# Randomized Algorithms

## Continuous Distributions and the Poisson Process

Joseph Chuang-Chieh Lin
Dept. CSIE, Tamkang University

# Outline

- Continuous Random Variables
- The Uniform Distribution
- The Exponential Distribution
- The Poisson Process

# Recall: Probability function

- $\Pr(\Omega) = 1$
- For any event $E$, $0 \le \Pr(E) \le 1$.
- For any finite or enumerable collection **B** of disjoint events,

$$\Pr\left(\bigcup_{E \in \mathbf{B}} E\right) = \sum_{E \in \mathbf{B}} \Pr[E].$$

# $k$ distinct points in [0, 1)

- $p$: the probability of any given point is in [0, 1).
- $S(k)$: a set of $k$ distinct points in [0, 1).

# $k$ distinct points in [0, 1)

- $p$: the probability of any given point is in [0, 1).
- $S(k)$: a set of $k$ distinct points in [0, 1).

$$\Pr[x \in S(k)] = kp \leq 1.$$

# $k$ distinct points in [0, 1)

- $p$: the probability of any given point is in [0, 1).
- $S(k)$: a set of $k$ distinct points in [0, 1).

$$\Pr[x \in S(k)] = kp \leq 1.$$

- $k$ can be any number!

# *k* distinct points in [0, 1)

- *p*: the probability of any given point is in [0, 1).
- *S*(*k*): a set of *k* distinct points in [0, 1).

$$\Pr[x \in S(k)] = kp \leq 1.$$

- *k* can be any number!

$$\therefore p = 0$$

# Continuous distribution

- Probabilities are assigned to **intervals** rather than to individual values.
- For any $x \in \mathbf{R}$, $\quad F(x) = \Pr[X \leq x] = \Pr[X < x]$.
  - $X$ is continuous if $F(x)$ is a continuous function.

# Continuous distribution

- Probabilities are assigned to **intervals** rather than to individual values.
- For any $x \in \mathbf{R}$, $F(x) = \Pr[X \leq x] = \Pr[X < x]$.
  - $X$ is continuous if $F(x)$ is a continuous function.
- For all $-\infty < a < \infty$,

$$F(a) = \int_{-\infty}^{a} f(t)dt.$$

$$f(x) = F'(x)$$

# Continuous distribution

- Probabilities are assigned to **intervals** rather than to individual values.
- For any $x \in \mathbf{R}$, $F(x) = \Pr[X \le x] = \Pr[X < x]$.
  - $X$ is continuous if $F(x)$ is a continuous function.
- For all $-\infty < a < \infty$,

$$F(a) = \int_{-\infty}^{a} f(t)dt.$$

$$f(x) = F'(x)$$

$$\begin{aligned} \Pr[x < X \le x + dx] &= F(x+dx) - F(x) \\ &\approx f(x)dx. \end{aligned}$$

# Continuous distribution

$$\Pr[a \le X \le b] = \int_a^b f(x)dx.$$

$$\mathbf{E}[X^i] = \int_{-\infty}^{\infty} x^i f(x)dx.$$

$$\mathbf{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)dx.$$

$$\begin{aligned}
\mathbf{Var}[X] &= \mathbf{E}[(X - \mathbf{E}[X])^2] \\
&= \int_{-\infty}^{\infty} (x - \mathbf{E}[x])^2 f(x)dx \\
&= \mathbf{E}[X^2] - (\mathbf{E}[X])^2.
\end{aligned}$$

# Exercise

- **Lemma**. Let $X \geq 0$ be a continuous random variable. Then

$$\mathbf{E}[X] = \int_0^\infty \Pr[X \geq x] dx.$$

# Exercise

$$
\begin{aligned}
\mathbf{E}[X] &= \int_0^\infty y \cdot f(y)dy \\
&= \int_{y=0}^\infty f(y) \int_{x=0}^{y} dx dy \\
&= \int_{x=0}^\infty \int_{y=x}^\infty f(y) dy dx \\
&= \int_{x=0}^\infty (1 - F(x))dx \\
&= \int_{x=0}^\infty \Pr[X \geq x]dx.
\end{aligned}
$$

$$0 \leq x \leq y < \infty$$

# Joint Distribution

- The joint distribution function of *X* and *Y*:

$$F(x, y) = \Pr[X \leq x, Y \leq y].$$

- *X* and *Y* have joint density function *f* if for all *x*, *y*,

$$F(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(u, v) du dv.$$

$$\text{We denote that } \; f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y).$$

# Marginal Distribution Function

- Given a joint distribution function $F(x, y)$ over $X$ and $Y$, we have the marginal distribution functions:

$$F_X(x) = \Pr[X \leq x], \quad F_Y(y) = \Pr[Y \leq y].$$

The corresponding marginal density functions:

$$f_X(x) \text{ and } f_Y(y).$$

# Independence

- The random variables *X* and *Y* are independent if, for all *x* and *y*,

$$\Pr[(X \leq x) \cap (Y \leq y)] = \Pr[X \leq x] \Pr[Y \leq y].$$

$$\Updownarrow$$

$$F(x, y) = F_X(x) F_Y(y).$$

$$\Updownarrow$$

$$f(x, y) = f_X(x) f_Y(y).$$

# Example

- For $a$, $b > 0$, consider the joint distribution function of two random variables $X$ and $Y$:

$$F(x, y) = 1 - e^{-ax} - e^{-by} + e^{-(ax+by)}, \text{ over } x, y \geq 0.$$

# Example

- For $a$, $b > 0$, consider the joint distribution function of two random variables $X$ and $Y$:

$$F(x,y) = 1 - e^{-ax} - e^{-by} + e^{-(ax+by)}, \text{ over } x, y \geq 0.$$

$$F_X(x) = F(x,\infty) = 1 - e^{-ax}, F_Y(y) = F(\infty, y) = 1 - e^{-by}.$$

# Example

- For $a$, $b > 0$, consider the joint distribution function of two random variables $X$ and $Y$:

$$F(x, y) = 1 - e^{-ax} - e^{-by} + e^{-(ax+by)}, \text{ over } x, y \geq 0.$$

$$F_X(x) = F(x, \infty) = 1 - e^{-ax}, F_Y(y) = F(\infty, y) = 1 - e^{-by}.$$

$$f(x, y) = abe^{-(ax+by)}.$$

# Example

- For $a$, $b > 0$, consider the joint distribution function of two random variables $X$ and $Y$:

$$F(x, y) = 1 - e^{-ax} - e^{-by} + e^{-(ax+by)}, \text{ over } x, y \geq 0.$$

$$F_X(x) = F(x, \infty) = 1 - e^{-ax}, F_Y(y) = F(\infty, y) = 1 - e^{-by}.$$

$$f(x, y) = abe^{-(ax+by)}.$$

$$F(x, y) = (1 - e^{-ax})(1 - e^{-by}) = F_X(x)F_Y(y).$$

$$f_X(x) = ae^{-ax}, f_Y(y) = be^{-by}, \therefore f(x, y) = f_X(x)f_Y(y).$$

# Conditional Probability

$$\Pr[X \leq x \mid Y = y] = \lim_{\delta \to 0} \Pr[X \leq x \mid y \leq Y \leq y + \delta].$$

Why?

# Conditional Probability

$$
\begin{aligned}
\Pr[X \le x \mid Y = y] &= \lim_{\delta \to 0} \Pr[X \le x \mid y \le Y \le y + \delta] \\
&= \lim_{\delta \to 0} \frac{\Pr[(X \le x) \cap (y \le Y \le y + \delta)]}{\Pr[y \le Y \le y + \delta]} \\
&= \lim_{\delta \to 0} \frac{F(x, y + \delta) - F(x, y)}{F_Y(y + \delta) - F_Y(y)} \\
&= \lim_{\delta \to 0} \int_{u=-\infty}^{x} \frac{\partial F(u, y + \delta)/\partial x - \partial F(u, y)/\partial x}{F_Y(y + \delta) - F_Y(y)} du \\
&= \int_{u=-\infty}^{x} \lim_{\delta \to 0} \frac{(\partial F(u, y + \delta)/\partial x - \partial F(u, y)/\partial x)/\delta}{(F_Y(y + \delta) - F_Y(y))/\delta} du \\
&= \int_{u=-\infty}^{x} \frac{f(u, y)}{f_Y(y)} du.
\end{aligned}
$$

# Conditional Probability

- For example,

$$F(x,y) = 1 - e^{-ax} - e^{-by} + e^{-(ax+by)},$$

$$\Pr[X \leq 3 \mid Y = 4]$$
$$= \int_{u=0}^{3} \frac{abe^{-au+4b}}{be^{-4b}} du = 1 - e^{-3a}.$$

For $a$, $b > 0$, consider the joint distribution function of two random variables $X$ and $Y$:

$$F(x,y) = 1 - e^{-ax} - e^{-by} + e^{-(ax+by)}, \text{ over } x, y \geq 0.$$

$$F_X(x) = F(x,\infty) = 1 - e^{-ax}, F_Y(y) = F(\infty, y) = 1 - e^{-by}.$$

$$f(x,y) = abe^{-(ax+by)}.$$

$$F(x,y) = (1 - e^{-ax})(1 - e^{-by}) = F_X(x)F_Y(y).$$

$$f_X(x) = ae^{-ax}, f_Y(y) = be^{-by}, \therefore f(x,y) = f_X(x)f_Y(y).$$

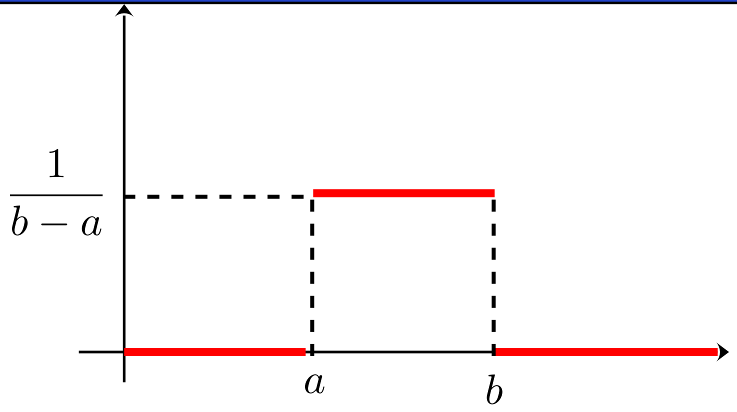
# Conditional Density Function

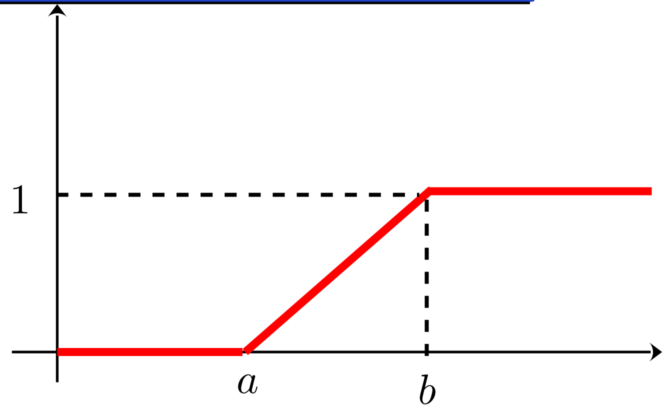
- Assume that $f_Y(y) \neq 0$ (resp., $f_X(x) \neq 0$),

$$\begin{aligned} f_{X|Y}(x,y) &= \frac{f(x,y)}{f_Y(y)} \\ f_{Y|X}(x,y) &= \frac{f(x,y)}{f_X(x)}. \end{aligned}$$

$$\mathbf{E}[X \mid Y = y] = \int_{x=-\infty}^{\infty} x \cdot f_{X|Y}(x,y)dx.$$

# Uniform Distribution




$$f(x) = \begin{cases} 0 & \text{if } x \le a, \\ \frac{1}{b-a} & \text{if } a \le x \le b, \\ 0 & \text{if } x \ge b. \end{cases}$$

$$F(x) = \begin{cases} 0 & \text{if } x \le a, \\ \frac{x-a}{b-a} & \text{if } a \le x \le b, \\ 1 & \text{if } x \ge b. \end{cases}$$

$$\mathbf{E}[X] = \int_a^b \frac{x}{b-a} dx = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2}.$$

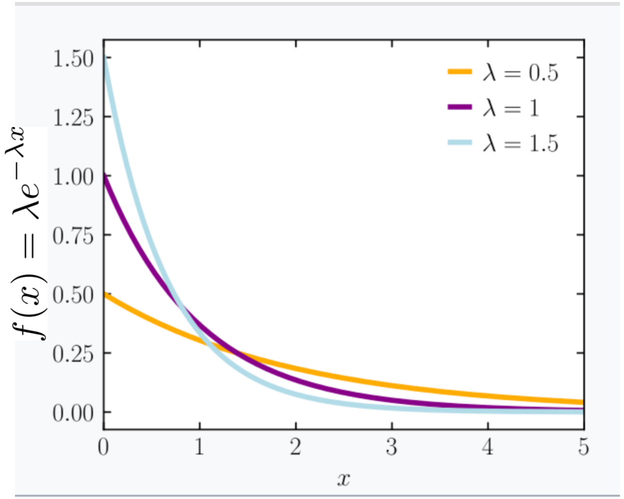
$$\mathbf{E}[X^2] = \int_a^b \frac{x^2}{b-a} dx = \frac{b^3 - a^3}{3(b-a)} = \frac{b^2 + ab + a^2}{3}.$$

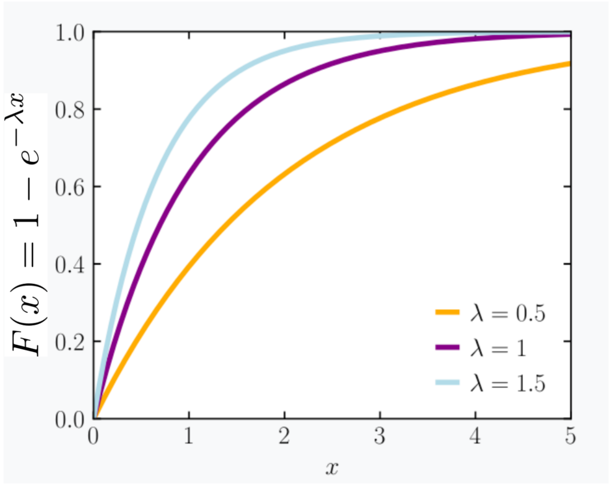
$$\mathbf{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \frac{(b-a)^2}{12}.$$

# Exponential Distribution

- **Definition**. An exponential distribution with parameter $\lambda$ is given by the following probability distribution function:

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{for } x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$




wikipedia

# Exponential Distribution

**Note:** $\Pr[X > t] = 1 - F(t) = e^{-\lambda t}$.

$$\mathbf{E}[X] = \int_0^\infty t\lambda e^{-\lambda t} dt = \frac{1}{\lambda}.$$ (Integration by parts)

$$\mathbf{E}[X^2] = \int_0^\infty t^2 \lambda e^{-\lambda t} dt = \frac{2}{\lambda^2}.$$

$$\mathbf{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \frac{1}{\lambda^2}.$$

# Exponential Distribution (memoryless)

- For an exponential random variable $X$ with parameter $\lambda$,

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

# Exponential Distribution (memoryless)

- For an exponential random variable $X$ with parameter $\lambda$,

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

$$\begin{aligned}
\Pr[X > s + t \mid X > t] &= \frac{\Pr[X > s + t]}{\Pr[X > t]} \\
&= \frac{1 - \Pr[X \leq s + t]}{1 - \Pr[X \leq t]} \\
&= \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} \\
&= e^{-\lambda s} = \Pr[X > s].
\end{aligned}$$

# Min (exponential random variables)

- **<u>Lemma</u>**. If $X_1$, $X_2$, …, $X_n$ are independent exponential random variables with parameters $\lambda_1$, $\lambda_2$, …, $\lambda_n$, respectively, then **min($X_1$, $X_2$, …, $X_n$)** is exponential random variable with parameter $\sum_{i=1}^{n} \lambda_i$

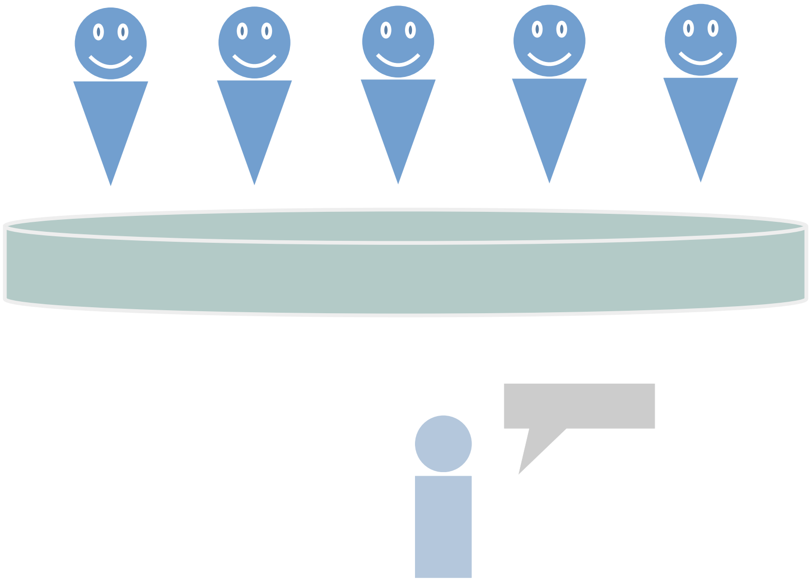
# Min (exponential random variables)

- **<u>Lemma</u>**. If $X_1, X_2, \ldots, X_n$ are independent exponential random variables with parameters $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively, then **min($X_1, X_2, \ldots, X_n$)** is exponential random variable with parameter $\sum_{i=1}^{n} \lambda_i$

$$
\begin{aligned}
\Pr[\min(X_1, X_2, \ldots, X_n)] &= \Pr[(X_1 > x) \cap (X_2 > x) \cap \cdots \cap (X_n > x)] \\
&= \Pr[X_1 > x] \cdot \Pr[X_2 > x] \cdots \Pr[X_n > x] \\
&= e^{-\lambda_1 x} \cdot e^{-\lambda_2 x} \ldots e^{-\lambda_n x} \\
&= e^{-\sum_{i=1}^{n} \lambda_i}.
\end{aligned}
$$

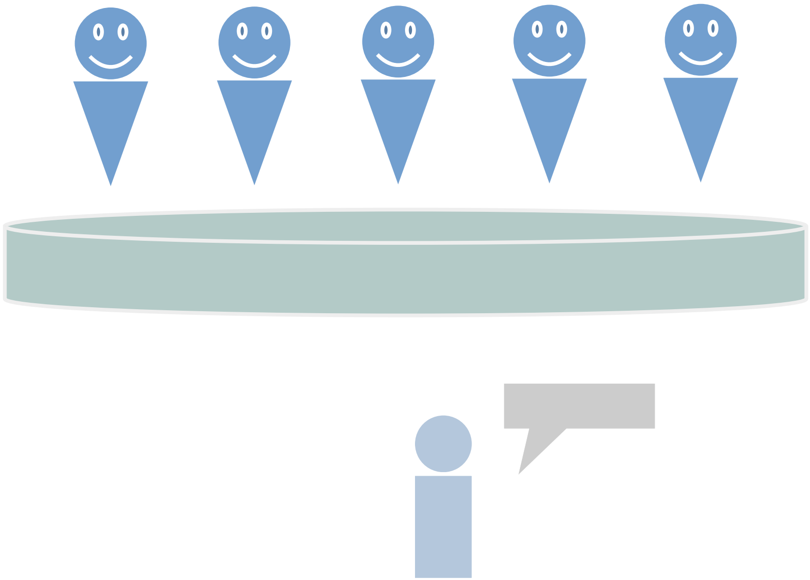
# Scenario

- An airline ticket counter with *n* service agents.
- The time agent *i* takes per customer:
  - Exponential distribution, parameter $\lambda_i$
- You are at the head of the line and wondering how long, in average, you wait for an agent to serve you...



# Scenario

- An airline ticket counter with *n* service agents.
- The time agent *i* takes per customer:
  - Exponential distribution, parameter $\lambda_i$
- You are at the head of the line and wondering how long, in average, you wait for an agent to serve you...



The time until the first agent is free

- Exponential distribution with parameter $\sum_{i=1}^{n} \lambda_i$
- Expected waiting time: $1/\sum_{i=1}^{n} \lambda_i$

# The Poisson Process

Siméon Poisson
(1781−1840)
Wikipedia

- Counting process.
  - E.g., arrivals of customers to a queue, emissions of radioactive particles, price surges in the stock markets, etc.
- $N(t)$: the number of events in the interval (say $[0, t]$).
- A **stochastic counting process**: $\{N(t):\ t \geq 0\}$

# The Poisson Process

Siméon Poisson
(1781−1840)
Wikipedia

- A **Poisson process** $\{N(t)\colon\ t \geq 0\}$ with parameter $\lambda$ is a stochastic counting process such that the following conditions hold.

  1. $N(0) = 0$.
  2. (Independent & stationary increments) For any $t$, $s > 0$,
     - the distribution of $N(t+s) - N(s)$ is identical to the distribution of $N(t)$;
     - for disjoint intervals $[t_1, t_2]$ and $[t_3, t_4]$, the distribution of $N(t_2) - N(t_1)$ is independent of the distribution of $N(t_4) - N(t_3)$.
  3. $\displaystyle \lim_{t \to 0} \frac{\Pr[N(t) = 1]}{t} = \lambda.$
  4. $\displaystyle \lim_{t \to 0} \frac{\Pr[N(t) \geq 2]}{t} = 0.$

# The Poisson Process

Siméon Poisson
(1781−1840)
Wikipedia

- A **Poisson process** $\{N(t)\colon\ t \geq 0\}$ with parameter $\lambda$ is a stochastic counting process such that the following conditions hold.

1. $N(0) = 0$.
2. (Independent & stationary increments) For any $t, s > 0$,
   - the distribution of $N(t+s) - N(s)$ is identical to the distribution of $N(t)$;
   - for disjoint intervals $[t_1, t_2]$ and $[t_3, t_4]$, the distribution of $N(t_2) - N(t_1)$ is independent of the distribution of $N(t_4) - N(t_3)$.
3. $\displaystyle\lim_{t\to 0} \frac{\Pr[N(t)=1]}{t} = \lambda.$
4. $\displaystyle\lim_{t\to 0} \frac{\Pr[N(t)\geq 2]}{t} = 0.$

The number of events in a given time interval follows the **Poisson distribution**!

# Poisson Process → Poisson Distribution

- **Theorem**. Let $\{N(t)\colon\ t \geq 0\}$ be a Poisson process with with parameter $\lambda$. For any $t$, $s \geq 0$ and any integer $n \geq 0$,

$$P_n(t) := \Pr[N(t+s) - N(s) = n] = e^{-\lambda t}\frac{(\lambda t)^n}{n!}.$$

# Poisson Process → Poisson Distribution

- **Theorem**. Let $\{N(t)\colon\ t \geq 0\}$ be a Poisson process with with parameter $\lambda$. For any $t$, $s \geq 0$ and any integer $n \geq 0$,

$$P_n(t) := \boxed{\Pr[N(t+s) - N(s) = n]} = e^{-\lambda t}\frac{(\lambda t)^n}{n!}.$$

$$= \Pr[N(t) - N(0)] = \Pr[N(t) - 0]$$

  - The probability that $n$ events happen during time interval of length $t$.

# Stochastic Process + Poisson = ?

- **<u>Theorem</u>**. Let $\{N(t):\ t \geq 0\}$ be a **stochastic process** such that
  1. $N(0) = 0$.
  2. (Independent increments) For disjoint intervals $[t_1, t_2]$ and $[t_3, t_4]$, the distribution of $N(t_2) - N(t_1)$ is independent of the distribution of $N(t_4) - N(t_3)$.
  3. For any $t, s \geq 0$, $N(t+s) - N(s)$ has a Poisson distribution with mean $\lambda t$.

  Then $\{N(t):\ t \geq 0\}$ is a Poisson process with rate $\lambda$.

# Stochastic Process + Poisson = ?

- **<u>Theorem</u>**. Let $\{N(t): t \geq 0\}$ be a **stochastic process** such that
  1. $N(0) = 0$.
  2. (Independent increments) For disjoint intervals $[t_1, t_2]$ and $[t_3, t_4]$, the distribution of $N(t_2) - N(t_1)$ is independent of the distribution of $N(t_4) - N(t_3)$.
  3. For any $t, s \geq 0$, $N(\boldsymbol{t}+s) - N(s)$ has a Poisson distribution with mean $\lambda \boldsymbol{t}$.

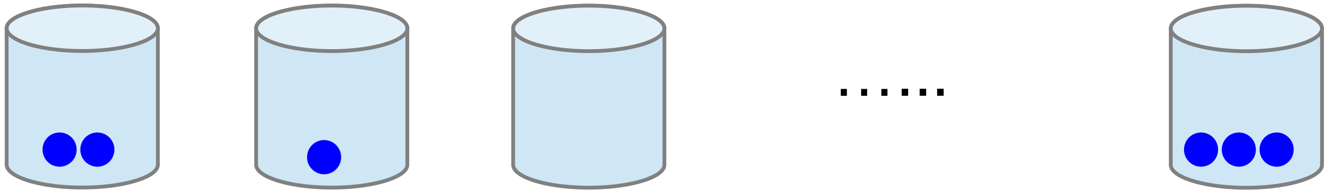
  Then $\{N(t): t \geq 0\}$ is a Poisson process with rate $\lambda$.

$$\lim_{t\to 0} \frac{\Pr[N(t) = 1]}{t} = \lim_{t\to 0} \frac{e^{-\lambda t}\lambda t}{t} = \lambda.$$

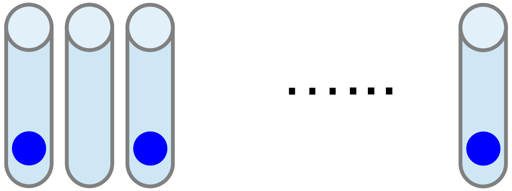
$$\lim_{t\to 0} \frac{\Pr[N(t) \geq 2]}{t} = \lim_{t\to 0} \sum_{k\geq 2} \frac{e^{-\lambda t}(\lambda t)^k}{k!t} = 0.$$

# An Intuitive Idea

- Balls: events, bins: time slots

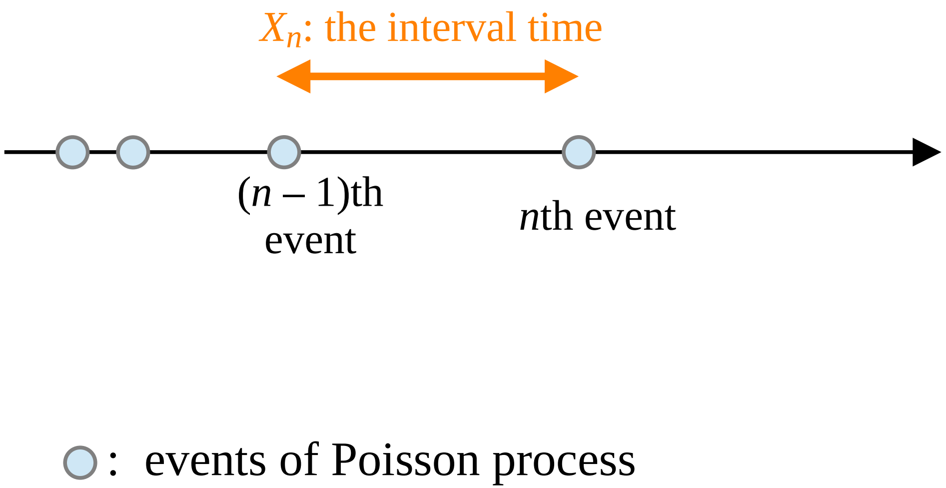


- A lot of balls into a lot of (infinitely small) bins...



# Another Viewpoint: Interarrival

- Surprising fact: All of the $X_n$ have the same distribution and this distribution is **exponential**!



# Interarrival times

- **Theorem**. $X_1$ has an exponential distribution with parameter $\lambda$.

$$\Pr[X_1 > t] = \Pr[N(t) = 0] = e^{\lambda t}.$$

$$F(X_1) = 1 - \Pr[X_1 > t] = 1 - e^{\lambda t}.$$

# Interarrival times

- **Theorem**. $X_i$ , i = 1, 2, …, are i.i.d exponential random variables with parameter parameter $\lambda$.

$$
\begin{aligned}
& \Pr[X_i > t_i \mid (X_0 = t_0) \cap (X_1 = t_1) \cap \cdots \cap (X_{i-1} = t_{i-1})] \\
= \;& \Pr\left[N\left(\sum_{k=0}^{i} t_k\right) - N\left(\sum_{k=0}^{i-1} t_k\right) = 0\right] \\
= \;& e^{-\lambda t_i}.
\end{aligned}
$$

# Interarrival times

- **Theorem**. $X_i$ , i = 1, 2, …, are i.i.d exponential random variables with parameter parameter $\lambda$.

$$
\begin{aligned}
& \Pr[X_i > t_i \mid (X_0 = t_0) \cap (X_1 = t_1) \cap \cdots \cap (X_{i-1} = t_{i-1})] \\
= \quad & \Pr\left[ N\left(\sum_{k=0}^{i} t_k\right) - N\left(\sum_{k=0}^{i-1} t_k\right) = 0 \right] \\
= \quad & e^{-\lambda t_i}.
\end{aligned}
$$

$\because$ Poisson process & $\sim$ distribution of $N(t_i) - N(0) = N(t_i)$

# So, why the Poisson process makes $P_n(t)$ Poisson distributed?

## The Poisson Process

Siméon Poisson
(1781–1840)
Wikipedia

- A **Poisson process** $\{N(t):\ t \geq 0\}$ with parameter $\lambda$ is a stochastic counting process such that the following conditions hold.
  1. $N(0) = 0$.
  2. (Independent & stationary increments) For any $t, s > 0$,
     - the distribution of $N(t+s) - N(s)$ is identical to the distribution of $N(t)$;
     - for disjoint intervals $[t_1, t_2]$ and $[t_3, t_4]$, the distribution of $N(t_2) - N(t_1)$ is independent of the distribution of $N(t_4) - N(t_3)$.
  3. $\lim_{t\to 0} \frac{\Pr[N(t)=1]}{t} = \lambda.$
  4. $\lim_{t\to 0} \frac{\Pr[N(t)\geq 2]}{t} = 0.$

The number of events in a given time interval follows the **Poisson distribution**!

Randomized Algorithms, CSIE, TKU, Taiwan

36

## Poisson Process → Poisson Distribution

- **Theorem**. Let $\{N(t):\ t \geq 0\}$ be a Poisson process with with parameter $\lambda$. For any $t, s \geq 0$ and any integer $n \geq 0$,

$$P_n(t) := \Pr[N(t+s) - N(s) = n] = e^{-\lambda t}\frac{(\lambda t)^n}{n!}.$$

# $P_0(t)$

- The number of events in $[0, t]$ and $(t, t+h]$ are independent.

$$P_0(t+h) = P_0(t) \cdot P_0(h).$$

# $P_0(t)$

- The number of events in [0, $t$] and ($t$, $t$+$h$] are independent.

$$P_0(t+h) = P_0(t) \cdot P_0(h).$$

$$\begin{aligned}
& \therefore \frac{P_0(t+h) - P_0(t)}{h} \\
= & \quad \frac{P_0(t)P_0(h) - P_0(t)}{h} \\
= & \quad P_0(t)\frac{P_0(h) - 1}{h} \\
= & \quad P_0(t)\frac{1 - \Pr[N(h) = 1] - \Pr[N(h) \geq 2] - 1}{h} \\
= & \quad P_0(t)\frac{-\Pr[N(h) = 1] - \Pr[N(h) \geq 2]}{h}.
\end{aligned}$$

# $P_0(t)$

- The number of events in [0, $t$] and ($t$, $t$+$h$] are independent.

$$P_0(t+h) = P_0(t) \cdot P_0(h).$$

$$\begin{aligned} \therefore \frac{P_0(t+h) - P_0(t)}{h} & \\ = & \frac{P_0(t)P_0(h) - P_0(t)}{h} \\ = & P_0(t)\frac{P_0(h) - 1}{h} \\ = & P_0(t)\frac{1 - \Pr[N(h)=1] - \Pr[N(h) \geq 2] - 1}{h} \\ = & P_0(t)\frac{-\Pr[N(h)=1] - \Pr[N(h) \geq 2]}{h}. \end{aligned}$$

$$\begin{aligned} \therefore P_0'(t) &= \lim_{h \to 0} \frac{P_0(t+h) - P_0(t)}{h} \\ &= \lim_{h \to 0} P_0(t) \frac{-\Pr[N(h)=1] - \Pr[N(h) \geq 2]}{h} \\ &= -\lambda P_0(t). \end{aligned}$$

Solve $P_0'(t) = -\lambda P_0(t)$

# $P_0(t)$

- The number of events in [0, $t$] and ($t$, $t$+$h$] are independent.

$$P_0(t+h) = P_0(t) \cdot P_0(h).$$

$$\begin{aligned}
&\therefore \frac{P_0(t+h) - P_0(t)}{h} \\
= \quad & \frac{P_0(t)P_0(h) - P_0(t)}{h} \\
= \quad & P_0(t)\frac{P_0(h) - 1}{h} \\
= \quad & P_0(t)\frac{1 - \Pr[N(h)=1] - \Pr[N(h)\geq 2] - 1}{h} \\
= \quad & P_0(t)\frac{-\Pr[N(h)=1] - \Pr[N(h)\geq 2]}{h}.
\end{aligned}$$

$$\begin{aligned}
\therefore P_0'(t) \quad &= \quad \lim_{h\to 0}\frac{P_0(t+h) - P_0(t)}{h} \\
&= \quad \lim_{h\to 0} P_0(t)\frac{-\Pr[N(h)=1] - \Pr[N(h)\geq 2]}{h} \\
&= \quad -\lambda P_0(t).
\end{aligned}$$

Solve $P_0'(t) = -\lambda P_0(t)$

$$\frac{P_0'(t)}{P_0(t)} = -\lambda. \Rightarrow \int \frac{P_0'(t)}{P_0(t)}dt = \int -\lambda dt$$

$$\ln P_0(t) = -\lambda t + C \quad \therefore P_0(t) = e^{-\lambda t} \; (\because P_0(0) = 1)$$

# $P_n(t),\ n \geq 1$

$$
\begin{aligned}
P_n(t+h) &= \sum_{k=0}^{n} P_{n-k}(t) P_k(h) \\
&= P_n(t)P_0(h) + P_{n-1}(t)P_1(h) + \sum_{k=2}^{n} P_{n-k}(t) \cdot \Pr[N(h) = k]
\end{aligned}
$$

$$
\begin{aligned}
P_n'(t) &= \lim_{h \to 0} \frac{P_n(t+h) - P_n(t)}{h} \\
&= \lim_{h \to 0} \frac{1}{h} \left( P_n(t)(P_0(h) - 1) + P_{n-1}(t) P_1(h) + \sum_{k=2}^{n} P_{n-k}(t) \Pr[N(h) = k] \right) \\
&= -\lambda P_n(t) + \lambda P_{n-1}(t).
\end{aligned}
$$

# $P_n(t), n \geq 1$

$$\begin{aligned} P_n(t+h) &= \sum_{k=0}^{n} P_{n-k}(t)P_k(h) \\ &= P_n(t)P_0(h) + P_{n-1}(t)P_1(h) + \sum_{k=2}^{n} P_{n-k}(t)\cdot \Pr[N(h)=k] \end{aligned}$$

$$\begin{aligned} P_n'(t) &= \lim_{h\to 0} \frac{P_n(t+h)-P_n(t)}{h} \\ &= \lim_{h\to 0} \frac{1}{h}\left( P_n(t)(P_0(h)-1) + P_{n-1}(t)P_1(h) + \sum_{k=2}^{n} P_{n-k}(t)\Pr[N(h)=k] \right) \\ &= -\lambda P_n(t) + \lambda P_{n-1}(t). \end{aligned}$$

Solve $P_n'(t) = -\lambda P_n(t) + \lambda P_{n-1}(t). \;\Rightarrow\; e^{\lambda t}(P_n'(t) + \lambda P_n(t)) = e^{\lambda t}(\lambda P_{n-1}(t).)$

# $P_n(t)$, $n \geq 1$

Solve $P'_n(t) = -\lambda P_n(t) + \lambda P_{n-1}(t)$. $\quad \Rightarrow \quad e^{\lambda t}(P'_n(t) + \lambda P_n(t)) = e^{\lambda t}(\lambda P_{n-1}(t))$.

$$\Rightarrow \quad \frac{d}{dt}(e^{\lambda t} P_n(t)) = \lambda e^{\lambda t} P_{n-1}(t).$$

# $P_n(t), n \geq 1$

Solve $P_n'(t) = -\lambda P_n(t) + \lambda P_{n-1}(t).$ $\Rightarrow$ $e^{\lambda t}(P_n'(t) + \lambda P_n(t)) = e^{\lambda t}(\lambda P_{n-1}(t)).$

$$\Rightarrow \frac{d}{dt}(e^{\lambda t} P_n(t)) = \lambda e^{\lambda t} P_{n-1}(t).$$

$$\therefore \frac{d}{dt}(e^{\lambda t} P_1(t)) = \lambda e^{\lambda t} P_0(t) = \lambda. \qquad (P_0(t) = e^{-\lambda t})$$

$$\Rightarrow \int d(e^{\lambda t} P_1(t)) = \int \lambda dt.$$

$$\Rightarrow e^{\lambda t} P_1(t) = \lambda t + C$$

$$\Rightarrow P_1(t) = e^{-\lambda t}(\lambda t + C) = t\lambda e^{-\lambda t}.$$

Note: $P_1(0) = 0$
1 event occurs in zero time period

# $P_n(t)$, $n \geq 1$

- Look back at what we want to have:

$$P_n(t) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$$

- We already have $P_0(t)$ and $P_1(t)$.

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

$$(P_1(t) = t\lambda e^{-\lambda t})$$

- Induction hypothesis:

$$P_{n-1}(t) = e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}$$

# $P_n(t)$, $n \geq 1$

- Look back at what we want to have:

$$P_n(t) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$$

- We already have $P_0(t)$ and $P_1(t)$.
- Induction hypothesis:

$$P_{n-1}(t) = e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}$$

Use $\dfrac{d}{dt}(e^{\lambda t}P_n(t)) = \lambda e^{\lambda t}P_{n-1}(t)$.

# $P_n(t), n \geq 1$

- Look back at what we want to have:

$$P_n(t) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$$

- We already have $P_0(t)$ and $P_1(t)$.
- Induction hypothesis:

$$P_{n-1}(t) = e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}$$

Use $\dfrac{d}{dt}(e^{\lambda t}P_n(t)) = \lambda e^{\lambda t}P_{n-1}(t) = \dfrac{\lambda^n t^{n-1}}{(n-1)!}.$

# $P_n(t), n \geq 1$

- Look back at what we want to have:

$$P_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$$

- We already have $P_0(t)$ and $P_1(t)$.
- Induction hypothesis:

$$P_{n-1}(t) = e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}$$

Use $\quad \dfrac{d}{dt}(e^{\lambda t} P_n(t)) = \lambda e^{\lambda t} P_{n-1}(t) = \dfrac{\lambda^n t^{n-1}}{(n-1)!}.$

$$\Rightarrow \quad \int d(e^{\lambda t} P_n(t)) = \int \frac{\lambda^n t^{n-1}}{(n-1)!} dt$$

$$\Rightarrow \quad e^{\lambda t} P_n(t) = \frac{\lambda^n t^{n-1}}{(n-1)!} \cdot \frac{t^n}{n} + C$$

$$\Rightarrow \quad P_n(t) = e^{-\lambda t} \frac{\lambda^n t^n}{n!}. \quad \text{(use } P_n(0) = 0\text{)}$$

# Further related topics

- Stochastic counting process: **point processes**
- Hawkes (self-exciting) processes.
  - Earthquake modeling, financial analysis.
  - A reference: https://arxiv.org/pdf/1507.02822.pdf