Branching Processes

E.g. 1. Survival of family names

In patriarchical society: inherited by sons only.

Suppose each individual has prob p_k of having k sons.

Start w/ $X_0 = 1$

X_n = # individuals w/ family name in n^{th} generation.

Want: Prob(name dies out)

E.g. 2. Neutron chain reaction

Initial # of neutrons $X_0 = 1$

Nucleus split by chance collision w/ neutron

$\Rightarrow$ obtain random # of new 'offspring' secondary neutrons.

X_n = # neutrons produced by chance collisions of X_{n-1} neutrons.

Want: distr. of X_n.

E.g.3. Survival of mutant genes

Each indiv. gene has: $P(k \text{ offspring}) = p_k \quad k = 0, 1, 2, \ldots$

" " " can also transform into mutant gene

Mutant gene is then first in generations of mutant genes.

Assume ~~P~~ p.m.f. of mutant gene's # of descendants $\sim$ $Poi(\lambda)$

$\lambda > 1$ if mutation has biological advantage

$\lambda < 1$ " " " " disadvantage.

Want: $P(\text{survival of mutant gene}) = ?$

Branching Processes

Consider a class of M.C.s. s.t. $X_0 = 1$, the size of the 'zeroth' generation.

Suppose each individual produces j offspring w/ prob. P_j for $j = 0,1,2,\ldots$ $P_j \geq 0$, $\sum_{j=0}^{\infty} P_j = 1$, indep. of others.

X_n = size of n^{th} generation.

The M.C. $\{X_n, n \geq 0\}$ is called a branching process.

Some Properties:

(i) $E(X_n) = \mu^n$ where μ = Exp. # of offspring of an individual

Let Z_i = # of offspring of i^{th} individual of $(n-1)$st generation

Then
$$\begin{aligned} E(X_n) &= E(E(X_n \mid X_{n-1})) \\ &= \mu E(X_{n-1}) \\ &= \mu^2 E(X_{n-2}) \\ &\ \ \vdots \\ &= \mu^n E(X_0) = \mu^n \end{aligned}$$

Obviously: $\mu = \sum_j j P_j$

End 2/27/07

(2) 0 is recurrent. all other states are transient

State space = $\{0,1,2,\dots\}$

Examine state = 0, i.e. popn. dies out.

This is an 'absorbing' state, i.e.

$P_{00} = P(X_{n+1}=0 \mid X_n = 0) = 1$

So 0 is recurrent.

Examine other states:

If $p_0 > 0$ (+ve prob. of no offspring)

then, starting w/ i individuals

$P(\text{no later generation has } i \text{ individuals}) \geq \underbrace{p_0^{\,i}}_{\text{Prob(all } i \text{ individuals have no offspring)}} > 0$

i.e, $P(\text{never returning to } i) > 0 \quad i = 1,2,\dots$

$\Rightarrow$ all other states are transient

$\Rightarrow$ any finite set of states $\{1,2,\dots,n\}$ only visited finitely often

(3) $\Rightarrow$ $\boxed{\text{If } p_0 > 0 \text{ popn. will either die out or converge to } \infty.}$

Note: Property of recurrence/transience has direct implications!

Prob. that popn will eventually die out

$= \pi_0 = \lim_{n\to\infty} P(X_n = 0 \mid X_0 = 1)$

$$\pi_0 = \sum_{j=0}^{\infty} P(\text{extinction} \mid X_1 = j)\, p_j$$

$$\pi_0 = \sum_{j=0}^{\infty} \underbrace{\pi_0^{\,j}}_{\text{prob of each of } j \text{ 'tree' dying out}} p_j$$

Can show : π_0 is smallest +ve # that solves this equation. when $\mu > 1$.

<u>Thm</u>: $\pi_0 = 1$ iff $\mu \le 1$

No proof here: see Ross Stoch Proc argument using prob. gen. fns.

Easy to see why $\mu < 1 \Rightarrow \pi_0 = 1$

Recall Markov's Inequality : $X \geq 0$ then $P(X \geq a) \leq \frac{E(X)}{a}$ for any $a > 0$.

$P(X_n \geq 1 \mid X_0 = 1) \leq E(X_n \mid X_0 = 1)/1 = \mu^n$

$\Rightarrow$ Taking limits: $\lim_{n\to\infty} P(X_n \geq 1 \mid X_0 = 1) \leq \lim_{n\to\infty} \mu^n = 0$ $(\because \mu < 1)$

But $\pi_0 = 1 - \lim_{n\to\infty} P(X_n \geq 1 \mid X_0 = 1)$

So, $\pi_0 = 1 - 0 = 1$.

E.g. 1. Branching process w/ $p_0 = 1/2$, $p_1 = 1/4$, $p_2 = 1/4$.

$\mu = E$ (# offspring for 1 individual) $= 3/4 < 1$

Hence, by previous result, $\pi_0 = 1$, i.e., process will be extinct w/ prob. 1.

Note: 1 always satisfies eqn. $\pi_0 = \sum_{j=0}^{2} \pi_0^j p_j$

RHS $= 1$

E.g. 2. $p_0 = 1/3$, $p_1 = 1/3$, $p_2 = 1/3$

$\mu = 1$ so again $\pi_0 = 1$

E.g. 3 $p_0 = 1/4$, $p_1 = 1/4$, $p_2 = 1/2$

$\mu = 5/4 > 1$ so $\pi_0 < 1$

Now, π_0 soln to $\pi_0 = \sum_{j=0}^{\infty} \pi_0^j p_j$

$$\pi_0 = \frac{1}{4} + \frac{1}{4}\pi_0 + \frac{1}{2}\pi_0^2$$

$$2\pi_0^2 - 3\pi_0 + 1 = 0 \quad \text{so } \pi_0 = 1 \text{ or } 1/2$$

$\therefore$ $1/2$ is smallest positive soln, $\pi_0 = 1/2$.

Poisson Processes

Counting process: A stochastic process $\{N(t), t \geq 0\}$ is a counting process if $N(t)$ represents the # of events that have occurred up to time t. $N(t) \in \mathbb{Z}^+$

$N(t)$ is non-decreasing in t.

Recall: independent increments: # events occurred in disjoint intervals are independent

stationary increments: distr. of # of events occurred in a time interval only depends on length of interval (not on time/position)

"First principles" defn. of a Poisson process.

The counting process $\{N(t), t \geq 0\}$ is said to be a Poisson process w/ rate λ, $\lambda > 0$ if:

1. $N(0) = 0$
2. Process has stationary and independent increments
3. $P\{N(h) = 1\} = \lambda h + o(h)$
4. $P\{N(h) \geq 2\} = o(h)$

Function f is said to be $o(h)$ if $\lim_{h \to 0} \frac{f(h)}{h} = 0$

$o(h)$: Any fn. going to 0 faster than h.

(More generally, f is $o(g)$ if $\lim_{x \to \infty} \frac{f(x)}{g(x)} = 0$)

E.g. h^2 and h^9 are $o(h)$.
$h^{1/2}$ and $2h$ are not $o(h)$.

Assumption 3.: $P(N(t+h) - N(t) = 1) = \lambda h + o(h)$
(and stationarity)

In any small interval of length h, Prob (1 event) approx. λh.

Assumption 4:

In any small " " " h, Prob (more than 1 event) approx. 0.

The process defined above is equivalent to the Poisson process as defined next.

A counting process $\{N(t), t \geq 0\}$ is a Poisson process w/ rate λ, $\lambda > 0$ if

1. $N(0) = 0$
2. The process has independent increments
3. The process has stationary increments and

$$N(t+s) - N(s) \sim \text{Poisson}(\lambda t)$$

$$E(N(t)) = \lambda t$$
$$E(N(t+s) - N(s)) = \lambda t$$

<u>Pf</u>. (that the two definitions are equivalent)

We will use Laplace transforms, $g(u) = E\{e^{-uX}\}$ for r.v. X, $u \geq 0$.
Recall Laplace transform of a non-negative r.v. uniquely determines its distribution. Let $P_n(t) = P(N(t)=n)$, $n = 0,1,2$

To obtain Laplace transform for $N(t)$:
For $n \geq 1$:

$$P(N(t+h)=n) = P(\text{no event in interval of length } h)\, P(N(t)=n) + P(1 \text{ event in } h)\, P(N(t)=n-1) + \sum_{i=0}^{n-2} P(n-i \text{ events in } h)\, P(N(t)=i)$$

Covers cases w/ 2 or more events

$$p_n(t+h) = (1-\lambda h - o(h))\, p_n(t) + (\lambda h + o(h))\, p_{n-1}(t) + (o(h)) \sum_{i=0}^{n-2} p_i(t)$$

Hence,
$$\frac{p_n(t+h) - p_n(t)}{h} = \frac{-\lambda h - o(h)}{h} p_n(t) + \frac{\lambda h + o(h)}{h} p_{n-1}(t) + \frac{o(h)}{h} \sum_{i=0}^{n-2} p_i(t)$$

Taking $h \to 0$:
$$\frac{d}{dt} p_n(t) = -\lambda p_n(t) + \lambda p_{n-1}(t) \qquad \text{for } n = 1, 2, \ldots \qquad -(1)$$

For $n=0$:
$$P(N(t+h)=0) = P(\text{no event in interval of length } h)\, P(N(t)=0)$$
$$p_0(t+h) = (1-\lambda h - o(h))\, p_0(t)$$
$$\frac{p_0(t+h) - p_0(t)}{h} = \frac{(-\lambda h - o(h))}{h} p_0(t)$$
$$\frac{d}{dt} p_0(t) = -\lambda p_0(t) \qquad -(2)$$

To derive Laplace transform for $N(t)$,

let $g(t,u) = \sum_{n=0}^{\infty} e^{-un} p_n(t)$

Assuming we can interchange summation and differentiation,

$$\frac{d}{dt} g(t,u) = \sum_{n=0}^{\infty} e^{-un} \frac{d}{dt} p_n(t)$$

From ① & ② RHS $= \sum_{n=1}^{\infty} e^{-un}\left[-\lambda p_n(t) + \lambda p_{n-1}(t)\right]$

$$+ \; e^{-u \cdot 0}(-\lambda p_0(t))$$

$$= \sum_{n=0}^{\infty} e^{-un}(-\lambda p_n(t)) + \sum_{n=1}^{\infty} e^{-un} \lambda p_{n-1}(t)$$

$$= -\lambda \, g(t,u) + \sum_{n=0}^{\infty} e^{-u(n+1)} \lambda p_n(t)$$

$$= -\lambda \, g(t,u) + \lambda e^{-u} g(t,u)$$

i.e, $\frac{d}{dt} g(t,u) = \left\{-\lambda + \lambda e^{-u}\right\} g(t,u)$

Note that $g(0,u) = \sum_{n=0}^{\infty} e^{-un} p_n(0)$

$= e^{-u \cdot 0} \cdot 1$ $\quad (\because N(0) = 0$ a.s.$)$

$= 1.$ $\quad$ (bdry condition)

$\therefore$ Soln. to diff. eqn, is $g(t,u) = e^{\int(-\lambda + \lambda e^{-u})dt}$

$$= e^{\lambda t(e^{-u} - 1)}$$

But, this is Laplace transform of $Poi(\lambda t)$

Hence $N(t) \sim Poi(\lambda t)$

Note: could have used mgfs instead of Laplace transform.

From this defn (w/o involving distr.), naturally find connections to
- Poisson
- Binomial
- Exponential
- Uniform

Ch.5. Exponential Distr.

The exponential distr. $X \sim \text{Expon}(\lambda)$ (special case of Gamma)

pdf $f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$

cdf $F(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$

$E(x) = \frac{1}{\lambda}$

$Var(x) = 1/\lambda^2$

Memorylessness property:

$P(X > s+t \mid X > t) = P(X > s)$

$$LHS = \frac{P(X > s+t, X > t)}{P(X > t)} = \frac{P(X > s+t)}{P(X > t)} = \frac{1 - (1 - e^{-\lambda(s+t)})}{1 - (1 - e^{-\lambda t})}$$

$$= e^{-\lambda s} e^{-\lambda t} / e^{-\lambda t} = e^{-\lambda s} = P(X > s)$$

$= R.H.S.$

E.g.: Suppose waiting time for service at a bank is $\text{Expon}(\lambda = \frac{1}{10})$ (mean = 10 mins.)

Prob (waiting time for service > 15m.) $= e^{-15\lambda} = e^{-3/2} \approx 0.22$

~~$P(X > 15 \mid X > 10) = P(X > 5) = e^{-5\lambda} = e^{-1/2} = 0.604$~~

Prob (waiting time for service > 25m. | X > 10) $= e^{-3/2} = 0.22$

= Prob(" " " " > 15m)

Failure rate (hazard rate) for a cntns positive r.v. X w/ distr. function F and density f is defined by

$$r(t) = \frac{f(t)}{1-F(t)}$$

Interpretation: Suppose an item with lifetime X has survived for t hours, probability it will fail within time dt (instantaneously) is

$$P(X \in (t, t+dt) \mid X > t) = \frac{P(X \in (t, t+dt),\ X > t)}{P(X > t)}$$

$$= \frac{P(X \in (t, t+dt))}{1-F(t)} \approx \frac{f(t)\,dt}{1-F(t)} = r(t)\,dt$$

$r(t)$ = conditl. prob. density that a t year old item (that has not failed so far) will fail.

Prop. If $X \sim Exp(\lambda)$ then $r(t) = \lambda$ (constant).

Intuition: If $X \sim Exp(\lambda)$, by memoryless property distr. of remaining life for t year old item is same as for new item. Hence $\lambda(t)$ should be constant.

Easy to see why: $r(t) = \dfrac{f(t)}{1-F(t)} = \dfrac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda$

Prop : $r(t)$ uniquely determines the distribution F.

$$r(t) = \frac{\frac{d}{dt}F(t)}{1-F(t)}$$

integrating both sides

$$\int_0^t r(x)\,dx + k = -\log(1-F(t))$$

$$\Rightarrow \quad 1-F(t) = e^{-k} \exp\left\{-\int_0^t r(x)\,dx\right\}$$

Setting $t=0$ gives

$$1-0 = e^{-k}\cdot 1 \quad \Rightarrow \quad k=0$$

Hence $1-F(t) = \exp\left\{-\int_0^t r(x)\,dx\right\}$

$$\Rightarrow \quad F(t) = 1-\exp\left\{-\int_0^t r(x)\,dx\right\}$$

Corr. : If X is cntns. positive r.v. w/ $r(t)$ constant, i.e., $r(t)=c$ $\forall t$, then X must be an exponential r.v.

Implication: if X is memoryless it must be exponential.

Argument: $F(t) = 1-\exp\left\{-\int_0^t c\,dx\right\} = 1-\exp(-ct)$, i.e.,

$$X \sim \text{Exp}(c).$$

Properties of expon. r.v.'s

1. $X_i \overset{iid}{\sim} Exp(\lambda) \quad i=1,\dots n$

$$Y = \sum_{i=1}^{n} X_i \sim Gamma(n, \lambda)$$

$$f(y) = \lambda e^{-\lambda y} \frac{(\lambda y)^{n-1}}{(n-1)!}$$

2. $X_1 \sim Exp(\lambda_1)$ indep. of $X_2 \sim Exp(\lambda_2)$

$$P(X_1 < X_2) = \frac{\lambda_1}{\lambda_1+\lambda_2}$$

3. $X_i \sim Exp(\lambda_i) \quad i=1,\dots,n$

a) If $Z = \min(X_1,\dots,X_n)$ $\quad Z \sim Exp\left(\sum_{i=1}^{n} \lambda_i\right)$

b) If $Y = \max(X_1,\dots,X_n)$ $\quad F_Y(y) = P(Y < y) = \prod_{i=1}^{n} \left(1-e^{-\lambda_i y}\right)$

Poisson approx. to Binomial (informal)

$X \sim \text{Bin}(n, p)$

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Let $np = \lambda$, i.e. $p = \lambda/n$

$$P(X=k) = \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{n!}{n^k(n-k)!} \frac{\lambda^k}{k!} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k}$$

Now, taking limit as $n \to \infty$

$$\lim_{n\to\infty} P(X=k) = \lim_{n\to\infty}\left[\frac{n!}{n^k(n-k)!}\right] \frac{\lambda^k}{k!} \, e^{-\lambda} \cdot 1$$

Now, $\displaystyle\lim_{n\to\infty} \frac{n!}{n^k(n-k)!} = \lim_{n\to\infty} \frac{n(n-1)\dots(n-(k-1))}{n \;\; n \;\dots\dots\; n} = 1$

So, $\displaystyle\lim_{n\to\infty} P(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}$

more formal pf: use prob. gen. fns.

Connecting "1st principles" defn. to the result that $N(t)$ has a Poisson distr. : this is a consequence of Poisson approx. to Binomial.

Subdivide interval $[0, t]$ into k equal parts where k is very large.

$\lim_{k\to\infty}$ Prob (2 or more ~~intervals~~ events in an sub-interval) $= 0$

$P($2 or more events in any sub-interval$) \le \sum_{i=1}^{k} P($2 or more in i^{th} subinterval$)$

$$= k\, o(t/k) = \frac{t}{t} k\, o(t/k) = t \frac{o(t/k)}{t/k}$$

$$\lim_{k\to\infty} t \frac{o(t/k)}{t/k} = 0.$$

Hence, $N(t)$ (w/ a prob. going to 1) will equal # of subintervals in which an event occurs.

However, by stationarity and independent increments, this number, $N(t) \sim$ Binomial$(k, p = \lambda(\frac{t}{k}) + o(t/k))$.

By Poisson approx. to Binomial, as $k \to \infty$ ($\because$ n and p get large) and small respectively)

$$N(t) \sim Poi\left(\lim_{k\to\infty} \left[k \left(\lambda\left(\frac{t}{k}\right) + o(t/k)\right)\right] \right)$$

$$\text{mean} = \lim_{k\to\infty} k \lambda t/k + \lim_{k\to\infty}\left[t \frac{o(t/k)}{t/k}\right]$$

$$= \lambda t$$

Poisson Process : <u>Connection to Exponential Distr.</u>

Consider a Poisson process $N_\lambda(t)$ w/ rate λ and let time of 1^{st} event be T_1.

T_n = elapsed time between $(n-1)$st and nth event.

$\{T_1, T_2, \ldots\}$ sequence of <u>interarival times.</u>

Now, $P(T_1 > t) = P(N(t) = 0)$

$= P(N(t) - N(0) = 0) = (\lambda t)^0 e^{-\lambda t}/0! = e^{-\lambda t}$

$\Rightarrow T_1 \sim Exp(\lambda)$

And, $P(T_2 > t) = E\{P(T_2 > t \mid T_1)\}$

$\overset{s}{\bullet}\ T_1 \quad t \quad \overset{s+t}{\bullet}\ T_2$

But $P(T_2 > t \mid T_1 = s) = P(0 \text{ events in } (s, s+t] \mid T_1 = s)$

(indep. incr.) $= P(0 \text{ events in } (s, s+t])$

(stat. incr.) $= e^{-\lambda t}$ So, $E\{P(T_2 > t \mid T_1)\} = e^{-\lambda t}$

Hence $T_2 \sim Exp(\lambda)$

Can repeat this argument to see:

$T_n \overset{i.i.d.}{\sim} Exp(\lambda)$.

Hence arrival time of n^{th} event, $S_n = \sum_{i=1}^{n} T_i$

$\sim Gamma(n, \lambda)$ $E(S_n) = n/\lambda$.

Alt. Defn. of Poisson process $N(t)$ w/ rate λ:

Let $T_1, \ldots, T_n, \ldots \overset{iid}{\sim} Expon(\lambda)$

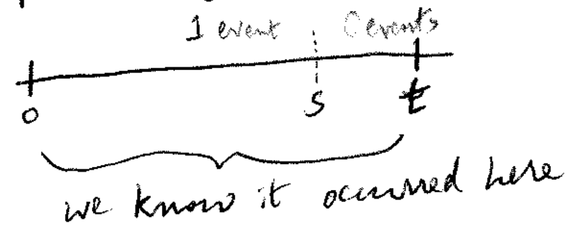
Define $N(t)$ s.t. n^{th} event occurs at time $S_n = \sum_{i=1}^{n} T_i$.

Connection to Uniform distr.

Condtl. distr. of arrival times

Let $\{N(t), t \geq 0\}$ be a Poisson process w/ rate λ. Let T_1 be arrival time of 1^{st} event.

Distr. of $(T_1 \mid \text{exactly one event has occurred before } t) = ?$



cdf: $s \leq t$

$$P(T_1 < s \mid N(t)=1) = \frac{P(T_1 < s, N(t)=1)}{P(N(t)=1)}$$

$$= \frac{P(N(t-s)=0, N(s)=1)}{\frac{\lambda t e^{-\lambda t}}{1!}} \overset{\text{indep. incr.}}{=} \frac{P(N(t-s)=0)P(N(s)=1)}{\lambda t e^{-\lambda t}}$$

$$= \frac{(\lambda(t-s))^0 e^{-\lambda(t-s)}}{0!} \frac{(\lambda s)^1 e^{-\lambda s}}{1!} \Big/ (\lambda t) e^{-\lambda t}$$

$$= \frac{e^{-\lambda t + \lambda s} \cdot \lambda s \cdot e^{-\lambda s}}{\lambda t e^{-\lambda t}} = \frac{s}{t} I(0 \leq s \leq t)$$

$\Rightarrow T_1 \mid \text{exactly 1 event before } t \sim \text{Unif}(0, t).$

Generalization: Let $S_n = \sum_{i=1}^{n} T_i$ = arrival time of n^{th} event

$(S_1, \ldots, S_n) \mid N(t) = n$ ~~has~~ same distr. as order stats. corresponding to n i.i.d. Unif$(0,t)$ r.v.'s. + Interpretation

$f((s_1, \ldots, s_n) \mid N(t) = n) = \frac{n!}{t^n} I(0 < s_1 \ldots < s_n < t)$

Recall: If $Y_i \overset{iid}{\sim} f$ cntns, $f_{Y_{(1)}, \ldots, Y_{(n)}}(y_1, \ldots, y_n) = n! \prod_{i=1}^{n} f(y_i) I(y_1 < \ldots$

Thinning of Poisson processes

Suppose $\{N(t), t \geq 0\}$ is a Poisson process w/ rate λ,
but events belong to two types: $Pr(\text{Type I}) = p$, $Pr(\text{Type II}) = 1-p$.
Type is indep. of everything else.
Let $N_1(t) =$ # Type 1 events up to time t
$N_2(t) =$ " " 2 " " " " "
$\{N_1(t), t \geq 0\}$ and $\{N_2(t), t \geq 0\}$ are indep. Poisson
processes w/ rates λp, $\lambda(1-p)$, resp.
(See Pf. in text)
Equivalent idea: a random thinning of a poisson process (rate λ)
is obtained by deleting events in a series of
mutually independent Bernoulli trials. (w/ prob. deletion $= 1-p$.)
Thinned process is Poisson w/ rate λp
Eg. spatial pt. process where $Pr(\text{observation of event}) = p$, results in thinned process.
E.g. Immigrants arrive according to Poi ($\lambda = 10/\text{week}$).
Each is of English descent w/ prob. $1/12$.
Pr(No English descent immigrants arrive in 4 weeks)

$$= Pr\left(N_{English}(4) = 0\right)$$

$$= \frac{\left(\frac{10}{12}4\right)^0 e^{-\left(\frac{10}{12}4\right)}}{\left(\frac{10}{12}\cdot 4\right)^0} = e^{-10/3}.$$

Generalizations of Poisson process

Nonhomogeneous Poisson process: Counting process $\{N(t), t \geq 0\}$ is a non-hom. P. proc. w/ intensity function, $\lambda(t), t \geq 0$ if

1. $N(0) = 0$
2. Process has <u>indep.</u> increments [still a model for independence]
3. $N(t+s) - N(t) \sim Poi(m(t+s) - m(t))$ where

$$m(t) = \int_0^t \lambda(x)\,dx$$

Non-stationary increments due to dependence on t.

If $\lambda(t) = \lambda$, constant, get homogeneous Poisson process.
Otherwise, expect more events at some times than at others.
Relate to other variables: e.g. $\log \lambda(t) = \beta X(t)$ modulated P.P. Stochastic reln $\log \lambda(t) = \beta X(t) + \varepsilon(t)$ "Cox process"

Compound Poisson Process:
$\{X(t), t \geq 0\}$ is a compound Poisson process if it can be represented as $X(t) = \sum_{i=1}^{N(t)} Y_i,\ t \geq 0$ where $\{N(t), t \geq 0\}$ is a Poisson process, and $\{Y_i, i \geq 1\}$ is a family of iid r.v.'s indep. of $N(t)$.

$Y_i = 1\ \forall i$ gives usual Poisson process.

E.g. $N(t)$, # buses, arrives according to a Poisson process.
Y_i, # people in each bus is iid f. e.g.
Total # people arrived by time t is a compound Poisson process.

Hierarchical: # observed $\sim$ P. Proc. Each obs $\overset{iid}{\sim} f$.