example, when you start watching at random time T_1 (time of first arrival)
- **(Q3)** The k -th arrival time Y_k ?
- k -th inter-arrival time $T_k = Y_k - Y_{k-1}$, $k \geq 2$, and $T_1 = Y_1$.
- $Y_k = T_1 + T_2 + \dots + T_k$ is sum of i.i.d. exponential rvs.
- $\mathbb{E}[Y_k] = k/\lambda$ and $\text{var}[Y_k] = k/\lambda^2$

- For a given δ , $\delta \cdot f_{Y_k}(y)$: prob of k -th arrival over $[y, y + \delta]$.
- When δ is small, only one arrival occurs. Thus,

$$\delta \cdot f_{Y_k}(y) = \mathbb{P}(\text{an arrival over } [y, y + \delta]) \times \mathbb{P}(k-1 \text{ arrivals before } y)$$

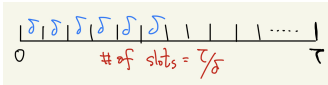
$$\approx \lambda \delta \times \mathbb{P}(k-1, y) = \lambda \delta \times \frac{\lambda^{k-1} y^{k-1} e^{-\lambda y}}{(k-1)!}$$

$$f_{Y_k}(y) = \frac{\lambda^k y^{k-1} e^{-\lambda y}}{(k-1)!}, \quad y \geq 0.$$
- This is called **Erlang** rv.
- Time of first arrival: geometric / exponential
- Time of k -th arrivals: Pascal / Erlang

Poisson Process vs. Bernoulli Process

Example: Poisson Fishing (Problem 10, page 329)

- $n = \tau/\delta$, $p = \lambda\delta$, $np = \lambda\tau$



| | POISSON | BERNOULLI |
|-----------------------|----------------------|----------------|
| Times of Arrival | Continuous | Discrete |
| PMF of # of Arrivals | Poisson | Binomial |
| Interarrival Time CDF | Exponential | Geometric |
| Arrival Rate | λ /unit time | p /per trial |

- Catching fish: Poisson process $\lambda = 0.6/\text{hour}$.

- Fish for 2 hours. Within 2 hours, if he catches at least one fish, then he stops at 2 hours. Otherwise, he fishes until one fish is caught.

(Q1) $\mathbb{P}(\text{fishing time} > 2 \text{ hours})$

Method 1: $\mathbb{P}(0, 2)$

Method 2: $\mathbb{P}(T_1 > 2)$

(Q2) $\mathbb{P}(2 < \text{fishing time} < 5)$

Method 1: $\mathbb{P}(0, 2)(1 - \mathbb{P}(0, 3))$

Method 2: $\mathbb{P}(2 < T_1 < 5)$

(Q3) $\mathbb{P}(\text{Catch at least two fish})$

Method 1: $\sum_{k=2}^{\infty} \mathbb{P}(k) = 1 - \mathbb{P}(0, 2) - \mathbb{P}(1, 2)$

Method 2: $\mathbb{P}(Y_k \leq 2)$

(Q4) $\mathbb{E}[\text{future fi. time} | \text{already fished for 3h}]$

Fresh-start. So,

$\mathbb{E}[\exp(\lambda)] = 1/\lambda = 1/0.6$

(Q5) $\mathbb{E}[F = \text{total fishing time}]$

$$2 + \mathbb{E}[F - 2] = 2 + \mathbb{P}(F = 2) \cdot 0 + \mathbb{P}(F > 2) \cdot \mathbb{E}[F - 2 | F > 2]$$

$$= 2 + \mathbb{P}(0, 2) \cdot \frac{1}{\lambda}$$

(Q6) $\mathbb{E}[\text{number of fish}] = \lambda \cdot 2 + \mathbb{P}(0, 2) \cdot 1$

- A lot of applications in engineering systems
- Basics on Random Process
- Bernoulli Process
- Poisson Process
- Use of Bernoulli and Poisson Processes
 - Coding of both processes
 - Merge and Split
- Markov Chain

- Inter-arrival times facilitates coding of both processes

Alternative Description of the Bernoulli Process

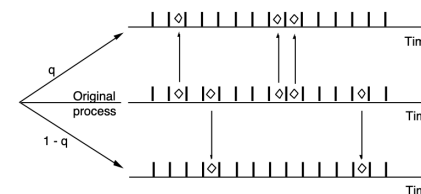
1. Start with a sequence of independent geometric random variables T_1, T_2, \dots , with common parameter p , and let these stand for the inter-arrival times.
2. Record a success (or arrival) at times $T_1, T_1 + T_2, T_1 + T_2 + T_3$, etc.

Alternative Description of the Poisson Process

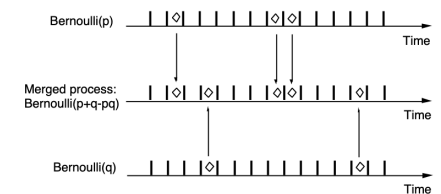
1. Start with a sequence of independent exponential random variables T_1, T_2, \dots , with common parameter λ , and let these stand for the interarrival times.
2. Record an arrival at times $T_1, T_1 + T_2, T_1 + T_2 + T_3$, etc.

- $X \sim \text{Poisson}(\mu), Y \sim \text{Poisson}(\nu)$,
- (Q1) $X \perp\!\!\!\perp Y$?
- (Q2) Distribution of $X + Y$?
 - Complex convolution, but any other easy way?
- X can be regarded as the number of arrivals of Poisson process with rate 1 over the time interval of length μ .
- Consecutive intervals of length μ and ν
- (Q1) $X \perp\!\!\!\perp Y$? Yes
- (Q2) Distribution of $X + Y$? $\text{Poisson}(\mu + \nu)$

- Split $\text{Bernoulli}(p)$ into two processes with biased coin of head probability q
- Split decisions are independent of arrivals
- Split processes: also Bernoulli processes
- $\text{Bernoulli}(pq)$ and $\text{Bernoulli}(p(1 - q))$



- Merge $\text{Bernoulli}(p)$ and $\text{Bernoulli}(q)$ into one.
- Collided arrival is regarded just one arrival in the merged process
- Merged process: $\text{Bernoulli}(1 - (1 - p)(1 - q))$



- Split Poisson process (λ) into two processes
 - Split based on the coin tossing with probability of head p
 - Poisson process $(p\lambda)$ and Poisson process $((1-p)\lambda)$
- Merge from Poisson process (λ_1) and Poisson process (λ_2)
 - Split based on the coin tossing with probability of head p
 - Poisson process $(\lambda_1 + \lambda_2)$
 - Bernoulli process of small interval δ

$$\mathbb{P}(0 \text{ arrivals in the merged process}) \approx (1 - \lambda_1\delta)(1 - \lambda_2\delta) = 1 - (\lambda_1 + \lambda_2)\delta + o(\delta)$$

$$\mathbb{P}(1 \text{ arrivals in the merged process}) \approx \lambda_1\delta(1 - \lambda_2\delta) + \lambda_2\delta(1 - \lambda_1\delta) = (\lambda_1 + \lambda_2)\delta + o(\delta)$$

1. Two independent light bulbs have life times T_a and T_b of exponential distributions with λ_a and λ_b .
 - (Q) Distribution of $Z = \min\{T_a, T_b\}$?
 - T_a and T_b are the first arrival times of two Poisson processes of λ_a and λ_b .
 - Z is the first arrival time of merged Poisson process $(\lambda_a + \lambda_b)$.
 - Thus, $Z \sim \exp(\lambda_a + \lambda_b)$

2. Three independent light bulbs have life times T of exponential distribution with λ .
 - (Q) $\mathbb{E}[\text{time until the last bulb burns out}]$?
 - Poisson process $(3\lambda) \xrightarrow{\text{1st burn out}} \text{Poisson process}(2\lambda) \xrightarrow{\text{2nd burn out}} \text{Poisson process}(\lambda)$
 - T_1 : time until the first burn-out, T_2 : time until the second burn-out, T_3 : time until the third burn-out
 - $T_1 \sim \exp(3\lambda)$, $T_2 \sim \exp(2\lambda)$, $T_3 \sim \exp(\lambda)$

$$\mathbb{E}[T_1 + T_2 + T_3] = \frac{1}{3\lambda} + \frac{1}{2\lambda} + \frac{1}{\lambda}$$

Questions?

1)