STAT 515 Course Outline

Stochastic Processes & Monte Carlo Methods

1st half: (pre spring break / midterm exam)

Condtl. prob. and expec. review

Markov chains mainly discrete time, discrete space

Poisson processes

~~Briefly (time permitting)~~ Contns time M.C. Birth-death processes.

2nd half:

Monte Carlo methods: basics

Importance sampling

Rejection sampling etc.

Markov chain Monte Carlo: M-H algorithm / Gibbs sampler etc.

Lots of computing (in "R")

Not a standard Stoch Proc. course: no queueing theory, renewal processes etc.

Useful for: techniques for calculating expec., prob.

Markov chain basics, learning about M.C. models

Monte Carlo methods: computing intractable expectations and MLEs / optimization via simulation-based techniques

Particularly useful for people fitting Bayesian models, but very useful across many areas of statistics / modeling.

INTRO

Conditional Probability, Condtl. Expec. : basics/review.

Condtl. prob.: For events E, F, probability of E given F, $P(E|F) = \dfrac{P(E \cap F)}{P(F)}$ if $P(F) > 0$,

i.e. prob. E and F happen divided by prob. F happens.

Discrete case: A r.v. X on Ω_X is discrete if it takes countably many values $\{x_1, x_2, \ldots\}$

Prob. mass fn. (pmf) for X, $f_X(x) = P(X = x)$ $\forall x \in \Omega_X$

Ω_X = range of X

Consider X, Y, both discrete r.v.s w/ joint pmf

$f_{XY}(x, y) = Pr(X = x, Y = y)$ $\quad f_{XY} : \mathbb{R}^2 \to \mathbb{R}$

then condtl pmf of X given $Y = y$ is

$$f_{X|Y}(X = x | Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{f_{XY}(x, y)}{f_Y(y)}$$

where $f_Y(y) = \sum_{\Omega_X} f_{XY}(x; y) = \sum_{\Omega_X} P(X = x, Y = y)$

, marginal pmf of Y.

Also, condtl. cumulative distr. fn, $F_{X|Y}(x|y) = P(X \leq x | Y = y)$

$= \sum_{a \leq x} f_{X|Y}(X = a | Y = y)$

E.g. 3.3 $X \perp\!\!\!\perp Y$ Poisson w/ means λ_1, λ_2

$Z = X+Y$

Find : $Pr(X=x \mid X+Y=n) = Pr(X=x \mid Z=n)$

$$= \frac{Pr(X=x,\ Z=n)}{Pr(Z=n)}$$

But, $Z \sim Poi(\lambda_1+\lambda_2)$ and $X=x \,\&\, Z=n \Leftrightarrow X=x \,\&\, Y=n-x$

$$\Rightarrow \frac{Pr(X=x)Pr(Y=n-x)}{\frac{(\lambda_1+\lambda_2)^n e^{-(\lambda_1+\lambda_2)}}{n!}} \quad \because X, Y \text{ are indep.}$$

$$= \frac{\frac{\lambda_1^x \cancel{e^{-\lambda_1}}}{x!} \frac{\lambda_2^{n-x} \cancel{e^{-\lambda_2}}}{(n-x)!}}{\frac{(\lambda_1+\lambda_2)^n \cancel{e^{-(\lambda_1+\lambda_2)}}}{n!}}$$

$$= \binom{n}{x} \frac{\lambda_1^x \lambda_2^{n-x}}{(\lambda_1+\lambda_2)^x (\lambda_1+\lambda_2)^{n-x}}$$

$$= \binom{n}{x} \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^x \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-x}$$

pmf. of $Bin\left(n, p=\frac{\lambda_1}{\lambda_1+\lambda_2}\right)$

$X \mid X+Y=n \sim Bin\left(n, \frac{\lambda_1}{\lambda_1+\lambda_2}\right)$

Continuous case: A r.v. X is cntns. if

$\exists$ a fn f_X s.t. $f_X(x) \geq 0 \quad \forall x$

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1 \quad \text{and},$$

for each $a \leq b$, $P(a < X < b) = \int_a^b f_X(x)\,dx$. ← 1/14/08

The function f_X is the prob. density fn. (pdf)

Cummulative density fn, $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$

and $f_X(x) = F_X'(x)$ [at all pts. x where F_X is differentiable.]

Note that $Pr(X=x) = 0 \quad \forall x$ ~~for almost all x~~

~~Then, condtl. pdf~~

Now consider X, Y both cntns r.v.'s w/

joint pdf $f_{XY}(x,y)$ $\qquad f_{XY}: \mathbb{R}^2 \to \mathbb{R}$

f_{XY} is a joint pdf of X, Y if for every $A \subset \mathbb{R}^2$

$$P((X,Y) \in A) = \iint_A f(x,y)\,dx\,dy.$$

Then condtl pdf $f_{X|Y}(x|y) = \dfrac{f_{XY}(x,y)}{f_Y(y)}$ if $f_Y(y) > 0$

where $f_Y(y) = \int_{\Omega_x} f_{XY}(x,y)\,dx$, marginal pdf of Y.

Also, condtl. cdf $F_{X|Y}(x|y) = \int_{-\infty}^{x} f_{X|Y}(x|y)\,dx$

<u>Expected value</u> of a r.v. X,

$$E(X) = \begin{cases} \sum_{\Omega_X} x \, P(X=x) & \text{if } X \text{ discrete} \\ \int_{\Omega_X} x \, f_X(x) \, dx & \text{if } X \text{ cntns} \end{cases}$$

Let $X \in \Omega_X$, $Y \in \Omega_Y$

<u>Condtl. expectation</u>, $E(X|Y=y) = \begin{cases} \sum_{\Omega_X} x \, f_{X|Y}(x|y) \\ \int_{\Omega_X} x \, f_{X|Y}(x|y) \, dx \end{cases}$

$E(X|Y)$ is the r.v. $g(Y)$ (it is a fn. of Y)

where $g(y) = E(X|Y=y)$ $\forall y \in \Omega_Y$

Notes: ① condtl pdf / pmf ~~just~~ has same properties as an ordinary pdf / pmf

② $f_{X|Y}(x|y)$ is a pdf for X for each value of y.

$\Rightarrow$ condtl. expec. $E(X|Y=y)$ is an ordinary expectation (at that fixed value of $Y=y$).

③ (In this case:) $\Pr(X \in A | Y=y) = \int_A f_{X|Y}(x|y)\,dx$ is not actually a condtl. prob. since $P(Y=y)=0$.

④ Joint pdf / pmf on a set of r.v.'s completely specifies all marginals and condtl. distr. of the r.v.'s.

eful Properties:

Linearity of expectations: $E\left\{\sum_{i=1}^{n} a_i X_i\right\} = \sum_{i=1}^{n} a_i E(X_i)$

for r.v. $X_1, \dots, X_n$ and constants $a_1, \dots, a_n$.

ll: $\operatorname{Cov}(X, Y) = E\{(X - EX)(Y - EY)\}$

$= EXY - EXEY$

Bilinearity of covariance: $\operatorname{Cov}\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{m} b_j Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j \operatorname{Cov}(X_i, Y_j)$

for r.v. $X_1, \dots X_n, Y_1, \dots, Y_m$, constants $a_1, \dots, a_n, b_1, \dots, b_m$.

Law of iterated expectations:

$$E\{E(X|Y)\} = EX$$

Application: calculate expectations (RHS) in stages:

(i) Find $g(y) = E_X(X|Y=y)$

(ii) Find $E_Y\{g(Y)\}$ * wtd avg. of $E(X|Y=y)$, weighted by $P(Y=y)$

Discrete: $EX = \sum_{y \in \Omega_Y} E(X|Y=y) f_Y(y)$

Continuous: $EX = \int_{\Omega_X} E(X|Y=y) f_Y(y)\, dy$

E.g. Insect lays Y eggs, each surviving w/ prob. p (indep. of one another).

$Y \sim \text{Poi}(\lambda)$.

What is the E(# surviving eggs)?

E.g. of a hierarchical model: complicated process modeled by a series of conditional specifications/models placed in a hierarchy.

Let X = # surviving eggs

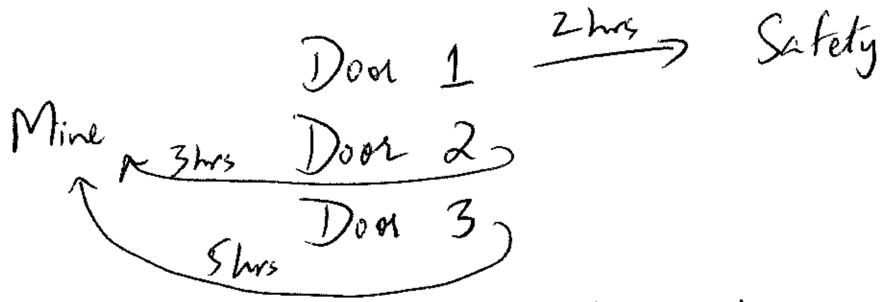
$$X \mid Y \sim \text{Binom}(Y, p)$$

$$\begin{aligned} E_X(X) &= E_Y\{E_{X|Y}(X \mid Y)\} \\ &= E_Y\{Yp\} \\ &= p\,E_Y\{Y\} = p\lambda. \end{aligned}$$

This is a natural situation to use L.I.E.

Often, we can use L.I.E. or, more generally, conditioning, as a tool/trick for simplifying calculation of expectations/probabilities.

E.g. 3.12 Miner trapped in mine w/ 3 doors



Miner chooses door w/ prob. 1/3 each.

T = time to safety.

Find $E(T)$. Use <u>1st step conditioning</u>

Let $X_i \in \{1,2,3\}$ = set of doors

X_i ↑ door chosen on i^{th} try

$$E(T \mid X_1 = 1) = 2$$
$$E(T \mid X_1 = 2) = 3 + E(T)$$
$$E(T \mid X_1 = 3) = 5 + E(T)$$

But, $E(T) = E(T \mid X_1 = 1) \times \frac{1}{3} + E(T \mid X_1 = 2) \times \frac{1}{3} + E(T \mid X_1 = 3) \times \frac{1}{3}$

$$= \frac{1}{3} 2 + \frac{1}{3}(3 + E(T)) + \frac{1}{3}(5 + E(T))$$

$$\Rightarrow E(T) = \frac{2}{3} + \frac{3}{3} + \frac{5}{3} + E(T)\frac{2}{3}$$

$$\Rightarrow E(T) = 3 \times \frac{10}{3} = 10$$

Conditioning is used as a trick to simplify calculations.

1/16/08

E.g. I.i.d. trials, success w/ prob p.
Repeat until k consecutive successes obtained.
N_k = # trials (" " " " ")

Find $E(N_k)$.

Recall: $E(N_1) = E(\text{geometric r.v.}) = \dfrac{1}{p}$

Define: $M_k = E(N_k) = E\big(E(N_k \mid N_{k-1})\big)$ by L.I.E.

Now, $E(N_k \mid N_{k-1}) = \underbrace{(N_{k-1}+1)}_{k^{th}\text{ successive success occurs immediately}}\, p + \underbrace{(N_{k-1}+1+E(N_k))}_{\text{start over } (\because\, k^{th}\text{ successive success did not occur})}(1-p)$

'last step conditioning'

$$\Rightarrow M_k = E(E(N_k \mid N_{k-1})) = pE(N_{k-1}) + p + \big(E(N_{k-1}) + 1 + E(N_k)\big)(1-p)$$

$$\Rightarrow M_k = \cancel{pE(N_{k-1})} + \cancel{p} + E(N_{k-1}) + 1 + E(N_k) - \cancel{pE(N_{k-1})} - \cancel{p} - pE(N_k)$$

$$\Rightarrow M_k = E(N_{k-1}) + 1 + (1-p)M_k$$

$$\Rightarrow pM_k = M_{k-1} + 1$$

$$\Rightarrow M_k = \frac{1}{p}(M_{k-1}+1)$$

We know, $M_1 = E(N_1) = \dfrac{1}{p}$

$M_2 = \dfrac{1}{p}\left(\dfrac{1}{p}+1\right) = \dfrac{1}{p^2} + \dfrac{1}{p}$

$M_3 = \dfrac{1}{p^3} + \dfrac{1}{p^2} + \dfrac{1}{p}$

$$M_k = \frac{1}{p^k} + \frac{1}{p^{k-1}} + \dots + \frac{1}{p}$$

Conditional Variance

Thm: $\mathrm{Var}X = E\{\mathrm{Var}(X|Y)\} + \mathrm{Var}\{E(X|Y)\}$

Pf: ~~First note that~~ $\cancel{\mathrm{Var}(X|Y) = E\{(X-E(X|Y))^2|Y\}}$

Use L.I.E. : $\mathrm{Var}\, X = E\{E(X^2|Y)\} - [E\{E(X|Y)\}]^2$

$$= E\{\mathrm{Var}(X|Y) + E(X|Y)^2\} - [E\{E(X|Y)\}]^2$$

$$= E\{\mathrm{Var}(X|Y)\} + \{E\{E(X|Y)^2\} - [E\,E(X|Y)]^2\}$$

$$= E\{\mathrm{Var}(X|Y)\} + \mathrm{Var}\,E(X|Y)$$

Computing variances by conditioning.

Two approaches:

1. Find EX^2 and $(EX)^2$ separately using L.I.E.
$\mathrm{Var}\, X = EX^2 - (EX)^2$

2. $\mathrm{Var}\, X = E\{\mathrm{Var}(X|Y)\} + \mathrm{Var}\{E(X|Y)\}$

E.g. Insect eggs (from before). $X|Y \sim \mathrm{Bin}(Y, p)$
$Y \sim \mathrm{Poi}(\lambda)$

What is $\mathrm{Var}\, X$?

$$\begin{aligned}
\mathrm{Var}(X) &= E_Y(\mathrm{Var}(X|Y)) + \mathrm{Var}_Y\{E(X|Y)\} \\
&= E_Y\{Yp(1-p)\} + \mathrm{Var}_Y\{Yp\} \\
&= p(1-p)\lambda + p^2\lambda \\
&= p\lambda - p^2\lambda + p^2\lambda = p\lambda
\end{aligned}$$

OR $\mathrm{Var}\, X = \cancel{EE(X|Y)}\; EX^2 - (EX)^2$

skip $EX = \lambda p$

$$\begin{aligned}
EX^2 &= E\{E(X^2|Y)\} \\
&= E_Y\{\mathrm{Var}(X|Y) + E(X|Y)^2\} \\
&= E_Y\{Yp(1-p) + Y^2p^2\} \\
&= p(1-p)\lambda + p^2 E\{Y^2\} \\
&= p(1-p)\lambda + p^2\{\mathrm{Var}\, Y + (EY)^2\} \\
&= p(1-p)\lambda + p^2\lambda + p^2\lambda^2 \\
&= p\lambda - \cancel{p^2\lambda} + \cancel{p^2\lambda} + p^2\lambda^2
\end{aligned}$$

$$\operatorname{Var} X = E X^2 - (EX)^2$$
$$= \lambda^2 p^2 + p\lambda - \lambda^2 p^2$$
$$= p\lambda$$

Computing probabilities by conditioning

equivalent to " expectations " "

$$\text{Prob. of event } E = E\{I(\text{event } E)\}$$

where $I(\text{event } E) = \begin{cases} 1 & \text{if } E \text{ occurs} \\ 0 & \text{else} \end{cases}$

or $1(E)$ or 1_E

Similarly, $E\{I \mid Y=y\} = Pr(E \mid Y=y)$

and $Pr(E) = \begin{cases} \sum Pr(E \mid Y=y) f_Y(y) & \text{discr.} \\ \int Pr(E \mid Y=y) f_Y(y)\,dy & \text{cntns.} \end{cases}$

E.g. $X \mid \lambda \sim Poi(\lambda)$

$\lambda \sim Gamma(2, 1)$ $\qquad f_\lambda(\lambda) = \lambda^{2-1} e^{-\lambda/1} I(\lambda > 0)$ $\qquad \boxed{\frac{1}{\Gamma(\alpha)\beta^\alpha}} x^{\alpha-1} e^{-x/\beta}$

Find $P(X=n)$

$$P(X=n) = \int_0^\infty P(X=n \mid \lambda)\, \lambda e^{-\lambda}\, d\lambda$$
$$= \frac{1}{n!}\int_0^\infty e^{-\lambda}\lambda^n (\lambda e^{-\lambda})\, d\lambda = \frac{1}{n!}\int_0^\infty \underbrace{\lambda^{n+1} e^{-2\lambda}}_{\propto\, G(n+2,\, 1/2)}\, d\lambda$$
$$= \frac{1}{n!}\Gamma(n+2)\left(\frac{1}{2}\right)^{n+2} = \frac{(n+1)!}{n!}\frac{1}{2^{n+2}}$$

useful integration trick: recognize known density

$$= \frac{n+1}{2^{n+2}}$$

Suggested reading Se[illegible] problem, Ballot problem

1/18/05

Bayesian Inference: a simple example

Suppose: $X_1, \dots, X_n \mid p \overset{\text{condtl. iid}}{\sim} \text{Ber}(p)$ Level 1

$p \sim \text{Unif}(0,1)$ Level 2

If we observe $S_n = \sum_{i=1}^{n} X_i$, how do we infer p?

Frequentist ('classical'): p is fixed though unknown (no Level 2)

Bayesian inference: p is a r.v. Level 2 = prior distr.

Note: if we have both levels, $X_1, \dots, X_n$ are not independent. They are conditionally indep.

Why: $P(X_1 = 1) = \int_0^1 P(X_1 = 1 \mid p)\, f(p)\, dp$

$$= \int_0^1 p \cdot 1\, dp = \left.\frac{p^2}{2}\right|_0^1 = \frac{1}{2}$$

Similarly, $P(X_2 = 1) = \frac{1}{2}$

$$P(X_1 = 1, X_2 = 1) = \int_0^1 P(X_1 = 1, X_2 = 1 \mid p)\, f(p)\, dp$$

$$\overset{\text{condtl. indep}}{=} \int_0^1 P(X_1 = 1 \mid p)\, P(X_2 = 1 \mid p) \cdot 1\, dp$$

$$= \int_0^1 p^2\, dp = \frac{1}{3}$$

Hence $P(X_1 = 1, X_2 = 1) \neq P(X_1 = 1)\, P(X_2 = 1)$

not indep.

Back to inference.

Frequentist: estimate of p, $\hat{p} = \dfrac{\sum_{i=1}^{n} x_i}{n}$ (MLE) $\left[\arg\max \mathcal{L}(p; x_1, \dots, x_n)\right]$

$S_n = \sum_{i=1}^{n} X_i = k \Rightarrow \hat{p} = \frac{k}{n}$

Bayesian: Distr. $f(p \mid S_n = k) = \dfrac{P(S_n = k,\ p)}{P(S_n = k)}$

$$= \underbrace{\frac{\binom{n}{k}}{P(S_n = k)}}_{\text{normalizing constant}} \underbrace{p^k (1-p)^{n-k}}_{\text{Beta density (w/o norm. constant)}} \cdot 1$$

$$= \text{Beta}(k+1,\ n-k+1)$$

Skip: Pt. estimate, $E(p \mid S_n = k) = \dfrac{k+1}{\cancel{k}+1+n-\cancel{k}+1} = \dfrac{k+1}{n+2}$

Prediction. What is $\Pr(X_{n+1} = 1 \mid S_n = k)$?

Frequentist: $\frac{k}{n}$

Bayesian: $\Pr(X_{n+1} = 1 \mid S_n = k) = \dfrac{P(X_{n+1} = 1, S_n = k)}{P(S_n = k)}$

$$P(S_n = k) = \int_0^1 \binom{n}{k} p^k (1-p)^{n-k} \cdot 1 \, dp$$

$$= \binom{n}{k} \int_0^1 \frac{\Gamma(k+1+n-k+1)}{\Gamma(k+1)\,\Gamma(n-k+1)} p^k (1-p)^{n-k} \, dp$$

$$\times \frac{\Gamma(k+1)\,\Gamma(n-k+1)}{\Gamma(\cancel{k}+1+n-\cancel{k}+1)}$$

$$= \frac{n!}{(n-k)!\,k!} \frac{k!\,(n-k)!}{(n+1)!} = \frac{1}{n+1}$$

$$P(X_{n+1}=1, S_n=k) = \cancel{P(X_{n+1}=1 \mid S_n=k)\, P(S_n=k)}$$

$$= \int_0^1 P(X_{n+1}=1, S_n=k \mid p) \cdot f(p)\,dp$$

condtl. indep.
$$= \int_0^1 P(X_{n+1}=1 \mid p)\, P(S_n=k \mid p) \cdot 1\,dp$$

$$= \int_0^1 p \binom{n}{k} p^k (1-p)^{n-k}\,dp$$

$$= \binom{n}{k} \int_0^1 p^{k+1} (1-p)^{n-k}\,dp$$

$$= \frac{n!}{(n-k)!\,k!} \frac{\Gamma(k+2)\,\Gamma(n-k+1)}{\Gamma(\cancel{k}+2+n-\cancel{k}+1)}$$

$$= \frac{n!}{\cancel{(n-k)!}\,k!} \frac{(k+1)!\,\cancel{(n-k)!}}{(n+2)!}$$

$$= \frac{k+1}{(n+1)(n+2)}$$

So, $$P(X_{n+1}=1 \mid S_n=k) = \frac{(k+1)}{\cancel{(n+1)}(n+2)} \times \cancel{(n+1)}$$

$$= \frac{k+1}{n+2}$$

Note: this = sample proportion w/ 1 success and 1 failure added (via the prior).

Think: answer when 0 successes observed in n trials.
" " n " " " " "

— 1/23/08

Above prediction for X_{n+1} (and inference for p) is different from frequentist prediction/inference

Stochastic Processes:

Defn.: A stochastic process is an indexed set of random variables, $\underset{\sim}{X} = \{X(t) : t \in T\}$ where

T : index set often thought of as a set of 'time' pts.

$X(t)$: state of process at time t.

When T is countable, usually $\{0,1,2,\ldots\}$ or $\{-2,-1,0,1,2,\ldots\}$ then $\underset{\sim}{X}$ is a discrete time process.

When T is an interval in $\mathbb{R}$, $\underset{\sim}{X}$ is a cntns-time process.

A realization of $\underset{\sim}{X}$ is called a <u>sample path</u>.

<u>State space</u> for the stochastic process: set of possible values for $X(t)$.

<u>Increments</u> of a stochastic process: $\{X_t : t \in T\}$ are $X_{t_1} - X_{t_0}$, $X_{t_2} - X_{t_1}$ etc. where $t_0 < t_1, \ldots < t_n \in T$.

Simple example: I.i.d. process

E.g. $X_1, X_2, \ldots \overset{iid}{\sim} Bin(n,p)$

$T = \{1,2,3,4,\ldots\}$: discrete time process

State space $= \Omega = \{0,1,\ldots,n\}$ discrete state space

$X_3 = 2 \Rightarrow$ process is in state 2 at time 3.

Non iid process

E.g. Bernoulli trials $X_1, X_2, \ldots \overset{iid}{\sim} Ber(p)$

New process: $Y_k = \sum_{t=1}^{k} X_t \sim Bin(k,p)$

But process is not ~~longer~~ iid.

It has iid increments $X_2 - X_1,\ X_3 - X_2, \ldots$ etc.

E.g. Simple random walk on integers

$$n-1 \xleftarrow{1-p} n \xrightarrow{p} n+1 \qquad p \in (0,1)$$

Define X_i: $Pr(X_i = 1) = p$

$Pr(X_i = -1) = 1-p$

for $i = 1, 2, \ldots$

Let $S_k = \sum_{i=1}^{k} X_i$

Sample paths: $\{1, 0, -1, -2, -1, -2, -3, -2, -1, 0, 1, 2, \ldots\}$

$\{-1, 0, 1, 2, 3, 2, 3, 4, \ldots\}$

More generally: Markov chains (prob. distr. of next state only depends on current state).

Important stochastic processes:

S. Ross Stoch. Proc. book

<u>Brownian motion</u>: A stochastic process $\{X_t \ t \geq 0\}$ is a Brownian motion process (or Wiener process) if

1. $X(0) = 0$
2. $\{X(t), t \geq 0\}$ has stationary independent increments.
3. For every $t > 0$, $X(t) \sim N(0, c^2 t)$.

Continuous time, continuous state space.

Botanist Robert Brown discovered it: motion exhibited by a small particle totally immersed in a liquid or gas.

Useful: stat. testing of goodness of fit, modeling price levels on stock markets, quantum mechanics etc.

<u>Independent increments</u>: Suppose $0 \leq t_0 < \ldots < t_n$ Then

$X_{t_1} - X_{t_0}$, $X_{t_2} - X_{t_1}$, ... are mutually independent r.v.'s

<u>Stationary increments</u>: Distr. of $X_{t_k} - X_{t_s}$ for any $t_k > t_s$ depends only on $t_k - t_s$, not on t_s.

· Continuous time, continuous state space.

~~Poisson process~~:

<u>Poisson process</u>: $\{N(t), t \geq 0\}$ is a Poisson process w/ rate λ if

1. $N(0) = 0$
2. Independent increments $N_{t_2} - N_{t_1}, N_{t_3} - N_{t_2}, \ldots$ are mutually indep.
3. If $s, t \geq 0$, then $\Pr\{N(s+t) - N(s) = n\} = \frac{e^{-\lambda t}(\lambda t)^n}{n!}$, $n = 0, 1, \ldots$

Note: For $s < t$, $N(t) - N(s) =$ # events that have occurred in $(s, t]$. (Counting process).

events in any interval is Poisson distr. w/ mean λt

Contin. time, discrete state space