Lemma: If $X_1, \dots, X_n$ are independent $X_i \sim \exp(-\lambda_i x)$, $\lambda_i \neq \lambda_j$

then $f_{X_1+\dots+X_n}(x) = \sum_{i=1}^{n} C_{i,n} \lambda_i \exp(-\lambda_i x)$

where $C_{i,n} = \prod_{i \neq j} \frac{\lambda_j}{\lambda_j - \lambda_i}$

Hence $1 - F(t) = \sum_{i=1}^{n} C_{i,n} \exp(-\lambda_i x)$

failure rate: $\frac{\sum_{i=1}^{n} C_{i,n} \lambda_i \exp(-\lambda_i x)}{\sum_{i=1}^{n} C_{i,n} \exp(-\lambda_i x)}$ $\Big| \; \frac{f(x)}{1-F(x)}$

If $\lambda_j = \min_i \lambda_i$, then multiplying numerator and denominator with $e^{\lambda_j t}$ and letting $t \to \infty$

$$\lim_{t \to \infty} r(t) = \lambda_j = \min(\lambda_i)$$

feb. 27-10:14

so if we know that the survival is up to t then the conditional probability of dying has approximately p.d.f $\lambda_j \exp(-\lambda_j t)$.

Coxian distributions have a random number of terms in the sum.

Let $X_1, \dots, X_n$ be independent $X_i \sim \exp(-\lambda_i x)$, $\lambda_i \neq \lambda_j$

Let N be independent of $X_1, \dots, X_n$ and $P_n = P(N = n)$. Then $Y = \sum_{j=1}^{N} X_j$ is a Coxian random variable.

The p.d.f of Y is

$$f_Y(t) = \sum_{n=1}^{N} f_Y(t \mid N=n) \cdot P_n$$

$$= \sum_{n=1}^{N} f_{X_1+\dots+X_n}(t \mid N=n) \cdot P_n$$

$$= \sum_{n=1}^{N} f_{X_1+\dots+X_n}(t) \cdot P_n$$

$$= \sum_{n=1}^{N} P_n \cdot \sum_{j=1}^{n} C_{i,n} \lambda_i \exp(-\lambda_i t)$$

feb. 27-10:23

Visualization

$\boxed{X_1} + \boxed{X_2} + \boxed{X_3} + \left(\boxed{X_4} + \dots + \boxed{X_N} \right)$

$N = 3$

5.3 The Poisson process

The Poisson process is the simplest process in continuous time. The process describes the occurrence of events and how the occurence can be described by continuously distributed random variables.

Examples: (i) Customers entering a shop

(ii) Children born in a city/country

(iii) goals scored by a soccer player.

feb. 27-10:32

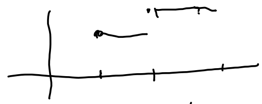
A basic tool is the counting process $N(t)$ which indicates the number of events up to and including t. Thus

(i) $N(t) \geq 0$

(ii) $N(t)$ integer valued

(iii) $N(t)$ is nondecreasing so if $s < t$ $N(s) \leq N(t)$

(iv) For $s < t$, $N(t) - N(s)$ is the number of events in the half open interval $(s, t]$. Thus $N(t)$ is continuous from the right



There are two imporant properties that a counting process may have or not

feb. 27-10:40

(i) <u>independent increments</u> which mean that number of events in <u>disjoint intervals</u> are indepent.

(ii) <u>stationary increments</u> mean that the number of events in an interval only depends on the <u>length of the interval</u>

Notation: A function is $o(h)$ if

$\frac{f(h)}{h} \to 0$ as $h \to 0$

Examples, x^2 is $o(h)$, x is not $o(h)$,

If f, g are $o(h)$ then $f + g$ is $o(h)$

cf is $o(h)$ if f is $o(h)$. All linear combination of function which are $o(h)$ are also $o(h)$.

Example: $X \sim f(x)$

$$P(t < X < t+h) = \frac{F(x+h) - F(x)}{h} \cdot h$$

$$= [f(t) + o(h)]h = f(x) \cdot h + o(h).$$

feb. 27-10:45

Definition: The counting process $\{N(t); t \geq 0\}$ is a Poisson process with rate λ if

(i) $N(0) = 0$

(ii) $\{N(t), t \geq 0\}$ has independent increments

(iii) $P(N(t+h) - N(t) = 1) = \lambda \cdot h + o(h)$

(iv) $P(N(t+h) - N(t) \geq 2) = o(h)$

Remark $P(N(t+h) - N(t) = 0) = 1 - \lambda h + o(h)$

Theorem 5.1 If $\{N(t), t \geq 0\}$ is a Poisson process with rate λ, then $N(s+t) - N(s)$ $s, t \geq 0$ i.e the number of events in $(s, s+t]$ is Poisson distributed with rate λ.

Proof. $g(t) = E[e^{-u N(t)}]$

Then $g(t+h) = E[\exp(-u N(t+h))]$

$$= E\, e^{-u[N(t+h) - N(t) + N(t)]}$$

$$= E\, e^{-u[N(t+h) - N(t)] - u \cdot N(t)}$$

feb. 27-10:59

By independent increments

$$= E[e^{-uN(t)}] \cdot E[e^{-u[N(t+h)-N(t)]}]$$

$$= g(t) \cdot \underline{E\, e^{-u[N(t+h)-N(t)]}}$$

But $E\, e^{-u[N(t+h)-N(t)]}$

$$= E\{e^{-u[N(t+h)-N(t)]} \cap N(t+h)-N(h)=0\}$$

$$+ E\{e^{\cdot} \text{ — '' — } \cap N(t+h)-N(h)=1\}$$

$$+ E\{ \text{ — '' — } \cap N(t+h)-N(h)\geq 2\}$$

$$= [1-\lambda h+o(h)] + e^{-u}[\lambda h+o(h)] + o(h)$$

$$= 1-\lambda h+e^{-u}\lambda h+o(h)$$

$$= 1+(e^{-u}-1)\lambda h+o(h)$$

feb. 27-11:27

Hence $g(t+h) = g(t)[1+(e^{-u}-1)\lambda h+o(h)]$

$$\frac{g(t+h)-g(t)}{h} = g(t)\lambda\cdot(e^{-u}-1) + \frac{o(h)}{h}$$

Let $h \to 0$, then

$$g'(t) = g(t)\lambda(e^{-u}-1) \text{ or } \frac{g'(t)}{g(t)} = \lambda\cdot(e^{-u}-1)$$

Hence $\frac{\partial}{\partial t}\log g(t) = \lambda\cdot(e^{-u}-1)$

$$\log g(t) = \lambda(e^{-u}-1)\cdot t + C$$

But $g(0)=1 \Rightarrow C=0$ and $g(t) = \underline{e^{\lambda t\cdot(e^{-u}-1)}}$

But if $X \sim Po(\lambda t)$, $E\, e^{-uX} = e^{-\lambda t}\sum \frac{(\lambda\cdot e^{-u})^i}{i!}$

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

feb. 27-11:34

Hence the two Laplace transforms are equal. And since the Laplace transform determines the distribution $N(t)$ is Poisson distributed with parameter λt.

Let $N_s(t) = N(s+t) - N(s)$. Then $N_s(t)$ is a counting process starting in 0 and satisfying the requirements in the definition of a Poisson process.

Hence $N(s+t)-N(s) = N_s(t)$ is Poisson distributed with parameter λt. ∎

feb. 27-11:40

Heuristic argument: ~~for~~

If $Y \sim Bin(n,p)$ then $P(Y=y) = \binom{n}{y}p^y(1-p)^{n-y}$

$$\rightarrow \frac{\lambda^y}{y!}e^{-\lambda} \quad np \to \lambda$$

see section 2.2.4 p29 in Ross

"success": event in interval of length t/k

$P(\text{success}) = P(\text{event in interval of length } t/k)$

$$= \lambda\cdot\frac{t}{k} + o(t/k)$$

by stationarity

Hence $P(y \text{ success}) = \binom{k}{y}[\lambda\cdot\frac{t}{k}+o(t/k)]^y [1-\lambda\frac{t}{k}+o(t/k)]^{k-y}$

feb. 27-11:45

since $k\cdot[\lambda\cdot\frac{t}{k}+o(\frac{t}{k})] \to \lambda t$ as $k\to\infty$

then $P(y \text{ success}) \to \frac{(\lambda t)^y}{y!}e^{-\lambda t}$

from the Poisson approximation to the Binomial distribution

Also, the distribution of $N(t+s)-N(s)$ is the same for all s, so $N(t+s)-N(s)$ has stationary increments. Thus also the Poisson process has stationary increments.

feb. 27-11:52

Next we will look into the distribution of time to first event, T_1, and the times between successive events which are denoted by $T_2, T_3, \ldots$. The variables $T_1, T_2, \ldots$ are called the sequence of interarrival times

feb. 27-11:58