

Exponential Distribution and Poisson Process

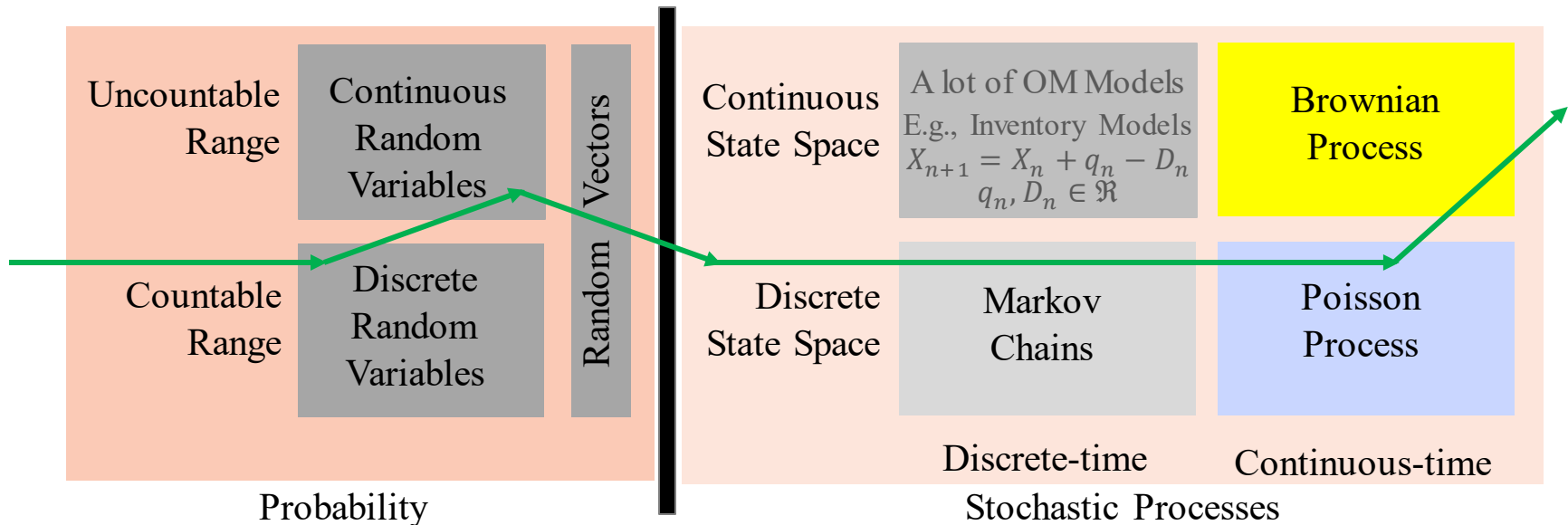
Outline

- ◆ **Continuous-time Markov Process**
- ◆ **Poisson Process**
- ◆ **Thinning**
- ◆ **Conditioning on the Number of Events**
- ◆ **Generalizations**

Probability and Stochastic Processes

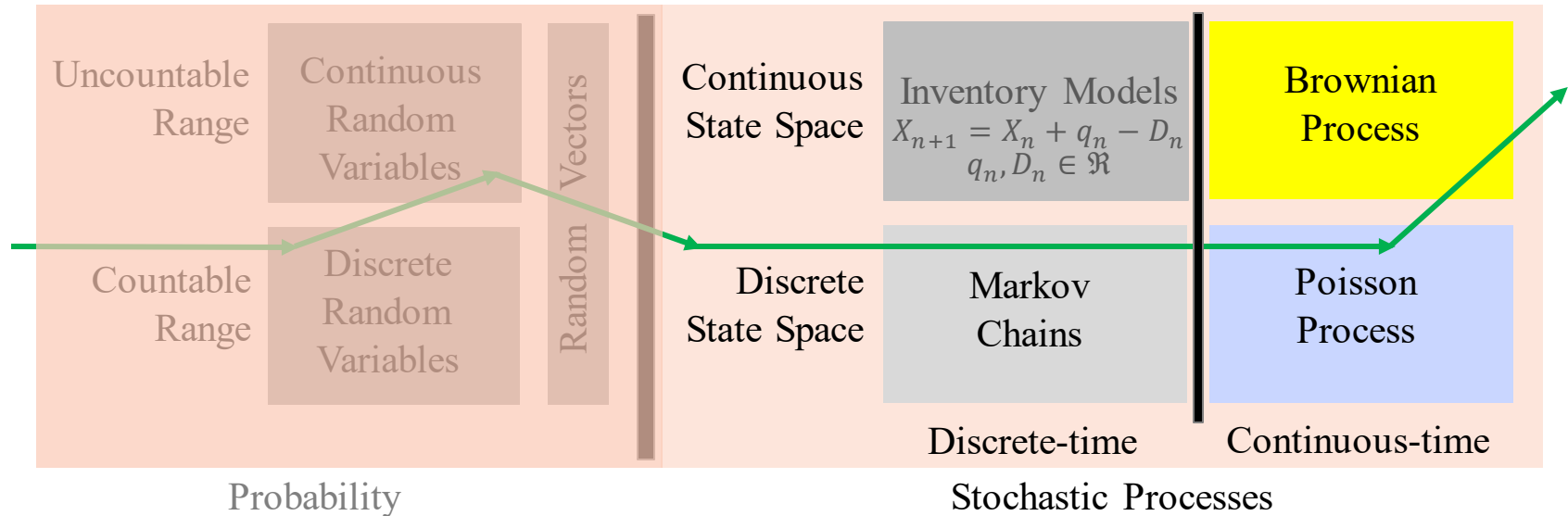
◆ Overview of Random Variables and Stochastic Processes

- Discrete-time: $\{X_n: n \geq 0, n \in \mathbb{Z}\}$; Continuous-time: $\{X(t): t \geq 0, t \in \mathbb{R}\}$
- Discrete state-space if each X_n or $X(t)$ has a countable range; Otherwise, Continuous state-space.



- Difference between an infinite-length random vector & a discrete-time stochastic process?
 - » Similarity, they are both denoted as $\{X_n: n \geq 0, n \in \mathbb{Z}\}$.
 - » Difference is in the treatment of **dependence**:
 - ◆ Random vectors
 - Only short lengths (2-3 rvs) can be studied if they are **dependent**;
 - Or assume **structured dependence** through copulas.
 - Or work with **independent** and identically distributed sequences.
 - ◆ In discrete-time stochastic processes, assume Markov Property to **limit the dependence**

Stochastic Processes in Continuous Time



- ◆ Difference between a discrete-time stochastic process & continuous-time stochastic process?
 - Similarity, **limited dependence** is still sought.
 - Difference is in the **continuity** of the process in time:
 - » Continuity is not an issue for processes with a discrete state space
 - » When the state space is continuous, we wonder how $X(t)$ & $X(t + \epsilon)$ differ.
 - » For random variables, almost sure continuity: $P\left(\lim_{\epsilon \rightarrow 0} [\omega \in \Omega: X(t) = X(t - \epsilon)]\right) = 1$.
- ◆ Ex: The exchange rate between Euro and US dollar is $EURUSD$. The stochastic process $EURUSD(t)$ can be generally continuous except when FED or ECB act to move the exchange rate in a certain direction.
 - Immediately after some bank board meetings, $EURUSD(t)$ can jump. Central bank interventions can cause jumps.

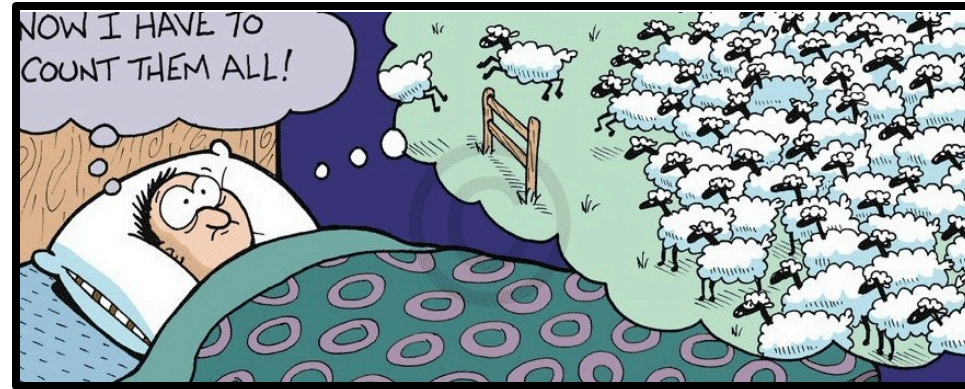
Seeking Convenience

- ◆ Convention: A jump that happens at time t is included in $X(t)$ but not in $X(t - \epsilon)$
- ◆ Two common properties
 - **Independent increments** $X(t_1) - X(s_1) \perp X(t_2) - X(s_2)$ for disjoint intervals $(s_1, t_1]$ and $(s_2, t_2]$.
 - **Stationary increments** $P(X(s + a) - X(s) = x) = P(X(t + a) - X(t) = x)$.
- ◆ Ex: Independent increments imply $P(X(t) = x_t | X(r) = x_r \text{ for } r \in (0, s]) = P(X(t) = x_t | X(s) = x_s)$, which is the Markov property.
- ◆ Continuity property
 - **Right-continuous evolution** $P\left(\lim_{t \downarrow 0} X(t + \epsilon) = X_t\right) = 1$, i.e., X_t is right-continuous almost surely.
- ◆ Ex: Let $X(t, \omega) = U_t(\omega)$ and $X(t + \epsilon, \omega) = U_{t+\epsilon}(\omega)$ for $U_t, U_{t+\epsilon}$ iid from $U(0,1)$.
 - Check for convergence in probability. $P(|X(t + \epsilon) - X_t| \geq 0.5) = \text{Area of two triangles in unit square} = \frac{1}{4}$
 - $X(t + \epsilon)$ does not converge to X_t in probability. So it does not converge almost surely.
 - $X(t)$ is not right continuous. It is not left-continuous either.
 - Can $X(t)$ be used to model a stock price?
- ◆ A process with **Independent & Stationary increments**, **Right-continuous evolution** is a Lévy process.
- ◆ Lévy-Itô Decomposition:
$$\text{Lévy Process} = \text{Poisson Process} + \text{Brownian Process} + \text{Martingale Process} + \text{Deterministic drift Linear in Time}$$
- ◆ A Martingale satisfies $E(X(t) | X(s) = x_s) = x_s$ for $s \leq t$.

Counting Process

- ◆ A counting process $\{N(t): t \geq 0\}$ is a continuous-time stochastic process with natural numbers as its state-space and $N(t)$ is the number of events that occur over interval $(0, t]$.

- Start with $N(0)=0$
- $N(t)$ is non-decreasing and right-continuous.



- ◆ Assume independent and stationary increments:

- The number of events happening over disjoint intervals are independent: $N(t_1) - N(s_1) \perp N(t_2) - N(s_2)$ or disjoint intervals $(s_1, t_1]$ and $(s_2, t_2]$.
- The number of events over two intervals of the same length have the same distribution: $P(N(s+a) - N(s) = k) = P(N(t+a) - N(t) = k)$.

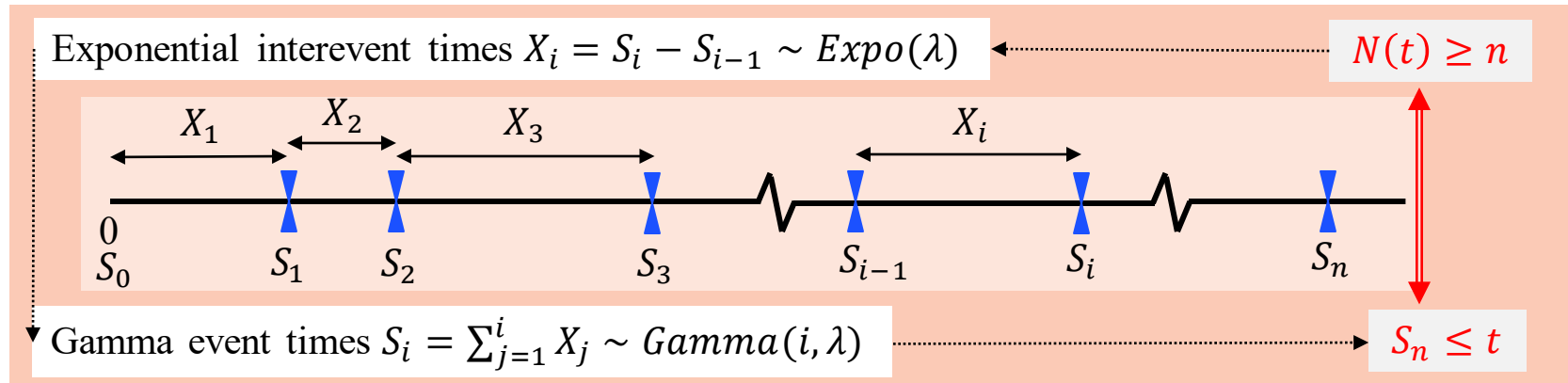
- ◆ Ex: For a car rental agency, let n_i be the number of scratches on returned car i . Then the associated counting process, which yields the total number of scratches in the first m returned cars, is

$$N(m) = \sum_{i=1}^m n_i.$$

- $N(m)$ is indexed by integer m but we can still ask if it has independent and stationary increments.
- Independence: Scratches generally occur when the cars are rented out to different drivers and driven on different roads at different times
- Stationary: Different cars have rental history of being driven under similar conditions.
- Assuming independent and stationary increments for $N(m)$ is equivalent to assuming iid for $\{n_1, n_2, \dots\}$.
- This highlights the similarity between stochastic processes and random vectors.

Poisson Process: A special counting process

- ◆ A counting process $N(t)$ is a Poisson process with rate λ if
 - i) $N(0) = 0$,
 - ii) $N(t)$ has independent and stationary increments,
 - iii) $P(N(t) = n) = \frac{\exp(-\lambda t)(\lambda t)^n}{n!}$.
- ◆ An alternative definition of the process is useful
- ◆ A function g is said to be order of number h and written as $g \sim o(h)$ if $\lim_{h \rightarrow 0} \frac{g(h)}{h} = 0$.
 - An $o(h)$ function **drops faster than the linear** function h as we get closer to 0.
- ◆ Alternative definition of Poisson process, maintain properties i) and ii) but replace iii) with
 - iii'a) $P(N(h) = 1) = \lambda h + o(h)$ and iii'b) $P(N(h) = 0) = 1 - \lambda h + o(h)$.
 - **At most 1 event** in a short interval with occurrence probability **proportional to the length** of interval
- ◆ Another alternative definition of Poisson process, maintain properties i) and ii) but replace iii) with
 - iii'') The interevent times are iid with $Expo(\lambda)$.



$$\text{iii)} \quad P(N(t) = n) = \exp(-\lambda t)(\lambda t)^n/n! \quad \longleftrightarrow$$

$$\text{iii')} \quad P(N(h) = 1) = \lambda h + o(h) \text{ \& } P(N(h) = 0) = 1 - \lambda h + o(h)$$

iii) \Rightarrow iii')

$$\blacklozenge \quad P(N(h) = 0) = \exp(-\lambda h) = 1 - \lambda h + \left(\frac{(\lambda h)^2}{2!} - \frac{(\lambda h)^3}{3!} \dots \right) = 1 - \lambda h + o(h)$$

$$\blacklozenge \quad P(N(h) = 1) = \exp(-\lambda h)\lambda h = \lambda h + \left(-(\lambda h)^2 + \frac{(\lambda h)^3}{2!} - \frac{(\lambda h)^4}{3!} \dots \right) = \lambda h + o(h)$$

iii') \Rightarrow iii)

$$\blacklozenge \quad \text{Let } p_i(t) = P(N(t) = i)$$

$$\blacklozenge \quad p_0(t+h) = P(N(t) = 0, N(t+h) - N(t) = 0) = P(N(t) = 0)P(N(h) = 0) = p_0(t)p_0(h) = p_0(t)(1 - \lambda h + o(h)),$$

which leads to

$$\frac{d}{dt}p_0(t) = \lim_{h \rightarrow 0} \frac{p_0(t+h) - p_0(t)}{h} = -\lambda p_0(t) + \lim_{h \rightarrow 0} \frac{o(h)}{h} = -\lambda p_0(t)$$

– This differential equation has the solution of the form $p_0(t) = c \exp(-\lambda t)$.

– Using the initial condition $p_0(0) = 1$, we obtain $c = 1$. Hence, $p_0(t) = \exp(-\lambda t)$.

$$\begin{aligned} \blacklozenge \quad p_i(t+h) &= P(N(t) = i, N(t+h) - N(t) = 0) + P(N(t) = i-1, N(t+h) - N(t) = 1) \\ &\quad + (P(N(t) = i-2, N(t+h) - N(t) = 2) + P(N(t) = i-3, N(t+h) - N(t) = 3) + \dots) \\ &= p_i(t)(1 - \lambda h + o(h)) + p_{i-1}(t)(\lambda h + o(h)) + o(h), \text{ which leads to, for } i \geq 1, \end{aligned}$$

$$\frac{d}{dt}p_i(t) + \lambda p_i(t) = \lim_{h \rightarrow 0} \frac{p_i(t+h) - p_i(t)}{h} + \lambda p_i(t) = \lambda p_{i-1}(t) + \lim_{h \rightarrow 0} \frac{o(h)}{h} = \lambda p_{i-1}(t)$$

\blacklozenge What remains to show: this equation has the solution $p_i(t) = \exp(-\lambda t)(\lambda t)^i/i!$. See the lecture notes.

$$\text{iii)} \quad P(N(t) = n) = \exp(-\lambda t)(\lambda t)^n/n! \quad \leftrightarrow$$

iii'') The interevent times are iid with $Expo(\lambda)$

- ◆ An algebraic equality used below: Differences of Gammas is Poisson

$$\begin{aligned} P(\text{Gamma}(n, \lambda) \leq t) &= \int_0^t e^{-\lambda x} \frac{(\lambda x)^{n-1}}{(n-1)!} \lambda dx = e^{-\lambda t} \frac{(\lambda t)^n}{n!} + \int_0^t e^{-\lambda x} \frac{(\lambda x)^n}{n!} \lambda dx \\ &= P(N(t) = n) + P(\text{Gamma}(n+1, \lambda) \leq t) \end{aligned}$$

It can be established via integration by parts; see lecture notes.

iii) \Rightarrow iii'')

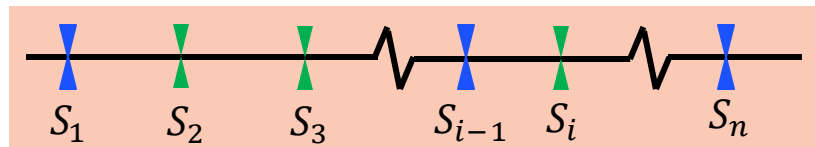
- ◆ $P(S_1 > t) = P(N(t) = 0) = \exp(-\lambda t)$ so $S_1 \sim \text{Gamma}(1, \lambda)$ and $X_1 \sim \text{Expo}(\lambda)$

- ◆ $P(S_n > t) = P(N(t) \in \{0, 1, \dots, n-1\}) = \sum_{i=0}^{n-1} P(N(t) = i)$
 $= P(N(t) = 0) + \sum_{i=1}^{n-1} P(N(t) = i)$
 $= e^{-\lambda t} + \sum_{i=1}^{n-1} \left(\int_0^t e^{-\lambda x} \frac{(\lambda x)^{i-1}}{(i-1)!} \lambda dx - \int_0^t e^{-\lambda x} \frac{(\lambda x)^i}{i!} \lambda dx \right)$
 $= e^{-\lambda t} + \int_0^t e^{-\lambda x} \lambda dx - \int_0^t e^{-\lambda x} \frac{(\lambda x)^{n-1}}{(n-1)!} \lambda dx$
 $= 1 - \int_0^t e^{-\lambda x} \frac{(\lambda x)^{n-1}}{(n-1)!} \lambda dx$ $S_n \sim \text{Gamma}(n, \lambda)$ and $X_n \sim \text{Expo}(\lambda)$

iii'') \Rightarrow iii)

- ◆ $P(N(t) = 0) = P(X_1 > t) = e^{-\lambda t}$ so $[N(t) = 0]$ has the Poisson probability
- ◆ $P(N(t) = n) = P(N(t) \geq n) - P(N(t) \geq n+1) \quad \longleftarrow P(N(t) \geq n) = P(N(t) = n) + P(N(t) \geq n+1)$
 $= P(S_n \leq t) - P(S_{n+1} \leq t)$
 $= e^{-\lambda t} \frac{(\lambda t)^n}{n!}$ so $[N(t) = n]$ has Poisson probability

Thinning of a Poisson Process



Some events are blue, some are green; with probabilities p_1, p_2 and $p_1 + p_2 = 1$. All events occur according to Poisson process. Occurrence of green or of blue events according to thinning of Poisson process.

- ◆ $N_1(t)$ is a the number of type 1 events; $N_2(t)$ is a the number of type 2 events.
- ◆ $N_1(t) + N_2(t) = N(t)$ and so $[N_1(t) = n_1 | N(t) = n]$ and $[N_2(t) = n_2 | N(t) = n]$ are dependent
 - $P(N_i(t) = n_i | N(t) = n) = P(\text{Bin}(n, p_i) = n_i)$ for $i \in \{1, 2\}$
 - Illustration of dependence
 - » $P(N_1(t) = n_1 | N(t) = n) P(N_2(t) = n_2 | N(t) = n) = P(\text{Bin}(n, p_1) = n_1) P(\text{Bin}(n, p_2) = n_2)$
 - » $P(N_1(t) = n_1, N_2(t) = n_2 | N(t) = n) = P(\text{Bin}(n, p_1) = n_1) = P(\text{Bin}(n, p_2) = n_2)$
- ◆ Conditional random variables are dependent; unconditional ones can still be independent, in 2 steps:

1. Unconditional random variable $N_1(t) \sim \text{Po}(\lambda p_1 t)$. Similarly $N_2(t) \sim \text{Po}(\lambda p_2 t)$.

$$\begin{aligned}
 - \quad P(N_1(t) = n_1) &= \sum_{n=n_1}^{\infty} P(N_1(t) = n_1 | N(t) = n) P(N(t) = n) = \sum_{n=n_1}^{\infty} P(\text{Bin}(n, p_1) = n_1) P(N(t) = n) \\
 &= \sum_{n=n_1}^{\infty} C_{n_1}^n p_1^{n_1} p_2^{n-n_1} e^{-\lambda t} \frac{(\lambda t)^n}{n!} = e^{-\lambda p_1 t} \frac{(\lambda p_1 t)^{n_1}}{n_1!} e^{-\lambda p_2 t} \sum_{n=n_1}^{\infty} \frac{1}{(n-n_1)!} (p_2 \lambda t)^{n-n_1} = e^{-\lambda p_1 t} \frac{(\lambda p_1 t)^{n_1}}{n_1!}
 \end{aligned}$$

2. Independence: $N_1(t) \perp N_2(t)$

$$\begin{aligned}
 - \quad P(N_1(t) = n_1, N_2(t) = n_2) &= P(N_1(t) = n_1, N_2(t) = n_2 | N(t) = n_1 + n_2) P(N(t) = n_1 + n_2) \\
 &= P(N_1(t) = n_1 | N(t) = n_1 + n_2) P(N(t) = n_1 + n_2) = C_{n_1}^{n_1+n_2} p_1^{n_1} p_2^{n_2} e^{-\lambda t} \frac{(\lambda t)^{n_1+n_2}}{(n_1 + n_2)!} \\
 &= e^{-\lambda p_1 t} \frac{(\lambda p_1 t)^{n_1}}{n_1!} e^{-\lambda p_2 t} \frac{(\lambda p_2 t)^{n_2}}{n_2!} = P(N_1(t) = n_1) P(N_2(t) = n_2)
 \end{aligned}$$

Number of Emails Until a Coupon

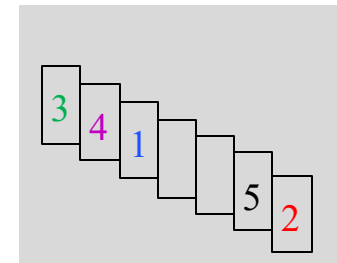
- ◆ Ex: A graduate student is interested in Texas style **horseback riding** but does not have much money to spend on this.
 - The student enrolls in livingsocial.com's Dallas email list to receive a discount offer for **horseback riding**.
 - Livingsocial sends offers for other events such as **piloting**, **karate lessons**, **gun carriage license training**, etc.
 - Livingsocial includes 1 discount offer in every email and that offer can be for any one of the events listed above.

What is the expected number of livingsocial emails that the student has to receive to find an offer for **horseback riding**?

- Probability p_1 is for each email to include a discount offer for **horseback riding**. Let N_1 be the index of the first email including a horseback riding offer. Then $P(N_1 = n_1) = p_1(1 - p_1)^{n_1-1}$, N_1 is a geometric random variable.

- ◆ Ex: Livingsocial sends offers for 5 types of events with probabilities $p_1 + p_2 + p_3 + p_4 + p_5 = 1$. What is the expected **number of emails** that the student has to receive **to obtain a discount offer for each event** type?

- Let N_i be the index of the first email including a discount offer for event type i .
- A particular realization can have $[N_1 = 3, N_2 = 7, N_3 = 1, N_4 = 2, N_5 = 5]$,
 - » the 1st email has an offer for **event 3**,
 - » the 2nd has an offer for **event 4**,
 - » the 3rd has an offer for **event 1**,
 - » the 4th and 6th must include offers for events **1**, **3** and **4**.
 - » the 5th has an offer for event **5**
 - » the 7th has an offer for **event 2**



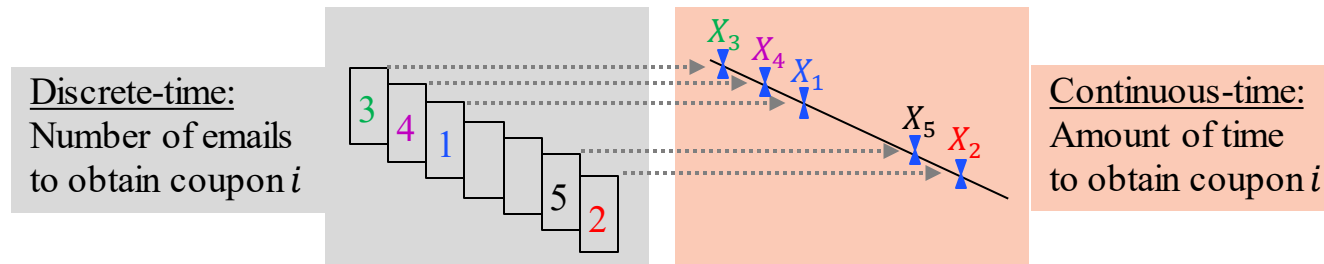
- In this realization, the number of emails to endure to obtain an offer for each event is $\max\{3, 7, 1, 2, 5\} = 7$.
- In general,

$$N = \max\{N_1, N_2, N_3, N_4, N_5\} \text{ for dependent } N_i\text{'s}$$

is the number emails to obtain an offer for each event.

- We want $E(N)$, but **N_i 's are dependent** even if they have marginal pmf's that are geometric distributions.
- How to break the **dependence**?

Amount of Time until a Coupon for Each Continuous-time Analog



- Suppose emails are sent at the rate of 1 per day. Number of emails $N(t) \sim Po(1t)$ with interevent time $Y_i \sim Expo(1)$.
- $N_i(t)$ is the number of coupons of type i obtained by time t . $N_i(t) \sim Po(1p_it)$
- $N_i(t)$ is obtained from $N(t)$ via thinning so $N_1 \perp N_2 \perp \dots \perp N_5$
- X_i is the interevent time for $N_i(t)$ process. $X_i \sim Expo(1p_i)$
 - X_i is the time between two emails both containing coupon i
- Amount of time until a coupon for each event

$$X = \max\{X_1, X_2, X_3, X_4, X_5\} \quad \text{for } X_1 \perp X_2 \perp \dots \perp X_5.$$
- Relating the number of emails to the time of the emails.
 - How many emails are received in X amount of time? How much time does it take to receive N emails?

$$N(t = X) \quad \text{or} \quad X = \sum_{i=1}^N Y_i$$

- Proceeding in either way: $E(N(X)) = 1E(X) = E(X)$ or $E(X) = E(Y_i)E(N) = 1E(N)$, so $E(N) = E(X)$.

- To finish, we need $E(X)$. First equality is from independence,

$$P(X \leq t) = \prod_{i=1}^5 P(X_i \leq t) = \prod_{i=1}^5 (1 - e^{-p_i t})$$

$$E(X) = \int_0^\infty P(X \geq t) dt = \int_0^\infty \left(1 - \prod_{i=1}^5 (1 - e^{-p_i t}) \right) dt$$

- For the number of event types approaching infinity, $E(X) \rightarrow \infty$. For a single event type, $E(X) = 1$.

Distribution of Event Times Conditioned on Number of Events

- Recall from order statistics: $\{X_1, X_2, \dots, X_n\}$ is an iid sequence, where $X_i \sim f_X$, and Y_i is the i th smallest element in the sequence. The joint pdf of $Y = [Y_1, Y_2, \dots, Y_n]$ is, for $y_1 \leq y_2 \leq \dots \leq y_n$,

$$f_Y(y_1, y_2, \dots, y_n) = n! \prod_{i=1}^n f_X(y_i)$$

- Ex: What is the joint pdf of the order statistics vector Y if each of underlying random variable $X_i \sim U(0, t)$?
 - The pdf of X_i is $f_X(x) = I_{x \in (0, t]} / t$. Then the joint pdf of the order statistics vector Y is, for $0 < y_1 < y_2 < \dots < y_n < t$,

$$f_Y(y_1, y_2, \dots, y_n) = n! \prod_{i=1}^n \frac{1}{t} = \frac{n!}{t^n}$$

- Regardless of the rate of Poisson, the conditional joint pdf of $([S_1, S_2, \dots, S_n] | N(t) = n)$ is

$$f_{[S_1, S_2, \dots, S_n] | N(t)=n}(s_1, s_2, \dots, s_n) = \frac{n!}{t^n} \quad \text{for } 0 < s_1 < \dots < s_n < t.$$

- First note

$$[S_1 = s_1, S_2 = s_2, \dots, S_n = s_n, N(t) = n] \quad \text{iff} \quad [X_1 = s_1, X_2 = s_2 - s_1, \dots, X_n = s_n - s_{n-1}, X_{n+1} \geq t - s_n].$$
- Hence,

$$\begin{aligned} f_{[S_1, S_2, \dots, S_n], N(t)}(s_1, s_2, \dots, s_n, n) &= f_X(s_1) f_X(s_2 - s_1) \dots f_X(s_n - s_{n-1}) (1 - F_X(t - s_n)) \\ &= \lambda e^{-\lambda s_1} \lambda e^{-\lambda(s_2 - s_1)} \dots \lambda e^{-\lambda(s_n - s_{n-1})} e^{-\lambda(t - s_n)} = \lambda^n e^{-\lambda t}. \end{aligned}$$

- Then going forward mechanically

$$f_{[S_1, S_2, \dots, S_n] | N(t)=n}(s_1, s_2, \dots, s_n) = \frac{f_{[S_1, S_2, \dots, S_n], N(t)}(s_1, s_2, \dots, s_n, n)}{P(N(t) = n)} = \frac{\lambda^n e^{-\lambda t}}{(\lambda t)^n e^{-\lambda t} / n!} = \frac{n!}{t^n}.$$

- From the last two examples, the conditional joint pdf of $([S_1, S_2, \dots, S_n] | N(t) = n)$ and the pdf of order statistics $[Y_1, Y_2, \dots, Y_n]$ with an underlying iid $X_i \sim U_i(0, t]$ have the same distribution.

$$(S_i | N(t) = n) = Y_i(U_1(0, t], \dots, U_n(0, t]) \text{ in distribution}$$

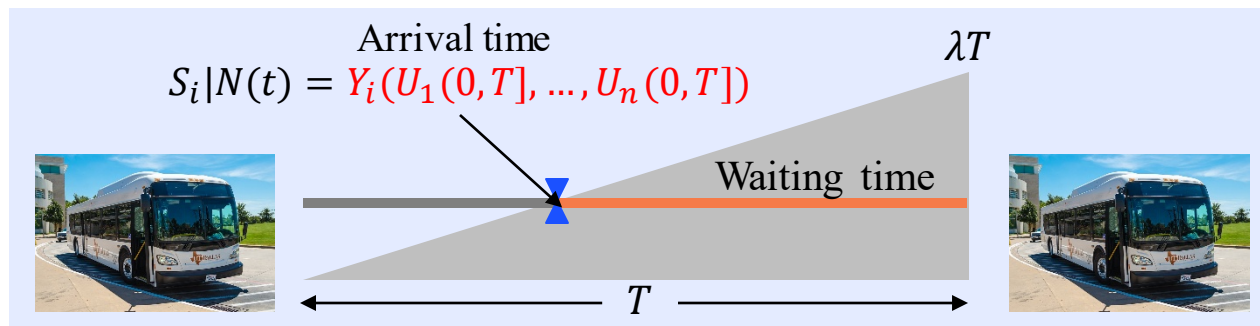
Expected Total Waiting Time for a Bus

- ◆ Ex: A DART bus departs from UTD to grocery markets every T minutes.
- Passengers arrive according to Poisson process $N(t)$ with rate λ and wait for the departure.
- E.g., the i th passenger that arrives over $(0, t]$ arrives at S_i and waits for $T - S_i$.
- The expected total waiting time is

$$\begin{aligned}
 E\left[\sum_{i=1}^{N(T)} (T - S_i)\right] &= E\left[E\left[\sum_{i=1}^{N(T)} (T - S_i) \mid N(T) = n\right]\right] = \sum_{n=0}^{\infty} E\left[\sum_{i=1}^{N(T)} (T - S_i) \mid N(T) = n\right] \frac{(\lambda T)^n e^{-\lambda T}}{n!} \\
 &= \sum_{n=0}^{\infty} E\left[\sum_{i=1}^n (T - Y_i(U_1(0, T], \dots, U_n(0, T]))\right] P(N(T) = n) \\
 &= \sum_{n=0}^{\infty} \left[\sum_{i=1}^n E(T - Y_i(U_1(0, T], \dots, U_n(0, T]))\right] P(N(T) = n) \\
 &= \sum_{n=0}^{\infty} \left[nT - E\left(\sum_{i=1}^n Y_i(U_1(0, T], \dots, U_n(0, T])\right)\right] P(N(T) = n) \\
 &= \sum_{n=0}^{\infty} \left[nT - \left(n \frac{T}{2}\right)\right] P(N(T) = n) = \frac{1}{2} T E(N(T)) = \frac{1}{2} T \lambda T
 \end{aligned}$$

Sum of order statistics = Sum of sample

$$\begin{aligned}
 E\left(\sum_{i=1}^n Y_i\right) &= E\left(\sum_{i=1}^n U_i\right) \\
 &= nE(U_i) = n \frac{T}{2}
 \end{aligned}$$



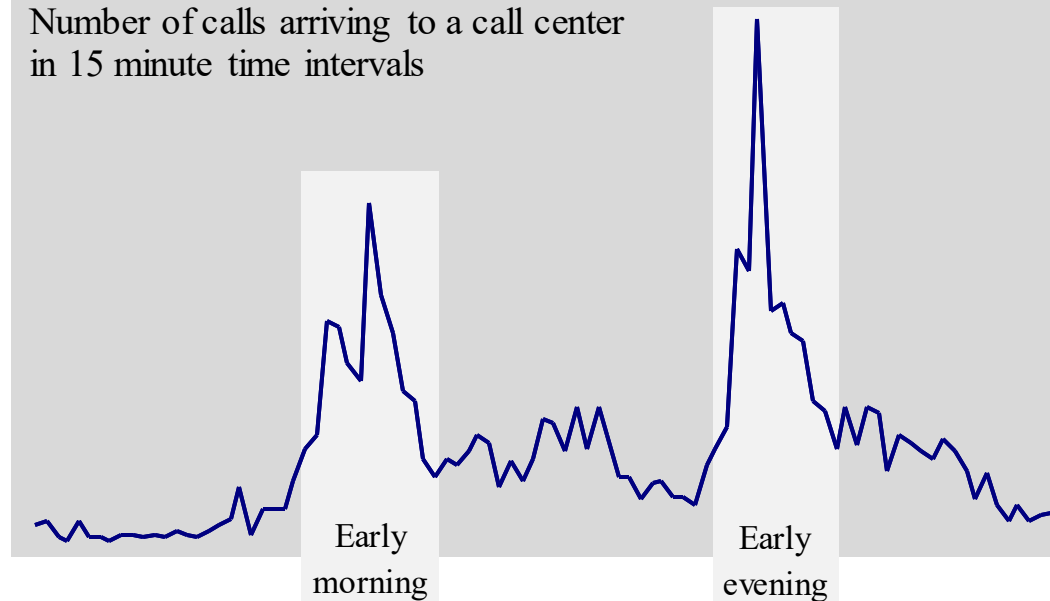
Area of the triangle
is the expected total
waiting time

Generalizations

◆ Nonhomogeneous Poisson process.

- Non-stationary ~~Stationary~~ Increments
- Time dependent event rate $\lambda(x)$ at time x
- The expected number of events over $(0, t]$ is $\int_0^t \lambda(x) dx$
- The number of events during $(0, t]$ has $Po\left(\int_0^t \lambda(x) dx\right)$
- Independent increments?
- To model varying traffic intensity

Number of calls arriving to a call center in 15 minute time intervals



- ◆ Compound Poisson process: A Poisson process $N(t)$ and an iid sequence $\{Y_i\}$ can be combined to obtain the compound Poisson process:

$$X(t) = \sum_{i=1}^{N(t)} Y_i$$

- Independent increments?
- To model batch arrivals of customers to a store (or buying multiple units at once)

Summary

- ◆ **Continuous-time Markov Process**
- ◆ **Poisson Process**
- ◆ **Thinning**
- ◆ **Conditioning on the Number of Events**
- ◆ **Generalizations**