M3S4/M4S4 Applied Probability

Emma J. McCoy

Room: 522 Huxley

email: `e.mccoy@ic.ac.uk`

www: `stats.ma.ic.ac.uk/~ejm/M3S4`

Department of Mathematics

Imperial College of Science Technology and Medicine

180 Queen's Gate, London SW7 2BZ

Spring Term, 2008

1 Introduction

Most of this course is about RANDOM PROCESSES or STOCHASTIC PROCESSES (SPs). SPs are systems which evolve in time (usually) whilst undergoing random fluctuations. This random variation is ubiquitous, one of the greatest steps in understanding how the world works is through understanding chance:

- by learning what can be said about inherently unpredictable events.
- by learning how to manipulate the occurrence of inherently unpredictable events.

This course is about the practical aspects of building useful models of real events and discovering properties of these models NOT the mathematical niceties involved.

1.1 Examples

1. **Epidemics:** One of you catches 'flu

   - what is the chance you'll catch it?
   - what is the chance you'll catch it just before an exam?
   - How many will catch it?
   - will it spread throughout the college?
   - will some people die of it? how many?
   - How can we prevent a potential epidemic?

2. **Genetics:** If a couple's first child is colour blind, what is the probability that subsequent children will be?

3. **Network traffic:** In a network which passes messages randomly between nodes:

   - what's the chance that two arrive together?
   - what if the links have different transmission speeds?

- is the behaviour different if the messages are of random lengths/different lengths?

4. **Simulated Annealing:** a stochastic optimization strategy which can be guaranteed to find the global maximum of a multimodal function.

   - how should the parameters be chosen to get the best results?

   - how long will it take?

5. **Markov Chain Monte Carlo (MCMC):** is a modern statistical tool which brings together results from SPs and simulation enabling potentially computationally intractable (Bayesian) statistics to be used in practice.

2 Revision

2.1 Notation

Ω = event space = set of possible outcomes

- Coin: $\Omega = \{\mathrm{H}, \mathrm{T}\}$.
- Die: $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- Number of people in a queue at a given time.
- Number of cars passing M1 bridge per hour.

Event is a subset of $\Omega : A \subset \Omega$.

Probability is a mapping $\mathrm{P} : \Omega \to \mathbb{R}$ such that

I: $0 \leq \mathrm{P}(A) \leq 1 \;\forall$ events A.

II: $\mathrm{P}(\Omega) = 1$.

III: If $A_i \cap A_j = \phi \;\forall\; i \neq j$ then

$$\mathrm{P}(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=1}^{n} \mathrm{P}(A_i).$$

Addition Law: $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$.

Conditional Prob:

$$\mathrm{P}(A \,|\, B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \quad \mathrm{P}(B) \neq 0.$$

Independence: $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$

Complement: $\mathrm{P}(A) = 1 - \mathrm{P}(\overline{A})$.

De Morgans Laws:

$$\begin{aligned} \mathrm{P}(\overline{A \cup B}) &= \mathrm{P}(\overline{A} \cap \overline{B}) \\ \mathrm{P}(\overline{A \cap B}) &= \mathrm{P}(\overline{A} \cup \overline{B}) \end{aligned}$$

Theorem of Total Probability:

$$\mathrm{P}(A) = \sum_i \mathrm{P}(A \,|\, B_i)\mathrm{P}(B_i) \quad \text{where} \quad B_i \cap B_j = \phi \;\forall\; i \neq j; \cup_i B_i = \Omega.$$

Def: A random variable is a function

$$X : \Omega \to \mathbb{R}$$

Upper case: the name of the random variable

Lower case: particular values....

$$\text{e.g.} \quad \mathrm{P}(X = x) = 0.6 \quad (\mathrm{P}(x)).$$

Def: The probability function of a r.v. X gives the probability that X takes a particular range of values in $\mathbb{R}$.

2.2 Discrete Random Variables

Range of X contains finite or countably infinite number of points.

e.g. 1: Discrete Uniform

$$\mathrm{P}(x) = \frac{1}{n} \qquad \Omega_X = n \text{ equally spaced points.}$$

e.g. 2: Bernoulli

$$\mathrm{P}(x) = p^x(1-p)^{1-x} = p^x q^{1-x} \qquad \Omega_X = \{0, 1\}.$$

e.g. 3: Binomial $p = \mathrm{P}(\text{success})$ $\quad X =$ Number of successes in n trials:

$$\mathrm{P}(x) = \binom{n}{x} p^x q^{n-x} \qquad \Omega_X = \{0, 1, \ldots, n\}, \qquad X \sim \mathit{Binomial}(n, p).$$

e.g. 4: Geometric

$p = \mathrm{P}(\text{success})$ $\quad X =$ Number of trials up to and including first success:

$$\mathrm{P}(x) = q^{x-1}p \qquad \Omega_X = \{1, 2, \ldots\} \qquad X \sim G_1(x)$$

$p = \mathrm{P}(\text{success})$ $\quad X =$ Number of trials before first failure:

$$\mathrm{P}(x) = qp^x \qquad \Omega_X = \{0, 1, \ldots\} \qquad X \sim G_0(x)$$

e.g. 5: Negative Binomial

$p = \mathrm{P}(\text{success})$ $\quad$ $X =$ Number of trials up to and including kth success:

$$\mathrm{P}(x) = \binom{x-1}{k-1} p^k q^{x-k} \qquad \Omega_X = \{k, k+1, \ldots\}$$

e.g. 6: Poisson

$$\mathrm{P}(x) = \frac{e^{-\mu}\mu^x}{x!} \qquad \Omega_X = \{0, 1, \ldots\} \qquad X \sim \mathit{Poisson}(\mu).$$

2.3 Continuous Random Variables

Continuous sample space e.g. time, length.

CDF:

$$F_X(x) = \mathrm{P}(X \leq x) \qquad x \in \mathbb{R}.$$

A continuous rv is one which has a continuous CDF.

PDF:

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

Expectation:

$$\mathrm{E}(X) = \int x f(x)\, dx.$$

If X can take only positive values

$$\mathrm{E}(X) = \int_0^\infty [1 - F(x)]\, dx.$$

e.g. 1: Uniform

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

e.g. 2: Exponential

$$f(x) = \lambda e^{-\lambda x} \qquad x \geq 0; \qquad X \sim \mathit{Exponential}(\lambda).$$

e.g. 3: Gamma

$$f(x) = \frac{x^{n-1}\lambda^n e^{-\lambda x}}{(n-1)!} \qquad x \geq 0; \qquad X \sim \mathit{Gamma}(n, \lambda).$$

If $X_i, i = 1, \ldots, n$ are iid exponential rvs. Then $X = \sum X_i \sim \mathit{Gamma}(n, \lambda)$.

e.g. 4: Normal

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \qquad x \in R; \qquad X \sim N(\mu, \sigma^2).$$

Standard normal distribution: $\mu = 0, \sigma = 1$.

Notation for normal: pdf: $f(x) = \phi(x)$, cdf: $F(x) = \Phi(x)$.

Central Limit Theorem:

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \approx N(0, 1).$$

3 Random/Stochastic Processes

A collection of rvs $\{X(t); t \in \mathbb{R}\}$ or $\{X_n; n = 0, 1, 2, \ldots\}$

Discrete Time random process - observed only at specific times.

e.g. Gambler's ruin:

$$\text{Player } A \text{ has } \pounds k$$
$$\text{Player } B \text{ has } \pounds(a-k), \quad a > k > 0$$

Play a series of games in which A has prob. p of winning and $q = 1 - p$ of losing.

Define rv X_n : A's money after n games.

Then $(X_1, X_2, \ldots, X_n)$ is a realisation of a discrete time random process.

Continuous Time random process.

e.g. Number of customers in queue for ATM

e.g. Angle of barometer pointer

Discrete State Space

e.g. Gambler's ruin.

e.g. ATM customers

Continuous State Space

e.g. Angle of barometer pointer

3.1 Some Fundamental Random Processes

1. **Bernoulli Process**

   A sequence of Bernoulli trials: $Y_1, Y_2, \ldots$. Each trial independent, same p.

   Some questions of interest:

   Q1: $X_n = \sum_{i=1}^n Y_i$: What is the distribution of X_n?

   Q2: What is the distribution of X_n given X_{n-1}?

A1: X_n is number of successes in n trials $\Rightarrow X_n \sim \mathit{Binomial}(n, p)$

A2: Only two possible values

$X_n = X_{n-1}$ or $X_n = X_{n-1} + 1$

$\mathrm{P}(X_n = x \,|\, X_{n-1} = x) = \mathrm{P}(Y_n = 0) = 1 - p$

$\mathrm{P}(X_n = x + 1 \,|\, X_{n-1} = x) = \mathrm{P}(Y_n = 1) = p$

2. **Poisson Process**

Continuous time analogue of Bernoulli. Events occur at "random" times.

I: P(exactly 1 event occurs in any time interval of length δt) $= \lambda\delta t + o(\delta t)$

$[o(\delta t)/\delta t \to 0$ as $\delta t \to 0]$

II: $\mathrm{P}(2 \text{ or more} \ldots \delta t) = o(\delta t)$

III: Occurrence of events after time t is independent of occurrence of events before t.

Let $X(t)$ = number of events by t. Then $X(t) \sim \mathit{Poisson}(\lambda t) \quad (X(0) = 0)$.
(proof later....)

Let T_k be the time between $(k-1)$th and kth events,

$$\begin{aligned} \mathrm{P}(T_1 > t) &= \mathrm{P}(\text{no of events in } [0, t]) \\ &= \mathrm{P}(X(t) = 0) \\ &= \frac{e^{-\lambda t}(\lambda t)^0}{0!} = e^{-\lambda t}. \end{aligned}$$

Let cdf of T_1 be $F(t)$. Then $F(t) = 1 - \mathrm{P}(T_1 > t) = 1 - e^{-\lambda t}$

So $f(t) = \lambda e^{-\lambda t}$ (Exponential).

By **III**, T_k has the same distribution.

Let W_n be the time to the nth event.

So, $W_n = T_1 + \ldots + T_n$

Since $T_i \sim \mathit{Exponential}(\lambda)$ iid, $W_n \sim \mathit{Gamma}(n, \lambda)$.

3. **Simple Birth Process**

A population of individuals, such that each one gives birth to new individuals at rate β. e.g. bacterial colony, cancer cells.

If population is of size x at time t then overall birth rate is βx

I: P(exactly 1 birth in short interval of length $\delta t) = \beta x \delta t + o(\delta t)$

II: P(2 or more births in interval $\delta t) = o(\delta t)$

III: Any individuals probability of giving birth after t is independent of events before t.

Might be interested in:

- $X(t) =$ size of population at t
- $W_n =$ time to nth birth
- $T_n =$ time between $(n-1)$th and nth birth

For some processes (e.g. Poisson) the maths is simple. For others the maths is difficult/impossible, in which case we can resort to either simulation methods or adopt a deterministic approach (model average, large population).

3.2 Poisson Deterministic Model

$$\begin{aligned} \mathrm{P}(1 \text{ event in } [t, t+\delta t]) &= \lambda \delta t + o(\delta t) \\ \mathrm{P}(0 \text{ event in } [t, t+\delta t]) &= 1 - \lambda \delta t + o(\delta t) \end{aligned}$$

$$\begin{aligned} &\Rightarrow \text{Expected number of events in } [t, t+\delta t] \\ &= 1 \times (\lambda \delta t + o(\delta t)) + 0 \times (1 - \lambda \delta t + o(\delta t)) \\ &= \lambda \delta t + o(\delta t) \end{aligned}$$

Deterministic model: number of events in $[t + \delta t]$ is taken to be $\lambda \delta t + o(\delta t)$.
Let $D(t) =$ number events by t in deterministic model. Then

$$\begin{aligned} D(t+\delta t) &= D(t) + \lambda \delta t + o(\delta t) \\ \Rightarrow \frac{D(t+\delta t) - D(t)}{\delta t} &= \lambda + \frac{o(\delta t)}{\delta t} \\ \Rightarrow \frac{dD}{dt} &= \lambda \\ \Rightarrow D(t) &= \lambda t + c \end{aligned}$$

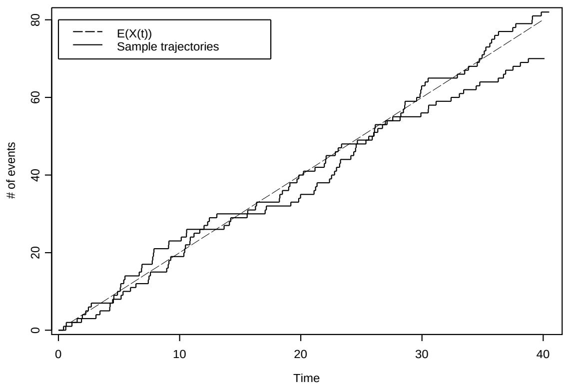

Figure 1: Sample trajectories vs. $E(X(t)) = \lambda t$, $\lambda = 2$

and since $D(0) = 0$

$$D(t) = \lambda t$$

Note: the trajectory of the average may have a different shape from any possible trajectory.

e.g. $X(t) \sim Poisson(\lambda t) \Rightarrow \mathrm{E}(X(t)) = \lambda t$ (see Figure 1).

Note: This is a general method for finding the deterministic solution.

In general:

$$\{X(t); t \geq 0\} \quad \text{is a stochastic process in continuous time}$$

$$\begin{aligned}
X(t) &= \text{number of events by } t \\
D(t) &= \text{number of events by } t \text{ in deterministic model} \\
\Rightarrow D(t+\delta t) &= D(t) + h(D,t)\delta t + o(\delta t) \\
\Rightarrow \frac{dD}{dt} &= h(D,t)
\end{aligned}$$

Same general approach applies in discrete case:

e.g. A simple branching process

Some cultures pass surname via male offspring only. Let the number sons a man has be a rv taking values $0, 1, 2, \ldots$ and assume each man reproduces independently of the others.

Start with 1 man (generation 0) and let X_n be the number in the nth generation ($x_0 = 1$).

In a deterministic approximation, assume each man produces s sons.

Then

$$\begin{aligned} x_n &= s x_{n-1} \\ &= s^2 x_{n-2} \\ &\vdots \\ &= s^n x_0 \\ &= s^n. \end{aligned}$$

3.3 Point Processes

Stochastic processes consisting of events occurring in time.

Stationary: distribution of number of events occurring in $(u, u+t]$ is the same as the distribution in $(0, t]$ for all $t > 0, u > 0$.

e.g: stationary – Poisson process

Some rvs of interest:

- number of events by time t, $X(t)$.

- time between $(n-1)$th and nth, T_n

- time to nth event, W_n

Multivariate point process: each event may be one of several types.

e.g. occurrence of death in a population is a point process, but can categorise according to cause of death.

e.g. Poisson process with rate λ, each event one of k types.

$$\mathrm{P}(\text{type } i) = p_i \qquad \sum p_i = 1.$$

Occurrence of each type independent of others.

Consider occurrence of type i events in $[t, t+\delta t]$.

1.

$$\begin{aligned}
\mathrm{P}(1 \text{ type } i \text{ event in } [t, t+\delta t]) &= \mathrm{P}(1 \text{ event in } [t, t+\delta t] \ \underline{\text{and}} \text{ it's type } i) \\
&= \mathrm{P}(1 \text{ in } [t, t+\delta t]) \times \mathrm{P}(\text{type } i \,|\, 1 \text{ event}) \\
&= (\lambda \delta t + o(\delta t)) \times p_i \\
&= \lambda p_i \delta t + o(\delta t).
\end{aligned}$$

2.

$$\begin{aligned}
\mathrm{P}(> 1 \text{ event in } [t, t+\delta t]) &= o(\delta t) \\
\mathrm{P}(> 1 \text{ event of type } i \text{ in } [t, t+\delta t]) &= o(\delta t)
\end{aligned}$$

3. Events in a Poisson process are independent and the different types are independent $\Rightarrow$ type i events are mutually independent and occurrence of type i event after t is independent of occurrence before t.

$\boxed{\text{Occurrence of type } i \text{ events is a Poisson process, rate } \lambda p_i}$

e.g. k independent Poisson processes, with rates $\lambda_1, \ldots, \lambda_k$, occur simultaneously and independently.

Consider events occurring in $[t, t+\delta t]$:

1.

$$\begin{aligned}
\mathrm{P}(1 \text{ event}) &= \sum_{i=1}^{k} \mathrm{P}(1 \text{ event of type } i \text{ and no other}) \\
&= \sum_{i=1}^{k} [\lambda_i \delta t + o(\delta t)] \prod_{j=1; j \neq i}^{k} [1 - \lambda_j \delta t + o(\delta t)] \\
&= \sum_{i=1}^{k} [\lambda_i \delta t + o(\delta t)] = (\sum \lambda_i) \delta t + o(\delta t).
\end{aligned}$$

2.

$$\begin{aligned}
\text{P}(>1 \text{ event}) &= 1 - \text{P}(1 \text{ event}) - \text{P}(0 \text{ event}) \\
\text{P}(0 \text{ event}) &= \prod_{i=1}^{k}[1 - \lambda_i \delta t + o(\delta t)] \\
&= 1 - (\sum \lambda_i)\delta t + o(\delta t) \\
\Rightarrow \text{P}(>1 \text{ event}) &= 1 - [(\sum \lambda_i)\delta t + o(\delta t)] - [1 - (\sum \lambda_i)\delta t + o(\delta t)] \\
&= o(\delta t)
\end{aligned}$$

3. since the processes are independent, the 3rd postulate is satisfied.

$\boxed{\text{Pooled process is a Poisson process with rate } \sum \lambda_i}$

Theorem (without proof) analogous to CLT

A superposition of k point processes in which 'no single process dominates the rest', is asymptotically a Poisson process.

If the k processes are Poisson, the result is exact (c.f. sum k normals).

$\Rightarrow$ huge importance of Poisson process.

3.4 Non-homogeneous Poisson process

- λ changes with time: $\lambda(t)$

Consider $X(t)$ - the number of events which have occurred by t.

$$p_n(t) = \text{P}(X(t) = n) \quad t > 0$$

$$p_o(0) = 1 \quad p_n(0) = 0,\ n = 1, 2, 3 \ldots \quad \text{(Init. Cond.)}$$

$$\begin{aligned}
p_n(t + \delta t) &= \text{P}(n \text{ events in } (0, t] \text{ and } 0 \text{ in } [t, t + \delta t]) \\
&\quad + \text{P}(n - 1 \text{ events in } (0, t] \text{ and } 1 \text{ in } [t, t + \delta t]) \\
&\quad + \text{P}(n - 2 \text{ events in } (0, t] \text{ and } 2 \text{ in } [t, t + \delta t]) \\
&\quad + \ \vdots \\
&\quad + \text{P}(0 \text{ events in } (0, t] \text{ and n in } [t, t + \delta t]) \\
&= p_n(t) \times [1 - \lambda(t)\delta t + o(\delta t)] \\
&\quad + p_{n-1}(t) \times [\lambda(t)\delta t + o(\delta t)]
\end{aligned}$$

$$
\begin{aligned}
& +p_{n-2}(t) \times o(\delta t) \\
& + \ \vdots \\
& +p_0(t) \times o(\delta t) \\
&= p_n(t) + (p_{n-1}(t) - p_n(t))\lambda(t)\delta t + o(\delta t) \\
\Rightarrow \frac{p_n(t+\delta t) - p_n(t)}{\delta t} &= (p_{n-1}(t) - p_n(t))\lambda(t) + \frac{o(\delta t)}{\delta t}
\end{aligned}
$$

As $\delta t \to 0$ we get

$$\frac{d}{dt}p_n(t) = (p_{n-1}(t) - p_n(t))\lambda(t) \quad n = 1, 2, \ldots$$

For $n = 0$, similarly

$$\frac{d}{dt}p_0(t) = -p_0(t)\lambda(t) \ \Rightarrow \ \frac{1}{p_0(t)}\frac{d}{dt}p_0(t) = -\lambda(t)$$

Define

$$\mu(t) = \int_0^t \lambda(u)\, du \quad t \geq 0 \ \ (\text{so } \mu(0) = 0)$$

Then,

$$\ln p_0(t) = -\mu(t) + c$$

Using $p_0(0) = 1, \mu(0) = 0 \Rightarrow c = 0$ so

$$p_0(t) = e^{-\mu(t)}$$

From this

$$
\begin{aligned}
\frac{d}{dt}p_1(t) &= (p_0(t) - p_1(t))\lambda(t) \\
&= e^{-\mu(t)}\lambda(t) - p_1(t)\lambda(t) \\
\frac{d}{dt}p_1(t) + p_1(t)\lambda(t) &= e^{-\mu(t)}\lambda(t)
\end{aligned}
$$

Multiply by integrating factor $e^{\mu(t)}$:

$$
\begin{aligned}
e^{\mu(t)}\frac{d}{dt}p_1(t) + p_1(t)\lambda(t)e^{\mu(t)} &= \lambda(t) \\
\frac{d}{dt}(p_1(t)e^{\mu(t)}) &= \lambda(t)
\end{aligned}
$$

$$\Rightarrow p_1(t)e^{\mu(t)} = \mu(t) + c$$

Using $p_1(0) = 0$ and $\mu(0) = 0$ gives $c = 0$, so

$$p_1(t) = e^{-\mu(t)}\mu(t)$$

By induction

$$p_n(t) = \frac{e^{-\mu(t)}[\mu(t)]^n}{n!}$$

That is: The number of events occurring in a non-homogeneous Poisson process with rate $\lambda(t)$ during the interval $(0, t]$ has a Poisson process with parameter

$$\mu(t) = \int_0^t \lambda(u)\, du \quad t \geq 0$$

Q: Given a non-homog. Poisson process with rate $\lambda(t)$, what is the distribution of the time to the first event?

A: Let T_1 = time to first event.

Then

$$\begin{aligned} \mathrm{P}(T_1 > t) &= \mathrm{P}(0 \text{ events in } (0, t]) \\ &= p_0(t) \\ &= e^{-\mu(t)} \end{aligned}$$

So,

$$\begin{aligned} F(t) &= 1 - \mathrm{P}(T_1 > t) = 1 - e^{-\mu(t)} \\ f(t) &= \lambda(t) e^{-\mu(t)} \end{aligned}$$

In general:

The number of events occurring in a non-homogeneous Poisson process with rate $\lambda(t)$ during any interval $(t_1, t_2]$ $0 \leq t_1 < t_2$, has a Poisson distribution with parameter

$$\mu(t_1, t_2) = \int_{t_1}^{t_2} \lambda(t)\, dt = \mu(t_2) - \mu(t_1).$$

The time until the first event after t_1 has pdf

$$f(t) = \lambda(t) e^{-\mu(t_1, t)} \quad t > t_1.$$

3.5 Compound Poisson Process

Events arrive according to a Poisson process, $X(t)$.

Each event is associated with another event, for which Y_i occurrences occur, Y_i iid.

e.g. Cars arriving at airport: number occupants.

e.g. Earthquakes: no. killed.

e.g. Insurance claims: amount claimed.

Let Poisson process have rate λ and $\mathrm{E}(Y_i) = \mu$, $\mathrm{var}(Y_i) = \sigma^2$.
Let $S(t) =$ no. occurrences by t, then

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

Difficult to find distribution of $S(t)$ because $X(t)$ is a rv. So, let's look at mean and variance. (Later use pgfs).

1.

$$\begin{aligned}
\mathrm{E}(S(t)) &= \sum_{s=0}^{\infty} s\mathrm{P}(S(t) = s) \\
&= \sum_{s} s \sum_{x=0}^{\infty} \mathrm{P}(S(t) = s \mid X(t) = x)\mathrm{P}(X(t) = x) \\
&= \sum_{x} \mathrm{P}(X(t) = x) \sum_{s} s\mathrm{P}(S(t) = s \mid X(t) = x) \\
&= \sum_{x} \mathrm{P}(X(t) = x) x\mu \\
&= \mu\lambda t \qquad (\text{since } \mathrm{E}(X(t)) = \lambda t)
\end{aligned}$$

$\longrightarrow$ A general method $\mathrm{E}(X) = \mathrm{E}_Y[\mathrm{E}(X \mid Y)]$.

2.

$$\mathrm{var}(S(t)) = \lambda t(\sigma^2 + \mu^2) \quad \text{Problem sheet.}$$

3.6 Doubly-stochastic Poisson process

$\lambda(t)$ is a rv.

e.g. I use a particular piece of software sporadically. When I do, I refer to the manual at times which follow a Poisson process with rate λ.

So the rate of referral is randomly 0 or λ.

3.7 General Point Processes: Notation

Mean time function of a point process is

$$\mu(t) = \mathrm{E}(X(t)).$$

Variance time function

$$\sigma^2(t) = \mathrm{var}(X(t)).$$

Index of dispersion

$$I(t) = \frac{\sigma^2(t)}{\mu(t)}$$

e.g. Poisson process:

$$\mu(t) = \lambda t; \quad \sigma^2(t) = \lambda t; \quad I(t) = 1.$$

So: if events occur with mean time function λt, but more regularly that Poisson, $I(t) < 1$.

e.g. Compound Poisson process

$$\mu(t) = \mu\lambda t; \quad \sigma^2(t) = \lambda t(\sigma^2 + \mu^2); \quad I(t) = \mu + \frac{\sigma^2}{\mu}.$$

Let Y be the number of occurrences at each compound event, then

$$\begin{aligned} I(t) &= \mu + \frac{\sigma^2}{\mu} = \frac{\mu^2 + \sigma^2}{\mu} = \frac{\mathrm{E}(Y^2)}{\mathrm{E}(Y)} \\ &= \frac{\sum y^2 \mathrm{P}(y)}{\sum y \mathrm{P}(y)} \\ &= \frac{\sum y^2 \mathrm{P}(y) - \sum y \mathrm{P}(y) + \sum y \mathrm{P}(y)}{\sum y \mathrm{P}(y)} \\ &= 1 + \frac{\sum y(y-1)\mathrm{P}(y)}{\sum y \mathrm{P}(y)} \end{aligned}$$

and this is > 1 unless $\mathrm{P}(y) = 0$ for $y > 1$.

That is:

the index of dispersion of a compound Poisson process is greater that 1 unless Y has a Bernoulli distribution.

3.8 The autocorrelation function

The correlation between the number of events occurring in the intervals $(0, t]$ and $(kt, (k+1)t]$.

Denote the number of events in $(kt, (k+1)t]$ by $X(kt, (k+1)t)$, $k = 0, 1, \ldots$.

For **stationary** processes

$$\text{var}\,(X(kt, (k+1)t)) = \sigma^2(t) \quad \forall k$$

i.e. NOT a function of k. Also, the autocorrelation function of order k is

$$\rho_k(t) = \frac{\text{cov}\,(X(0,t), X(kt, (k+1)t))}{\sigma^2(t)} \qquad t > 0$$

For a Poisson process, $\rho_k(t) = 0$.

In fact, the autocorrelation function contains no information not already in the variance-time function:

e.g. Let $k = 0$. Then

$$\begin{aligned}
\sigma^2(2t) &= \text{var}\,(X(0, 2t)) \\
&= \text{var}\,(X(0,t) + X(t, 2t)) \\
&= \text{var}\,(X(0,t)) + \text{var}\,(X(t, 2t)) + 2\text{cov}\,(X(0,t), X(t, 2t)) \\
&= \sigma^2(t) + \sigma^2(t) + 2\rho_1(t)\sigma^2(t)
\end{aligned}$$

So that

$$\rho_1(t) = \frac{\sigma^2(2t)}{2\sigma^2(t)} - 1.$$

i.e. $\rho_1(t)$ is a function of variance-time function.