

(Oct 28, Nov 25) 2 Tests No finals

Textbook: Probability & Random processes, 3rd Edition

by G. Grimmett & D. Stirzaker U.P.

One Thousand Exercises in Probability by same ppl.

Chapter 5 (some) Chapter 6. Chapter 8, topics from then on.

Assigned reading Sections 5.1 and 5.2 PRP.

Review the defn. of generating function

Generating function: The generating function (g_f) of a sequence a_0, a_1, \dots is $G_a(s) = \sum_{i=0}^{\infty} a_i s^i$ where $G(0) = a_0$.

Probability Generating function:

 X : nonnegative integer valued RV. with p.f. given by $f(i)$, $i=0, 1, 2, \dots$ Then $G_X(s) = \sum_{i=0}^{\infty} f(i) s^i$ P.g.f.or, $G_X(s) = E(s^X)$

Convolution: Given two sequences:

$$a = \{a_0, a_1, a_2, \dots\}$$

$$b = \{b_0, b_1, b_2, \dots\}$$

We define the convolution of a and b to be

$$c = \{c_0, c_1, c_2, \dots\} \text{ where}$$

$$c_n = \sum_{i=0}^n a_i b_{n-i}$$

We write $c = a * b$. Then

$$G_c(s) = G_a(s) \cdot G_b(s)$$

$$\text{or } G_{a * b}(s) = G_a(s) G_b(s)$$

$$\text{In particular, } G_{a * a}(s) = G_a(s)^2 \quad G_{a * a * a} = G_a(s)^3 \quad \dots \text{etc.}$$

Radius of convergence of g.f's.

Basic facts about P.g.f of X :

$$\textcircled{1} \quad P(X=n) = \frac{G_x^{(n)}(0)}{n!}$$

$$\textcircled{2} \quad E(X) = G'(1) \quad \boxed{\text{left derivative}}$$

$$= \lim_{s \uparrow 1} G'_x(s) \quad \text{left-derivative}$$

$$\textcircled{3} \quad [E[X(X-1)(X-2)\dots(X-k+1)]] = G^{(k)}(1)$$

$$= \lim_{s \uparrow 1} G^{(k)}(s)$$

Review some basic examples: constant RV's

Bernoulli RV, Binomial RV's, geometric Negative Binomial, Poisson.

Review moment generating function (m.g.f.)

$$M_X(t) = G_x(e^t) \text{ or more generally.}$$

$$M_X(t) = E(e^{tx})$$

(We will do the characteristic func. later in this chapter)

Review joint p.q.f. in (28) Definition on page 154.

Read (21) Theorem on page 153.

Do problem 5.1.1 on page 155.

Read (11) Example on page 158.

STAT 833 Sept 11, 2008.

Prob 5.1.2 page 155

Let X be ≥ 0 an integer valued random variable with P.g.f. G

Write $t(n) = P(X > n)$

For the tail probabilities of X , show that the gf of sequence t_1, t_2, \dots is given by

$$T(s) = \frac{1 - G(s)}{1 - s} \quad \text{show that } E(X) = T(1) \text{ and } \text{Var}(X) = 2T'(1) + T(1) - T(1)^2$$

s is in the interval of convergence of $T(s)$

$$\text{Solution: } T(s) = \sum_{n=0}^{\infty} t(n)s^n = \sum_{n=0}^{\infty} s^n P(X > n)$$

$$\text{think here } = \sum_{n=0}^{\infty} s^n \cdot E[\mathbb{1}_{(X>n)}]$$

$$= E \sum_{n=0}^{\infty} s^n \mathbb{1}_{(X>n)} = E \sum_{n=0}^{X-1} s^n = E\left(\frac{1-s^X}{1-s}\right)$$

$$= \frac{IE(1_{S^X})}{1-S} = \frac{1-G(S)}{1-S} \quad (|S| < 1)$$

$$(ii) T(1) = \frac{1-G(1)}{1-1} = \frac{0}{0} \quad \text{undefined}$$

Define T' to be the limit from below

$$T(1) := \lim_{\substack{s \nearrow 1 \\ s > 1}} T(s) = \lim_{s \nearrow 1} \frac{1-G(s)}{1-s} = \lim_{s \nearrow 1} G'(s) = EX$$

$$T'(1) = \lim_{s \nearrow 1} \frac{(1-s)(1-G(s))' - (1-G(s))(1-s)'}{(1-s)^2}$$

$$= \lim_{s \nearrow 1} \frac{(s-1)G'(s) - G(s) + 1}{(1-s)^2} = \lim_{s \nearrow 1} \frac{G(s) + (s-1)G''(s)}{f(s)} \quad (\text{L'Hopital's rule})$$

$$= \frac{1}{2} \lim_{s \nearrow 1} G'(s) = \frac{1}{2} IE[X(X-1)]$$

I'll leave it to you to finish the calculation.

(iii) Example page 158

Given a set of events A_1, A_2, \dots, A_n . Let X be the number of events which occur. Find the probability function of X .

The solution to $P(X=k)$ will be written in terms of probabilities of the form $P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_m})$ over all subsets of A_1, A_2, A_n

$$\text{Defn. } S_m = \sum_{1 \leq i_1 < i_2 < \dots < i_m} P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_m})$$

where the sum is over all subset of size $\{1, 2, \dots, n\}$

$$S_m = IE[\binom{X}{m}] = IE \frac{\binom{X}{m}(X-m)!}{m!} \quad \begin{matrix} \text{expectation, does not depend on } \\ \text{sum} \end{matrix}$$

$$\text{Proof: } S_m = \sum_{1 \leq i_1 < i_2 < \dots < i_m} P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_m})$$

$$= \sum_{1 \leq i_1 < i_2 < \dots < i_m} IE(1_{A_{i_1}} 1_{A_{i_2}} \dots 1_{A_{i_m}}) = IE \left(\sum_{1 \leq i_1 < i_2 < \dots < i_m} 1_{A_{i_1}} 1_{A_{i_2}} \dots 1_{A_{i_m}} \right)$$

$$= IE[\binom{X}{m}] \quad \text{GEP?} \quad \text{Hilfsy}$$

next. consider the (q_f) of the sequence

s_0, s_1, s_2, \dots where we shall define $s_0 = 1$

$G_S(x) = \sum_{m=0}^n x^m \cdot s_m$ we shall also need the (pqf) of x .

$$G_x(s) = \sum_{i=0}^n x^i P(x=i)$$

$$G_S(x) = \sum_{m=0}^n x^m s_m = \sum_{m=0}^n x^m E(m)$$

$$= \sum_{m=0}^n x^m \left[\sum_{i=0}^n \binom{i}{m} / P(x=i) \right]$$

$$= \sum_{i=0}^n P(x=i) \sum_{m=0}^n x^m \cdot \binom{i}{m} \cdot i^{-m}$$

$$= \sum_{i=0}^n P(x=i) (1+x)^i$$

$$= G_x(1+x)$$

$$\boxed{G_x(x) = G_S(x-1)}$$

two ways of expansion

$$S_0 \quad G_x(x) = P(x=0) + x P(x=1) + x^2 P(x=2) + \dots + x^n P(x=n)$$

$$G_S(x-1) = s_0 x^{-1} + s_1 (x-1) + (x-1)^2 s_2 + \dots + (x-1)^n s_n$$

$$= (s_0 - s_1 + s_2 - \dots + (-1)^n s_n) + x(s_1 - (\frac{1}{2})s_2 + (\frac{1}{2})s_3 + \dots + (-1)^{n+1} (\frac{1}{n+1}) s_n)$$

$$+ x^2 (\dots + \dots)$$

Equating the coefficients of the two expansions

of $G_x(x)$ we see that

$$P(x=0) = s_0 - s_1 + s_2 - \dots + (-1)^n s_n$$

$$P(x=1) = s_1 - (\frac{1}{2})s_2 + (\frac{1}{2})s_3 + \dots + (-1)^{n+1} (\frac{1}{n+1}) s_n$$

$$\boxed{P(x=i) = \sum_{j=0}^n (-1)^{j-1} \binom{j}{i} s_j \quad 0 \leq i \leq n}$$

Warming's Theorem.

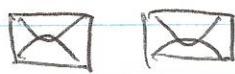
Special Case:

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(X > 0) = 1 - P(X = 0)$$

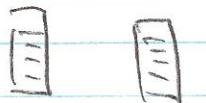
$$= 1 - [S_0 - S_1 + S_2 - \dots + (-1)^n S_n]$$

$$= 1 - 1 + S_1 - S_2 - \dots + (-1)^{n+1} S_n$$

$$= \sum_{i=1}^n P(A_i) - \sum_{i,j} P(A_i \cap A_j) + \dots + (-1)^{n+1} P(A_1 \cap A_2 \cap \dots \cap A_n)$$



n envelopes



n letters

Letters are randomly assigned to envelopes.

A_i : the i th letter is assigned to i th letter. (event)

Read Section 5.3 Random Walks.

Sept 16, 2008 STAT 833

Assigned problems 5.1: 1, 2, 3, 5, 6. (use generating func) 9.

5.2: 1, 2, 4, 5, 8 (use g.f.'s) 5.3: 1, 2, 3, 4

5.3 Random Walks

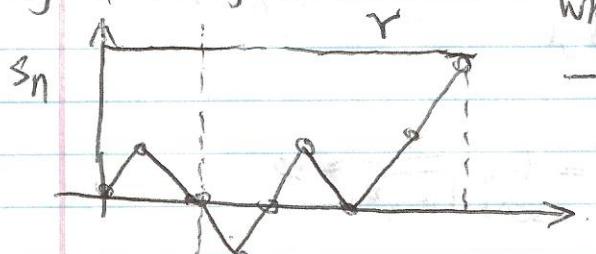
X_1, X_2, X_3, \dots independent with $X_i = \begin{cases} 1 & \text{with prob. } p \\ -1 & \text{with prob. } 1-p \end{cases}$

The random walk is defined as S_n , $n \geq 0$, such that

$$S_0 = 0 \quad S_1 = X_1, \quad S_2 = X_1 + X_2, \quad \dots \quad S_n = X_1 + X_2 + \dots + X_n = \sum_{i=1}^n X_i$$

Then S_n is a simple random walk starting at zero. When $p = \frac{1}{2}$, we

say S_n is symmetric.



What is distribution of S_n for fixed n ?

- fix the value of r , ask for the distribution of the first time to hit r

Hilary

Define $P_c(n) = P(S_n = 0)$

$$f_0(n) = P(S_1 \neq 0, S_2 \neq 0, \dots, S_{n-1} \neq 0, S_n = 0)$$

$P_c(n)$ is the probability of returning to zero at time n .

$f_0(n)$ is the probability of the first return to zero at time n .

We set $P_c(0) = 1$ and $f_0(0) = 0$.

The g.f' are

$$P_0(s) = \sum_{n=0}^{\infty} P_c(n) s^n \quad F_0(s) = \sum_{n=0}^{\infty} f_0(n) s^n$$

convolution

Theorem 1: ① $P_0(s) = 1 + P_0(s) F_0(s)$

② $P_0(s) = \frac{1}{\sqrt{1-4pq s^2}}$ where $q=1-p$

③ $F_0(s) = 1 - \sqrt{1-4pq s^2}$

Proof: Page 163.

(4) Corollary: (a) The probability of the random walk never returns to zero is $1 - |p-q|$, so when $p=q=1/2$, this prob. is one.

(b) When this prob is one (i.e., $p=q=1$) the expected time to return is infinite

proof (a) $P(\text{eventual return to zero})$

$$= P(S_1 \neq 0) + P(S_1 \neq 0, S_2 \neq 0) + P(S_1 \neq 0, S_2 \neq 0, S_3 \neq 0) + \dots$$

$$= f_0(1) + f_0(2) + f_0(3) + \dots$$

$$= f_0(1) s + f_0(2) s^2 + f_0(3) s^3 + \dots \Big|_{s=1} = F_0(s) \Big|_{s=1}$$

$$= 1 - \sqrt{1-4pq} = 1 - \sqrt{(p-q)^2} = 1 - |p-q|$$

when $p=q$ $F_0(1)=1$

(b) If $p=q$, so that $F_0(1)=1$, then $F_0(s)$ is the p.p.f. of the random variable $N = \min \{ n \geq 1 : S_n = 0 \}$

so $P(N < \infty) = 1$. Then

$$E(N) = \lim_{s \rightarrow 1} F_0'(s) = \lim_{s \rightarrow 1} (1 - \sqrt{1-s^2})' = \lim_{s \rightarrow 1} \frac{s}{\sqrt{1-s^2}} = \infty$$

We call the random walk S_n persistent (or recurrent) if the probability of return to zero are 1 (i.e., $P_0(1) = 1$ here) if not, S_n is called transient.

More generally,

$$f_{r,n} = P(\text{first visit to } r \text{ at time } n), r \neq 0$$

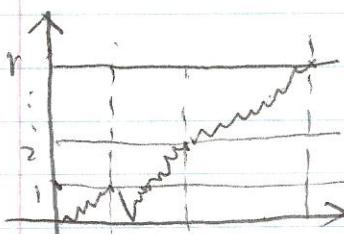
$$= P(S_1 + r, S_2 + r, \dots, S_{n-1} + r, S_n = r)$$

and define

$$F_r(s) := \sum_{n=1}^{\infty} f_{r,n} s^n$$

Theorem (5) (a) $F_r(s) = [F_1(s)]^r$ for all $r \geq 1$

$$(b) F_1(s) = \frac{1 - \sqrt{1 - 4ps^2}}{2qs}$$



To calculate $F_r(s)$ when $r > 0$, we use the idea of reflection: reversing the roles of p and q .

$$F_{-1}(s) = \frac{1 - \sqrt{1 - 4ps^2}}{2qs}$$

$$F_{-r}(s) = [F_{-1}(s)]^r, \text{ for all } r \geq 1$$

Read (7) The hitting time theorem. You are responsible for the statements, but not the proof. We shall not cover topics in the section after that.

(pp 166 — top of 170)

Problem 52 # 4. Let X have a binomial distribution with parameters n and p . Show that

$$E\left(\frac{1}{1+X}\right) = \frac{1 - (1-p)^{n+1}}{(n+1)p}$$

Solution: First attempt

$$\sum_{x=0}^n \frac{1}{1+x} \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=0}^n \frac{1}{1+x} \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^n \frac{n!}{(x+1)!(n-x)!} p^x (1-p)^{n-x}$$

$$E = \frac{1}{(n+1)p} \sum_{y=1}^{n+1} \binom{n+1}{y} p^y (1-p)^{n+1-y}$$

$$= \frac{1}{(n+1)p} [1 - (1-p)^{n+1}]$$

attempt 2: $E\left(\frac{1}{1+x}\right) = E(1-x+x^2-x^3+x^4\dots)$

- ① doesn't converge ??
- ② $1 - EX + EX^2 - EX^3 \dots$

attempt 3: $\frac{1}{1+x} = \int_0^1 t^x dt$

$\therefore E\left(\frac{1}{1+x}\right) = E \int_0^1 t^x dt = \int_0^1 E(t^x) dt$

 $= \int_0^1 [(1-p) + pt]^n dt$
 $u = (1-p) + pt \quad du = pdt$
 $\Rightarrow \int_{1-p}^1 u^n \frac{du}{p} = \frac{u^{n+1}}{(n+1)p} \Big|_{1-p}^1 = \frac{1 - (1-p)^{n+1}}{(n+1)p}$

Sept 18th, 2008 STAT 833 Thursday

Read Section 5-4 We will skip 5-5 and go to 5-6

Assigned problems for 5-4 1, 3, 4, 5.

Branching processes:

suppose $Z_n, n \geq 0$ represents the size of a population in generation n .

for simplicity, we set $Z_0 = 1$ with probability one. To obtain Z_{n+1} , the size of generation $n+1$, we suppose that each individual in generation n gives rise to $X_{n,i}$ offsprings (and then dies). So

$$Z_{n+1} = X_{n,1} + X_{n,2} + \dots + X_{n,n} \geq n$$

We assume that, conditionally on Z_n ,

- ① $X_{n,1}, X_{n,2}, \dots, X_{n,n}$ are independent
- ② $X_{n,1}, X_{n,2}, \dots, X_{n,n}$ are identically distributed.

pgf $G(s)$

(1) Theorem

Let Z_n have pgf $G_n(s)$, then $(q-1)^n q^{\sum_{i=1}^n X_{n,i}}$

$$G_n(s) = G(s)$$

$$q^{\sum_{i=1}^n X_{n,i}} (q-1)^n s^{\sum_{i=1}^n X_{n,i}} =$$

$$G_2(s) = G(G(s))$$

$$G_3(s) = G(G(G(s)))$$

$$G_n(s) = \underbrace{G(\dots G}_{n\text{-times}}(s)\dots)$$

Proof: We prove by induction on n .

case $n=1$ $Z_1 = X_{0,1}$ so $G_1(s) = f(s)$

Induction step: Suppose this is true for n , we shall prove for $n+1$

$$G_{n+1}(s) = E(s^{Z_{n+1}}) = E[E(s^{Z_{n+1}} | Z_n)]$$

$$= E(E(s^{X_{n,1} + \dots + X_{n,n}} | Z_n))$$

$$= E[\underbrace{G(s)^{Z_n}}_{\text{Expected value where } t=G(s)}] = E[t^{Z_n}] = G_n(t) \text{ where } t = G(s)$$

$$= \underbrace{G(G\dots G(t)\dots)}_{n\text{-times}} \text{ where } t = G(s)$$

(by induction hypothesis)

$$= \underbrace{G(G\dots G(G(s))\dots)}_{(n+1)\text{-times}}$$

This completes the induction step. QED.

so $G_{m+n}(s) = G_m(G_n(s))$ for all $m, n \geq 1$

(2) Lemma Let $\mu = E(Z_1)$ that is $\mu = G'(1)$ let $\sigma^2 = \text{Var}(Z_1)$
Then $E(Z_n) = \mu^n$, $\text{Var}(Z_n) = \begin{cases} n\sigma^2 & \mu=1 \\ \frac{\sigma^2(\mu^n-1)}{\mu-1}\mu^{n-1} & \mu \neq 1 \end{cases}$

Proof, Page 172.

The population becomes extinct in generation n , if $Z_n = 0$
after that $Z_{n+1} = Z_{n+2} = \dots = 0$

So ultimate extinction $\eta = \bigcup_{n=1}^{\infty} \{Z_n = 0\}$

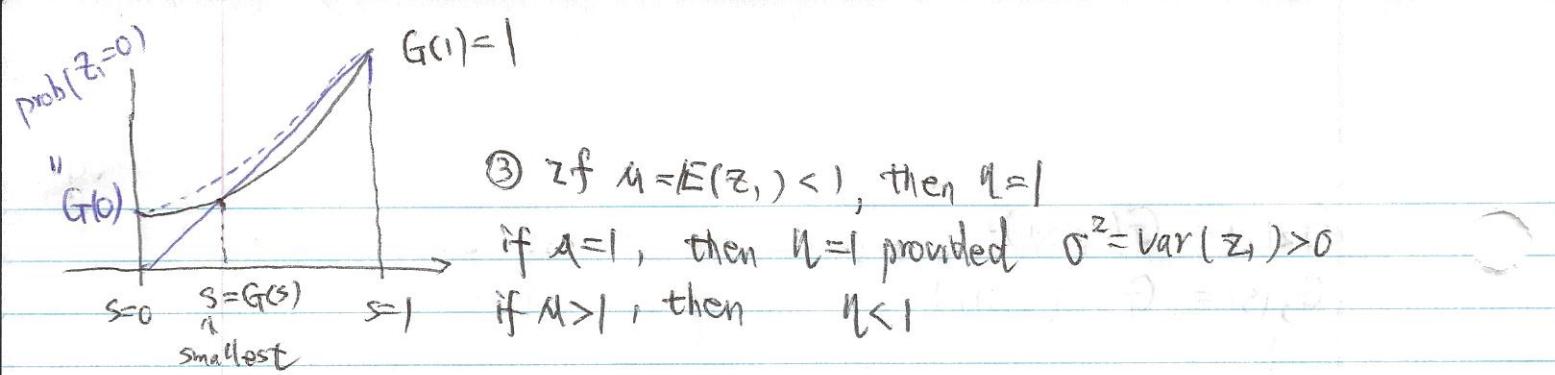
(3) Theorem ① as $n \rightarrow \infty$, $P(Z_n < 0) \rightarrow P(\text{ultimate extinction})$

② If we set $P(\text{ultimate extinction}) = \eta$, then η is the
smallest positive root of the equation

nonnegative

$$S = G(S)$$

Hilary



③ If $\mu = E(z_1) < 1$, then $\eta = 1$

if $\mu = 1$, then $\eta = 1$ provided $\sigma^2 = \text{Var}(z_1) > 0$

if $\mu > 1$, then $\eta < 1$

proof ① $\{z_1=0\} \subset \{z_2=0\} \subset \{z_3=0\} \subset \dots$

$$\text{So. } P(\text{ultimate extinction}) = P\left(\bigcup_{n=1}^{\infty} \{z_n=0\}\right) = \lim_{n \rightarrow \infty} P(z_n=0)$$

Cf. STAT 901.

② Let $\eta_n = P(z_n=0) \quad n \geq 1$

$$\text{Then } \eta_n = P(z_n=0) = G_n(0) = G(F_{n-1}(0))$$

$$= G(\eta_{n-1}) \quad G \text{ is continuous}$$

$$\text{So } \lim_{n \rightarrow \infty} \eta_n = \lim_{n \rightarrow \infty} G(\eta_{n-1}) = G\left(\lim_{n \rightarrow \infty} \eta_{n-1}\right)$$

$$= G(\lim_{n \rightarrow \infty} \eta_n) \quad \text{or} \quad \eta = G(\eta) \quad (\text{fixed pts for } G(s))$$

So η is a root of $G(s) = s$ on page 173, a proof given that this is the smallest non-neg. root.

③ see page 173, also QED.

$$\text{Surprisingly, } P(z_n \rightarrow \infty | z_1 \neq 0) = 1, (\sigma^2 > 0)$$

Sept 23 2008 STAT 833.

Section 5.6 Expectation revisited

Read

$$E(X) = \sum x f(x) \quad \text{where } f(x) \text{ is the pf of a discrete RV. } X.$$

$$E(X) = \int x f(x) dx \quad \text{where } f(x) \text{ is the pdf of a continuous RV } X.$$

We need a unified notation.

We define jump of F at x .

$$dF(x) = \begin{cases} F(x) - \lim_{y \rightarrow x^-} F(y) & F \text{ discrete df. for discrete RV.} \end{cases}$$

$$\left[\frac{dF(x)}{dx} \right] dx. \quad F \text{ df for continuous RV.}$$

$$\int = \begin{cases} \sum & \text{discrete} \\ \int & \text{continuous} \end{cases}$$

Then $E(X) = \int x dF(x)$ in both cases.

This terminology is justified as a Riemann-Stieltjes integral in the general case.

integrating function

$$\int g(x) dF(x) = \lim_{n \rightarrow \infty} \sum_{i=1}^n g(x_i) [F(x_{i+1}) - F(x_i)]$$

integrand $\sup |x_{i+1} - x_i| > 0$

Another shifted approach is to think of $E(X)$ as an integral over the sample space Ω .

$$E(X) = \int_{\Omega} X(\omega) dP(\omega)$$

integrand Lebesgue integral
 integrating prob. measure

Important properties of $E(X)$

① Monotone Convergence Theorem:

Suppose $P(X_n > x) = 1$ for all $n = 1, 2, 3, \dots$. $P(X_n \nearrow x) = 1$

Then $(E(X_n)) \nearrow E(X)$, which may be ∞ .

$\boxed{\lim E \neq E(\lim \text{ in general})}$

② Dominated Convergence theorem

Suppose $P(X_n > x) = 1$ and there exists a random variable Y , such that

$P(|X_n| \leq r) = 1$ for all n , and $(E(Y) < \infty)$

then $(E(X_n)) \rightarrow (E(X))$, which is finite.

$f_n(x) = \frac{1}{n}$ for all x .

$f(x) = 0$ for all x

$f_n(x) \rightarrow f(x)$ \Rightarrow dominated

$$\int_{-\infty}^{+\infty} f_n(x) dx \rightarrow \int_{-\infty}^{+\infty} f(x) dx = 0$$

$|f_n(x)| \leq g(x)$

$$f_n(x) = 1 - \frac{1}{n} \quad \text{for all } x$$

$$f(x) = 1 \quad \text{for all } x$$

$$\int_{-\infty}^{+\infty} f_n(x) dx \rightarrow \int_{-\infty}^{+\infty} f(x) dx = \infty$$

Assigned problems 5.6 #2 #4

5.7 characteristic functions. read section 5.7

Defn. The moment generating function (MGF) of X is

$$M(t) = E(e^{tX}) = \int e^{tx} dF(x)$$

If $M(t)$ is defined in some open interval about $t=c$, then it has a power series representation.

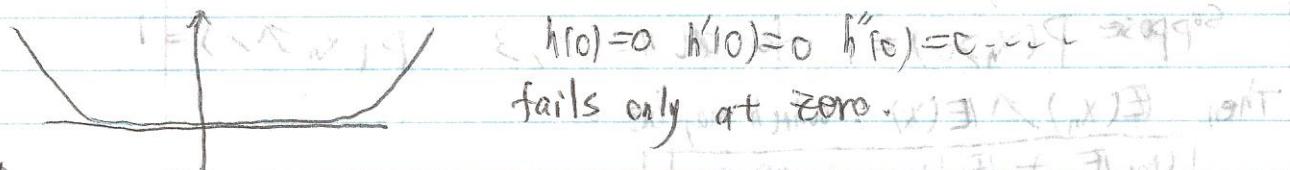
$$M(t) = 1 + E(X)t + \frac{E(X^2)t^2}{2!} + \frac{E(X^3)t^3}{3!} + \dots$$

Taylor Expansion of $M(t)$

$h(t) = it$ does not have a power series expansion in t .

$$h(t) = \begin{cases} e^{t/2} & t \neq 0 \\ 0 & t=0 \end{cases}$$

very smooth defined for all t except $t=0$
very flat at $t=0$



Brownian Motion is Not differentiable anywhere,

No Taylor expansions.

Let X have density $f(x) = \frac{1}{\pi} \frac{1}{1+x^2}$

$$f(x) = \frac{1}{\pi} \frac{1}{1+x^2} \quad -\infty < x < \infty \quad (\text{Cauchy density})$$

(Here, $E(X)$ does not exist)

$$M(t) = \int 1 \quad t=0$$

undefined $t \neq 0$

Define for $z = \sqrt{-t}$,

$$e^{iz} = \cos u + i \sin u \quad (\text{complex exponential})$$

$$e^{i\pi} = 0 \quad (\text{DeMoivre's theorem form})$$

$$\text{Properties: } e^{i(u+v)} = e^{iu} e^{iv}$$

$$e^{i(u+v)} = \cos(u+v) + i \sin(u+v)$$

$$= [\cos u \cos v - \sin u \sin v]$$

$$+ i[\sin u \cos v + \cos u \sin v]$$

$$e^{iu} \cdot e^{iv} = [\cos u + i \sin u] [\cos v + i \sin v]$$

$$= (\cos u \cos v - \sin u \sin v) + i[\sin u \cos v + \cos u \sin v]$$

$$\Re(e^{iu}) = 1 \quad \text{for all } u \in \mathbb{R}$$

Modulus

$$|x+iy|^2 = x^2+y^2 \quad \text{for } x, y \text{ real (imaginary parts)}$$

$$|\mathrm{e}^{iu}|^2 = |\cos u + i \sin u|^2 = \cos^2 u + \sin^2 u = 1$$

Definition: We define the characteristic function (ϕ) of X to be

$$\phi(t) = \mathbb{E}(e^{itX}) = \mathbb{E}(\cos tX) + \mathbb{E}(\sin tX) \cdot i \quad (-\infty < t < \infty)$$

Proposition: $|\phi(t)| \leq 1$ for all real t .

$$\text{Proof: } |\phi(t)| = |\mathbb{E}(e^{itX})| = \left| \int e^{itx} dF(x) \right|$$

$$\leq \int |e^{itx}| dF(x) = \int dF(x) = 1$$

Sept 25, 2008

$$X \text{ random variable. } \phi(t) = \mathbb{E}(e^{itX}) = \int e^{itx} dF(x)$$

Example: $X \sim \text{Exp}(1)$

$$f(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

$$\begin{aligned} \text{Then } \phi(t) &= \mathbb{E}(e^{itX}) = \int_{-\infty}^{+\infty} e^{itx} dF(x) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx \\ &= \int_{-\infty}^{+\infty} e^{itx} e^{-x} dx = \int_0^{+\infty} e^{(it-1)x} dx = \frac{e^{(it-1)x}}{it-1} \Big|_0^{+\infty} \\ &= \lim_{x \rightarrow \infty} \frac{e^{(it-1)x}}{it-1} - \frac{e^{(it-1)0}}{it-1} = \lim_{x \rightarrow \infty} \frac{e^{(it-1)x}}{(it-1)} - \frac{1}{it-1} \\ &= \lim_{x \rightarrow \infty} \frac{e^{itx} \cdot e^{-x}}{it-1} - \frac{1}{it-1} = \frac{1}{-it+1} = \frac{1}{1-it} \quad (-\infty < t < +\infty) \end{aligned}$$

On the other hand,

$$M(t) = \frac{1}{1-t} \quad \text{defined for } |t| < 1$$

In fact, $\phi(t) = M(it)$ where both are defined

about $t=0$.

Example: $X \sim N(\mu, \sigma^2)$

Here, $M(t)$ is defined for all real t . we have

$$M(t) = e^{\mu t + \sigma^2 t^2 / 2}$$

Hilary

$$\text{so } \phi(t) = M(it) = e^{iat} - \frac{\sigma^2}{2} t^2$$

If $M(t)$ is defined in an open interval about $t=0$, then $\phi(t)$ has a power series expansion.

$$\phi(t) = 1 + \frac{iE(x)(it)}{1!} + \frac{iE(x^2)(it)^2}{2!} + \frac{iE(x^3)(it)^3}{3!} + \dots$$

$$= 1 + iE(x)t - \frac{iE(x^2)t^2}{2!} - \frac{iE(x^3)t^3}{3!} + \frac{iE(x^4)t^4}{4!} + \dots$$

However when $M(t)$ is only defined at $t=c$, then

$\phi(t)$ does not have a power series expansion.

Proposition: If $Y = aX+b$, then

$$\phi_Y(t) = e^{ibt} \phi_X(at)$$

Proof: Basically the same as mgf case.

$$\phi_Y(t) = E(e^{ity}) = E(e^{it(ax+b)}) = e^{ibt} E(e^{itaX}) = e^{ibt} \phi_X(at)$$

Proposition: If $Y = X_1 + X_2 + \dots + X_n$, where X_1, X_2, \dots, X_n are independent QED.

$$\text{then } \phi_Y(t) = \phi_{X_1}(t) \phi_{X_2}(t) \dots \phi_{X_n}(t)$$

Proof: Basically the same as mgf case. QED.

Special case

X_1, X_2, \dots, X_n are i.i.d. with characteristic func. $\phi(t)$

$$\text{Then } \phi_Y(t) = \phi_X^n(t)$$

Definition We can define the joint cf. of X_1, \dots, X_n to be

$$\phi(t_1, t_2, \dots, t_n) = E(e^{it_1 X_1 + it_2 X_2 + \dots + it_n X_n})$$

$$= E(e^{i \sum_{j=1}^n t_j X_j}) \text{ for all } -\infty < t_1, \dots, t_n < \infty$$

Proposition: X_1, \dots, X_n are independent if and only if

$$\phi(t_1, t_2, \dots, t_n) = \phi_1(t_1) \phi_2(t_2) \dots \phi_n(t_n)$$

where X_j has cf $\phi_j(t)$

Proof: Basically the same as mgf.

Problems Section 3.7, 3, 4, 5, 7

Read 5.8

Examples of characteristic functions

① Bernoulli distribution $\text{Ber}(p)$

$$X = \begin{cases} 1 & \text{Prob. } p \\ 0 & \text{Prob. } 1-p=q \end{cases}$$

$$\Phi(t) = E(e^{itX}) = pe^{it-1} + qe^{it0} = q + pe^{\frac{it}{1-p}} = q + pe^{it}$$

② Binomial (n, p)

$X = Y_1 + Y_2 + \dots + Y_n$ where Y_1, \dots, Y_n are independent Bernoulli ($\text{Ber}(p)$)

$$\text{So } \phi_X(t) = [\phi_Y(t)]^n = [q + pe^{it}]^n$$

③ Exp (λ)

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

$$\Phi(t) = \frac{\lambda}{\lambda - it} \quad -\infty < t < \infty$$

④ Cauchy distribution

$$f(x) = \frac{1}{\pi} \frac{1}{1+x^2} \quad -\infty < x < \infty$$

M(t) only exists for $t=0$

$$\Phi(t) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx = \frac{1}{\pi} \int_{-\infty}^{+\infty} e^{itx} / (1+x^2) dx$$

$= e^{-|t|}$ by contour integration

Application X_1, \dots, X_n independent Cauchy random variables

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

$$\Phi(\bar{X}) = \Phi_{X_1 + X_2 + \dots + X_n} \left(\frac{t}{n} \right) = [\Phi_X(\frac{t}{n})]^n = e^{-|t/n| \cdot n} = e^{-|t|}$$

so \bar{X} has a cauchy distribution

STABLE LAW

Application We have seen that when X and R are independent,

$$\Phi_{X+R}(t) = \Phi_X(t)\Phi_R(t) \text{ for all } t.$$

The converse is not TRUE

Example: let X be cauchy, $R = X$, $X+R = 2X$

$$\Phi_X(t)\Phi_R(t) = [\Phi_X(t)]^2 = e^{-2|t|} = \Phi_{2X}(t) = \Phi_{X+R}(t)$$

Example: Gamma(λ, α)

\uparrow
scale shape

$$f(x) = \begin{cases} \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

$$\Phi(t) = \left(\frac{\lambda}{\lambda - it} \right)^\alpha \quad -\infty < t < +\infty$$

A couple of quick questions:

$$\textcircled{1} \quad \Phi(0) = 1 \quad \text{also } |\Phi(t)| \leq 1$$

$$\textcircled{2} \quad \Phi(-t) = E(e^{-itx})$$

$$= E[(e^{itx})^*] \quad \text{complex conjugate}$$

$x - i\bar{x}$ is a linear function.

$$\Rightarrow \Phi(-t) = [E(e^{itx})]^* = \Phi(t)^*$$

For you to check: $(e^{itx})^* = e^{-itx}$

If x and $-x$ have the same distribution

then $\Phi(t)$ is real for all t .

Sept 30, 2008 STAT 833

Assigned problems 5.7 3, 4, 5, 7, 11.

5.8: 1, 2, 4, 5, 6, 8, 9

5.9 Fourier Inversion Theorem

Theorem 1) If x is a continuous random variable, with density function $f(x)$ and characteristic function $\Phi(t)$, then

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \Phi(t) dt$$

for every value x where $f'(x)$ exists.

Note: if $\int_{-\infty}^{+\infty} |\Phi(t)| dt < \infty$, then X can be shown to be continuous.

(suff. condition)

example. when X is cauchy we found that

$$\Phi(t) = e^{-|t|} \quad \text{so.}$$

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} e^{-|t|} dt$$

$$= \frac{1}{2\pi} \int_0^{+\infty} e^{-itx} e^{-t} dt + \frac{1}{2\pi} \int_{-\infty}^0 e^{-itx} e^t dt = \int_{-\infty}^{+\infty} (1+e^{-itx}) e^{-t} dt = (1+e^{-itx}) \Phi(t)$$

$$\begin{aligned}
 &= \frac{1}{2\pi} \int_0^\infty e^{-(ix+1)t} dt + \frac{1}{2\pi} \int_{-\infty}^0 e^{-(ix-1)t} dt \\
 &= \frac{1}{2\pi} \left[\frac{e^{-(ix+1)t}}{-ix+1} \right]_{t=0}^{+\infty} + \frac{1}{2\pi} \left[\frac{e^{(1-ix)t}}{1-ix} \right]_{-\infty}^0 \\
 &= \frac{1}{2\pi} \frac{1}{1+ix} + \frac{1}{2\pi} \frac{1}{1-ix} \\
 &= \frac{1}{2\pi} \frac{2}{(1+ix)(1-ix)} = \frac{1}{\pi} \frac{1}{1+x^2}
 \end{aligned}$$

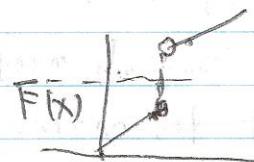
density func. for cauchy distr.

Theorem (2)

Let X be any random variable with distribution function $F(x)$ and characteristic func. $\phi(t)$.

Define

$$\begin{aligned}
 F(x) &= \frac{1}{2} [F(x) + \lim_{y \rightarrow x} F(y)] \\
 &= \frac{1}{2} [F(x) + F(x^-)]
 \end{aligned}$$



$$\begin{aligned}
 \text{Then } \bar{F}(b) - \bar{F}(a) &= \lim_{N \rightarrow \infty} \frac{1}{N} \int_a^b \frac{e^{-iat} - e^{-ibt}}{2\pi i t} \phi(t) dt \\
 \int_{-\infty}^{+\infty} g(x) dx, &= \lim_{c, d \rightarrow \infty} \int_c^d g(x) dx
 \end{aligned}$$

In fact, (2) implies (1)

Proof: When F is continuous at x , $\bar{F}(x) = F(x)$, so if $f'(x)$ exists, then F is continuous at x , and $\bar{F}(x) = f(x)$

$$\frac{\bar{F}(b) - \bar{F}(a)}{b-a} = \lim_{N \rightarrow \infty} \frac{1}{N} \int_a^b \frac{e^{-iat} - e^{-ibt}}{(b-a)i t} \frac{1}{2\pi} \phi(t) dt$$

Let $b \downarrow a$, so that $f(a) = \lim_{b \downarrow a} \frac{\bar{F}(b) - \bar{F}(a)}{b-a}$

$$\text{so } f(a) = \lim_{b \downarrow a} \lim_{N \rightarrow \infty} \int_{-N}^N \frac{e^{-iat} - e^{-ibt}}{(b-a)i t} \frac{1}{2\pi} \phi(t) dt$$

Non-trivial,

$$\stackrel{?}{=} \lim_{N \rightarrow \infty} \lim_{b \downarrow a} \int_N^N \frac{e^{-iat} - e^{-ibt}}{(b-a)i t} \frac{1}{2\pi} \phi(t) dt$$

$$= \lim_{N \rightarrow \infty} \lim_{b \downarrow a} \frac{1}{2\pi} \int_{-N}^N \frac{e^{-iat} - e^{-ibt}}{(b-a)i t} \frac{1}{2\pi} \phi(t) dt$$

Hilbert

$$\begin{aligned}
 &= \frac{1}{2\pi} \lim_{N \rightarrow \infty} \int_{-\infty}^N \left[\lim_{b \rightarrow a} \frac{e^{-ibt} - e^{-at}}{(b-a)t} \right] \phi(t) dt \\
 &= \frac{1}{2\pi} \lim_{N \rightarrow \infty} \int_{-\infty}^N e^{-iat} \phi(t) dt \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iat} \phi(t) dt
 \end{aligned}$$

Corollary of theorem 2.

If $\phi_x(t) = \phi_y(t)$ for all t , then $x \sim y$ (have the same distribution)

Proof: by theorem 2, if $\phi_x(t) = \phi_y(t)$ for all t ,

then $\bar{F}_x(x) = \bar{F}_y(x)$ for all real x .

so $f_x(x) = f_y(x)$ for all continuity points x of both F_x and F_y
(dis)

But all but countably many x are continuity points.

At a point of discontinuity x .

$$F_x(x) = \lim_{y \downarrow x} F_x(y) = \lim_{y \uparrow x} f_y(y) = f_y(x)$$

so $X \sim Y$ QED.

Definition: Let X have distribution function $F(x)$ and let $X_n, n \geq 1$ be a sequence of random variables with distribution functions $F_n(x), n \geq 1$ respectively.

We say that X_n converges in distribution to X

$$\text{if } \lim_{n \rightarrow \infty} F_n(x) = F(x)$$

For all x such that F is continuous at x

if this is true, the book writes $F_n \rightarrow F$

Continuity Theorem (5)

Let $F_n, n \geq 1$ be a sequence of distribution functions with corresponding characteristic functions $\phi_n(t)$, respectively.

(i) If $F_n \rightarrow F$ where F has characteristic function $\phi(t)$,

then $\phi_n(t) \rightarrow \phi(t)$ for all $t \in \mathbb{R}$.

(ii) inversely, if $\phi_n(t) \rightarrow \phi(t)$, for all $t \in \mathbb{R}$.

where $\phi(t)$ is continuous at $t=0$, then

Qualification for $\phi(t)$ to be a characteristic func.

ϕ is the characteristic function of some distribution F and $F_n \rightarrow F$. 19.

$(M_n(t) \rightarrow M(t) \text{ for all } t \text{ in a neighbourhood of } t=0)$

Example (6)

$$h: n \sqrt{2\pi}^{n+1/2} e^{-n} \xrightarrow[n \rightarrow \infty]{} \frac{n!}{n^{n+1/2} e^{-n}} = \text{constant.} = \sqrt{2\pi}$$

James Stirling & Abraham De Moivre

Oct 2 2008 . STAT 833 .

$\phi(t)$ is a continuous function of t

More assigned problems

3.9: 1, 3, 4 5.12, 5, 11, 14, 15

Ch 6 . Markov Chains

6.1 Markov process

We start with discrete time stochastic processes in a state space S which is countable.

time $T = \{0, 1, 2, \dots\}$ or $\{1, 2, 3, \dots\}$

process $X = \{X_0, X_1, X_2, \dots\}$ or $\{X_1, X_2, \dots\}$

Definition: X is called a Markov Chain if it satisfies the Markov property

$$P(X_n = s | X_0 = s_0, X_1 = s_1, \dots, X_{n-1} = s_{n-1})$$

$$= P(X_n = s | X_{n-1} = s_{n-1})$$

$$P(\text{immediate future} | \text{present, past}) = P(\text{immediate future} | \text{present})$$

Examples: ① simple random walks

② Branching processes.

Note the equivalent conditions (2), (3) on page 214.

(4) Definition. X is said to be (time) homogeneous if the probability

$$P(X_{n+1} = s | X_n = t) = P(X_1 = s | X_0 = t) \text{ for all } n, s, t$$

It is convenient to represent the state space S as a set of integers.

$$P(X_{n+1} = j | X_n = i) = P(X_1 = j | X_0 = i)$$

The transition matrix of a homogeneous M.C. is

$$P = (P_{ij})$$

$$P_{ij} = P(X_{n+1} = j \mid X_n = i)$$

When the MC is not homogeneous, then P will be a function of n also.

(5) Theorem: $P = (P_{ij})$ satisfies

$$(a) P_{ij} \geq 0 \text{ for all } i, j$$

$$(b) \sum_j P_{ij} = 1 \text{ for all } i. \text{ (the rows sum to 1)}$$

(6) Definition: The k -step transition matrix $P(m, m+n) = (P_{ij}(m, m+n))$

is defined by

$$P_{ij}(m, m+n) = P(X_{m+n} = j \mid X_m = i)$$

(7) Theorem (Chapman-Kolmogorov Equation)

$$P_{ij}(m, m+n+r) = \sum_k P_{ik}(m, m+n) \cdot P_{kj}(m+n, m+n+r) \quad (\text{law of total probability})$$

$$\begin{matrix} & \text{ok} \\ \text{m} & \text{m+n} & \text{m+n+r} \end{matrix}$$

$$P(m, m+n+r) = P(m, m+n) \cdot P(m+n, m+n+r) \quad \text{in matrix form } X$$

(Corollary) Since we are primarily interested in homogeneous MC, Theorem (7) implies that

$$P(m, m+n) = p^n \text{ in this case.}$$

Proof of Theorem 7 and Corollary is on page 215.

Define the vector $\pi^{(n)} = (\pi_i^{(n)})$ where $\pi_i^{(n)} = P(X_n = i)$

$$\text{where } \pi_i^{(n)} \text{ is } = P(X_n = i)$$

$$\text{Then } \pi^{(m+n)} = \pi^{(m)} \cdot P^n$$

$$\text{or } \pi^{(n)} = \pi^{(0)} \cdot P^n$$

We also define

$$P^n = (P_{ij}(n))$$

Example: Simple Random Walk.

$$S = \{0, \pm 1, \pm 2, \dots\}$$

$$P_{ij} = \begin{cases} \frac{1}{2} & \text{if } j = i+1 \\ \frac{1}{2} & \text{if } j = i-1 \\ 0 & \text{otherwise.} \end{cases}$$

$$(P, S) = \emptyset$$

21

$$P^n = (P_{ij}(n))$$

$$P_{ij}(n) = P(Z \text{ Bin}(n, p) - i = j - i)$$

$$j-i = \# \text{ up} - \# \text{ down}$$

$$= \# \text{ up} - (n - \# \text{ up})$$

$$= 2\# \text{ up} - n$$

$$= P(Z \text{ Bin}(n, p) = \frac{n+j-i}{2})$$

$$= \begin{cases} \binom{n}{\frac{n+j-i}{2}} p^{\frac{n+j-i}{2}} q^{\frac{n-j+i}{2}} & \text{even} \\ 0 & \text{odd} \end{cases}$$

Example: 6.1. #2.

A fair die is rolled successively, let

$$Y_n = \# \text{ that appears on roll } n \quad n \geq 1$$

which of the following are MC's?

(a) $X_n = \text{maximum value of first } n \text{ rolls.}$ (Yes, homogeneous)

$$S = \{1, 2, 3, 4, 5, 6\}$$

$$P_{ij} = \begin{cases} 0 & j < i \\ \frac{1}{6} & j = i \\ \frac{1}{6} & j > i \end{cases}$$

$$P_{24} = P(X_{n+1} = 4 \mid X_n = 2) = P(Y_{n+1} = 4 \mid X_n = 2)$$

$$= P(Y_{n+1} = 4) = 1/6.$$

(b) $N_n = \# \text{ of 6's in the first } n \text{ rolls.}$

$$S = \{0, 1_2, \dots\}$$

$$P_{ij} = \begin{cases} 1/6 & j = i+1 \\ \frac{5}{6} & j \leq i \\ 0 & j < i \end{cases}$$

Homogeneous - increasing seq.

(c) $C_n = \text{time since the most recent 6.}$

$$Y_n : 1 \ 3 \ 6 \ 4 \ 6 \ 2 \ 3 \ 1 \ 6$$

$$C_n : 1 \ 2 \ 0 \ 1 \ 0 \ 1 \ 2 \ 3 \ 0$$

pretend

$$S = \{c_1, c_2, \dots\} \quad P_{ij} = \begin{cases} 1/6 & j = 0 \quad \text{1-step transition} \\ 1/6 & j = i+1 \quad \text{probability} \\ 0 & \text{otherwise} \end{cases}$$

Hilary

(d) You do it.

Oct 7, 2008 STAT 833 10:00 am - 11:20 am

Assigned problems: 6.1: 1, 2, 7, 8, 12

6.2: 1, 2, 3 6.3: 1, 2, 3

Homogeneous MC

Review section 6.2 ~ 6.3

6.2 Classification of states

$P(X_n=i \text{ for some } n \geq 1 \mid X_0 = i)$

Probability (return to state i) start in state i)

≥ 1

< 1

persistent aka-recurrent

transient

First passage times

$f_{ij}(n) = P(X_n=j \text{ for the first time} \mid X_0 = i)$

$n \geq 0$

$= P(X_1 \neq j, X_2 \neq j, \dots, X_{n-1} \neq j, X_n = j \mid X_0 = i)$

$P_{ij}(n) = n\text{-step transition probability}$

$= P(X_n = j \mid X_0 = i) \text{ (i,j)th element of } P^n$

Generating functions:

$P_{ij}(s) = \sum_{n=0}^{\infty} s^n P_{ij}(n) \text{ where } P_{ij}(0) = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$

$F_{ij}(s) = \sum_{n=1}^{\infty} s^n f_{ij}(n)$

Theorem (a) $P_{ii}(s) = 1 + F_{ii}(s)P_{ii}(s)$

(b) if $i \neq j$ then

$$P_{ij}(s) = F_{ij}(s)P_{ij}(s)$$

Proof: page 221.

Corollary: (a) A state j is persistent if

$$\sum_{n=0}^{\infty} P_{jj}(n) = \infty$$

(b) A state j is transient if

$$\sum_{n=0}^{\infty} P_{jj}(n) < \infty$$

Furthermore, if j is persistent, then $\sum_{n=1}^{\infty} P_{jj}(n) = \infty$ and

if j is transient, then $\sum_{n=1}^{\infty} P_{jj}(n) < \infty$

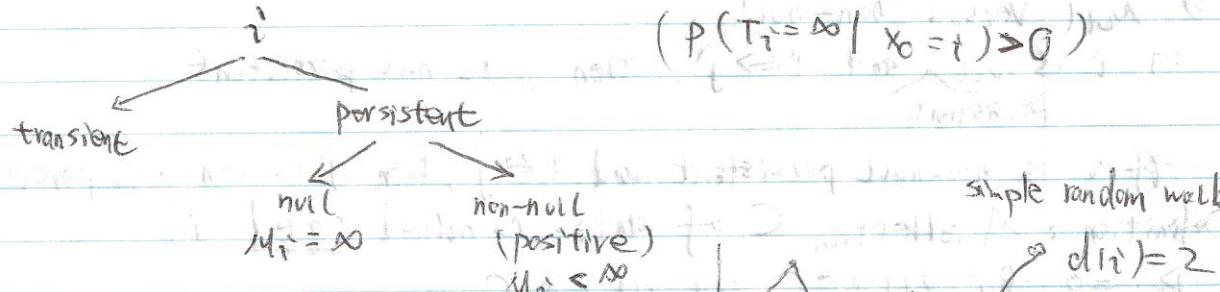
Corollary: if i is transient, then $P_{ij}(n) \rightarrow 0$ as $n \rightarrow \infty$ for all j .

Define $T_i = \min \{ n \geq 1 \text{ such that } X_n = i \}$

$$M_i = \text{Exp}[T_i | X_0 = i]$$

$$= \sum_{n=1}^{\infty} n f_{ii}(n) \quad \begin{cases} \text{if } i \text{ is persistent } (F_{ii}(1) = 1) \\ \infty \quad \text{if } i \text{ is transient } (F_{ii}(1) < 1) \end{cases}$$

$$(P(T_i = \infty | X_0 = i) > 0)$$



Definition: The period $d(i)$ of state i is

$$d(i) = \text{gcd} \{ n : P_{ii}(n) > 0 \}$$

We say that i is periodic, if $d(i) > 1$, we say that i is aperiodic if

$$d(i) = 1.$$

Definition: A state i which is persistent, non-null, aperiodic is called ergodic. (law of large numbers)

6.3 Classifying Chains

Definition: State i communicates with state j if $P_{ij}(n) > 0$ for some $n \geq 0$. We write $i \rightarrow j$.

State i intercommunicates with state j , if $i \rightarrow j$, and $j \rightarrow i$. we write $i \leftrightarrow j$.

Proposition: \leftrightarrow is an equivalence relation.

(a) (reflexivity) $i \leftrightarrow i$, for all $i \in S$

(b) $i \leftrightarrow j$ implies $j \leftrightarrow i$ for all $i, j \in S$ (symmetry)

(c) (transitivity) $i \leftrightarrow j$ and if $j \leftrightarrow k$ implies $i \leftrightarrow k$ for all $i, j, k \in S$

so \leftrightarrow breaks up S into equivalence classes.

Definition: A property of a state i is called a class property if j also has the property whenever $i \leftrightarrow j$

Theorem: The following are class properties

① The period $d(i)$

If $i \leftrightarrow j$, then $d(i) = d(j)$

② Transience

If i is transient, $i \leftrightarrow j$, then j is transient

③ Null versus non-null

If i is null and $i \leftrightarrow j$, then j is non-persistent

If i is non-null persistent and $i \leftrightarrow j$, then j is non-null persistent

Definition: A collection C of states is called closed if

$P_{ij} = 0$ for all $i \in C$ and $j \notin C$

C ($i \rightarrow j$)
one step. $P_{ij} = 0$ good enough for multiple steps

Definition: A collection C of states is called irreducible if $i \leftrightarrow j$

for all $i, j \in C$

Decomposition Theorem:

$$S = T \cup C_1 \cup C_2 \cup \dots$$

where T is the set of all transient states and C_i is a closed, irreducible collection of persistent states.

Lemma: S a finite MC (ie. finitely many states)

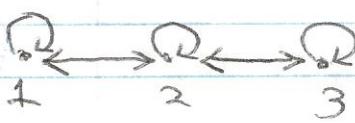
at least 1 state is persistent, and all persistent states are non-null.

Problem 3(a) Classify the states for

$$P = \begin{pmatrix} 1-2p & 2p & 0 \\ p & 1-2p & p \\ 0 & p & 1-2p \end{pmatrix}, 0 < p < 0.5$$

All states intercommunicate,

By the lemma, at least 1 state is persistent and non-null, \Rightarrow every state is non-null & persistent.



Also, $d(1) = d(2) = d(3) = 1$

$$d(1) = \gcd\{1 \geq 0 \text{ s.t. } P_{1j}(n) > 0\} = \gcd\{1, \dots\} = 1$$

STAT 833 Oct 9 2008 10:00 am - 11:20 am

Sec 6.4. Stationary distributions and the limit theorem

π - vector of probability weights.

$$\pi = (\pi_j : j \in S)$$

Definition: The vector π is called a stationary distribution if

$$(a) \pi_j \geq 0 \quad \sum_{j \in S} \pi_j = 1$$

$$(b) \pi = \pi P \text{ or } \pi_j = \sum_{i \in S} \pi_i P_{ij} \text{ for all } j \in S$$

Note: π is a left eigenvector.

Note: By mathematical induction, $\pi = \pi P^n$ for all $n \geq 0$.

(3) Theorem An irreducible Markov chain has a stationary distribution if and only if all states are non-null persistent.

In this case, the stationary distribution π is unique, and

$$\bar{\pi}_i = \frac{1}{\mu_i} \text{ where } \mu_i = E[\tau_i | X_0 = i] \text{ mean time to return to state } i \text{ from } i.$$

Note: An irreducible MC is one where all states intercommunicate.

Note: The condition for existence of π is "if and only if".

Sc, one way to prove that states are non-null persistent is to find π and show it is a probability vector.

What about MC's that are irreducible and null-persistent?

(6) Theorem

Suppose a MC is irreducible and persistent, then there exists a positive root $x = (x_i : i \in S)$ of the equation $x = xP$, which is unique up to a multiplicative constant. Also

$$(a) \sum_{i \in S} x_i < \infty \Leftrightarrow \text{MC is non-null}$$

$$(b) \sum_{i \in S} x_i = \infty \Leftrightarrow \text{MC is null.}$$

Note: x positive means $x_i > 0$ for all $i \in S$

Note: x is unique up to a multiplicative constant means $(x_i : i \in S)$ ^{this} means

a solution if and only if $(c x_p : t^T s) c > 0$ is a solution.

Note $\pi_p = \frac{x_p}{\sum_{j \in S} x_j}$ where MC is non-null.

What about MC's that are irreducible and transient?

(1c) Theorem. Let $s \in S$ be any states of an irreducible MC. The MC is transient if and only if there exists a non-zero solution to the following $(y_j : j \neq s)$ such that

$$(a) |y_j| \leq 1 \text{ for all } j \neq s.$$

$$(b) y_i = \sum_{j:j \neq s} p_{ij} y_j$$

Note: $(y_j : j \neq s)$ is a right eigenvector of the matrix

$$(p_{ij} : i, j \neq s)$$

sketches of proofs:

proof of theorem(3):

Step 1: Start a MC in state k and let

$N_t := \# \text{ of visits to state } i \text{ before return to } k$.

$$(N_k = 1)$$

Define $P_t(k) = E(N_t | x_0 = k)$

where $P_k(k) = 1$

Consider $\sum_{i \in S} P_i(k) = \text{Expected } \# \text{ of states visited before return to state } k$

$$= \text{Expected } (T_p | x_0 = k) = \mu_k$$

Step 2: Prove that if

$$P(k) = (P_i(k) : i \in S) \text{ then}$$

$$P(k) = P(k)P$$

Step 3: "Set" $\pi_i = \frac{P_i(k)}{\mu_k}$

existence

This proves that any non-null persistent irreducible MC has a stationary distribution.

Next: suppose a stationary distribution exists,

irreducible

i
↓
j

27.

then the argument proceeds by showing that for any solution $\pi = (\pi_i : i \in S)$ we have

$$\pi_i \mu_i = 1 \text{ for all } i \in S.$$

$$\text{so } \pi_i = 1/\mu_i \text{ where } \mu_i = E[T_i | x_0 = i]$$

Thus π must be unique and the MC is non-null persistent.

(15) Example: random walk with barrier

$$S = \{0, 1, 2, \dots\}$$

$$P_{00} = q, P_{01} = p, P_{10} = 1$$

$$P_{i,i-1} = q, P_{i,i+1} = p \text{ for } i \geq 1$$

The MC is irreducible.

Solving $\pi = \pi P$:

$$\pi_0 = \pi_0 P_{00} + \pi_1 P_{10} \quad \text{or} \quad \pi_1 = \pi_0 P + \pi_2 q$$

$$\pi_1 = \pi_0 P_{01} + \pi_2 P_{21}$$

$$\pi_2 = \pi_1 P_{12} + \pi_3 P_{32}$$

$$\pi_0(1-q) = \pi_1 q$$

$$\text{or } \pi_0 p = \pi_1 q \therefore \frac{\pi_1}{\pi_0} = \frac{p}{q} := P \text{ say}$$

$$\pi_2 q = \pi_1 - \pi_0 p = \pi_1 - (\pi_1 \frac{p}{q}) p = \pi_1 - \pi_1 q = \pi_1 p$$

$$\therefore \pi_2 / \pi_1 = p/q = P$$

proceeding by induction

$$\pi_j / \pi_{j-1} = P \text{ for } j \geq 1$$

To have a probability distribution, then we have

(a) $P < 1$, then using $\sum_{j=0}^{\infty} \pi_j = 1$ we get

$$\pi_j = P^j (1-P)$$

(b) $P = 1/2$ ($x = (1, 1, 1, \dots)$) solves $x = xP$ because $P = q = 0.5$

But x is not a probability distribution.

$\sum x_i = \infty$. By theorem (b), the MC is null-persistent.

(c) $p > 1$, MC is transient. See page 222.

(17) Theorem

When the MC is ergodic (irreducible, non-null persistent, aperiodic)

$$\pi_j = \lim_{n \rightarrow \infty} P_{ij}(n) \text{ for all } j \in S$$

In this case, π is the unique stationary distribution and $\pi_i = 1/m_i$ where m_i is the expected time to return to state i .

STAT 333 Oct 14 2008.

Assigned Problems

6.4: 1, 2, 4, 7

6.5. Time reversibility

Time reversible MC

Time irreversible MC clockwise

A MC is (time) reversible if it has the same properties in reverse as it has in the original time order, when the system is in equilibrium.

Definition: A MC X_n is called time reversible if it is

$$\pi_i P_{ij} = \pi_j P_{ji} \text{ for all } i, j.$$

// the rate of flow from state j to state i .
the rate of flow from state i to state j

in equilibrium.

(4) Theorem: Suppose X is an irreducible MC and $\pi = (\pi_i : i \in S)$ is the probability distribution such that

$$\pi_i P_{ij} = \pi_j P_{ji} \text{ for all } i, j$$

then π is a stationary distribution.

Assigned Problems 6.5: 1, 2, 4, 6.

We will skip sections 6.6 and 6.7.

6.8. Poisson process and birth processes

In this section, we consider continuous time processes $N(t)$ where $t \geq 0$

such that $N(t) \in \{0, 1, 2, \dots\}$

Counting process $\left\{ \begin{array}{l} N(0) = 0 \text{ or more generally } N(0) \geq 0 \\ N \text{ nondecreasing: if } s < t \text{ implies } N(s) \leq N(t) \end{array} \right.$

Definition (Poisson process)

A poisson process of intensity $\lambda > 0$ is a counting process

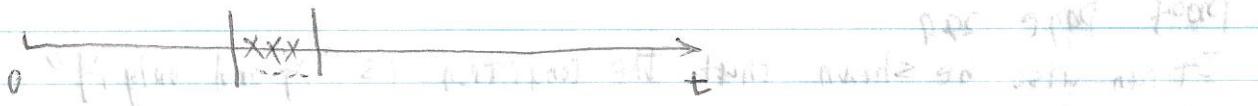
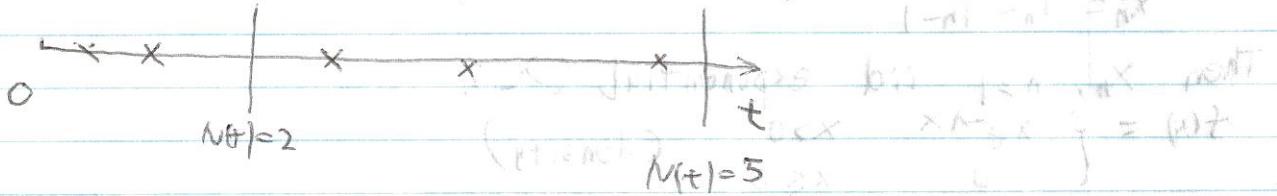
$N(t)$ such that

$$(a) N(0) = 0 \quad N(s) \leq N(t) \text{ for all } s < t$$

$$(b) P(N(t+h) = m \mid N(t) = n) = \begin{cases} \lambda h + o(h) & n=1 \\ 1 - \lambda h + o(h) & n=0 \\ o(h) = 0 & m > 1 \end{cases}$$

as $h \downarrow 0$.

(c) If $s < t$, then $N(t) - N(s)$ is independent of $(N(u), 0 \leq u \leq s)$



(2) Theorem $N(t) \sim \text{Poisson}(\lambda t)$

$$\begin{aligned} \text{Proof: (sketch)} \quad P(N(t+h) = j) &= \sum_i P(N(t) = i) P(N(t+h) = j) \\ &= \sum_i P(N(t) = i) P(N(t+h) - N(t) = j-i \mid N(t) = i) \\ &= \sum_{i=j}^{\infty} P(N(t) = i) P(N(t+h) - N(t) = j-i) \quad \text{by assumption (c)} \\ &= P(N(t) = j)(1 - \lambda h + o(h)) + P(N(t) = j-1)(\lambda h + o(h)) \\ &\quad + P(N(t) < j-1) \cdot o(h) \quad \text{by condition (b)} \end{aligned}$$

Let us define

$$P_j(t) = P(N(t) = j)$$

Let us consider the case where $j = 0$

In this case, the formula reduces to $P_0(t+h) = P_0(t) \cdot (1 - \lambda h + o(h))$

$$\frac{P_0(t+h) - P_0(t)}{h} = [\lambda + \frac{o(h)}{h}] P_0(t) \quad \text{let } h \downarrow 0 \text{ then}$$

$$\boxed{P'_0(t) = (\lambda t + 0) P_0(t) - \lambda P_0(t)} \quad \begin{array}{l} \text{Also } P_0(0) = 1 \\ \text{boundary condition} \end{array}$$

This is solved by

$$P_0(t) = e^{-\lambda t}$$

$$\text{For } j \neq 0, \quad P'_j(t) = \lambda P_{j-1}(t) \cdot t - \lambda P_j(t) \quad P_j(0) = 0$$

In the text book, these are solved by

- recursion on $j=0, 1, 2, \dots$

- generating functions.

$$P_j(t) = P(N(t)=j) = \frac{(at)^j e^{-at}}{j!} \quad j=0, 1, 2, \dots$$

(c) Theorem Define $T_0 = 0$

$$T_n = \inf \{ t : N(t) = n \}$$

$x_1 \ x_2 \ x_3 \ x_4 > 2 \rightarrow$

$$x_n = T_n - T_{n-1}$$

Then $x_n, n \geq 1$ iid exponential (λ).

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad (\text{density})$$

Proof Page 249.

It can also be shown that the condition is "if and only if".

$N(t)$ Poisson process of intensity λ .



$x_n, n \geq 1$ are iid. $\text{Exp}(\lambda)$

Definition: A birth process with intensity $\lambda_n, n=0, 1, 2, \dots$

is a process $N(t)$ such that

(a) $N(0) \geq 0$, $N(t)$ nondecreasing

(b) $P(N(t+h) = n | N(t) = m) = \int_{-\infty}^{\infty} (1 - \lambda_n h + o(h))^{m-0} e^{-\lambda_n h} dh$

(c) If $s < t$, then

$(N(t) - N(s) | N(s))$ whereas $U \leq s < t$

$\sim (N(t) - N(s) | N(s))$ Markov property

STAT 833 STAT Oct 16, 2008

Examples: ① $\lambda_n = \lambda$ for all n Poisson process.

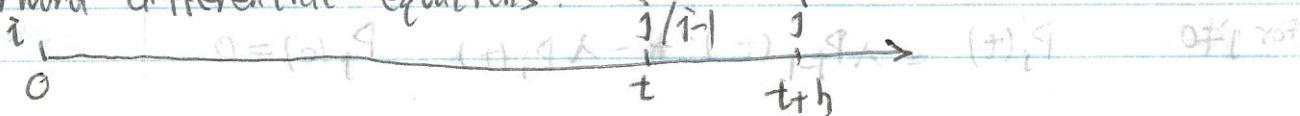
② $\lambda_n = n\lambda$ simple birth process, Yule process

③ $\lambda_n = n\lambda + \nu$ simple birth process with immigration

for $j \geq i$. Define $P_{ij}(t) = P(N(s+t) = j | N(s) = i)$

$$= P(N(t) = j | N(0) = i)$$

Forward differential equations.



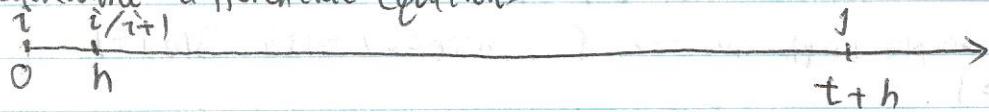
$$P_{ij}(t+h) = P_{ij}(t)(1 - \lambda_j h) + P_{i,j-1} \lambda_{j-1} h + o(h) \quad \text{as } h \rightarrow 0$$

$$\frac{P_{ij}(t+h) - P_{ij}(t)}{h} = P_{i,j-1}(t) \lambda_{j-1} - P_{ij}(t) \lambda_j + \frac{o(h)}{h}$$

$$P'_{ij}(t) = P_{i,j-1}(t) \lambda_{j-1} - P_{ij}(t) \lambda_j \quad P_{ij}(0) = \begin{cases} 1 & i=j \\ 0 & \text{otherwise} \end{cases} = \delta_{ij}$$

Forward equations

Backward differential equations



$$P_{ij}(t+h) = (1 - \lambda_i h) P_{ij}(t) + \lambda_i h P_{i+1,j}(t) + o(h)$$

$$P'_{ij}(t) = \lambda_i P_{i+1,j}(t) - \lambda_i P_{ij}(t) \quad P_{ij}(0) = \delta_{ij}$$

(14) Theorem: The forward equations have a unique solution which satisfies the backward equations.

(15) Theorem: If $P_{ij}(t)$, $j \geq i \geq 0$ solves the forward equations and $\pi_{ij}(t)$, $j \geq i \geq 0$ solves the backward equations.

Then $P_{ij}(t) \leq \pi_{ij}(t)$ for all i, j, t

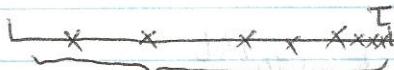
Idea: Among all solutions $\pi_{ij}(t)$ to the backward equations,

$$\sum_{j=0}^{\infty} P_{ij}(t) \leq \sum_{j=0}^{\infty} \pi_{ij}(t) \quad (?=1)$$

problem with this, $\sum_{j=0}^{\infty} P_{ij}(t) \neq 1$ in general.

because it is possible that

$$\lim_{t \rightarrow T} N(t) = \infty \quad \text{with positive probability}$$



infinitely many births

$$\sum_{j=i}^{\infty} P_{ij}(t) = P(N(t) \text{ is finite}) \sim \text{possibly less than one}$$

(18) Definition Let $T_n = \inf \{t : N(t) \geq n\}$ Let $T_{\infty} = \lim_{n \rightarrow \infty} T_n$ Hilroy

We say that $N(t)$ is honest if $P(T_{\infty} = \infty) = 1$

If $P(T_{\infty} = \infty) < 1$ then $N(t)$ is said to be dishonest.

(19) Theorem

$N(t)$ is honest if and only if $\sum \frac{1}{n^2} < \infty$

$$\sum_{n=0}^{\infty} \frac{1}{\lambda_n} = \infty$$

\downarrow

$u_i = \frac{1}{\lambda_{i+1}}$ $\lambda_{i+1} = \frac{1}{u_i}$

$$\sum \frac{1}{n} = \infty$$

The property is only defined

for a short interval of time.

The process cannot get to infinity in a finite time.

Example: problem 6.8.4 (we consider the case $I=1$)

$B(t)$ is a simple birth process (Yule process) with $B(0)=1$

Then for $k \geq 1$

$$P(B(t)=k) = e^{-\lambda t} (1-e^{-\lambda t})^{k-1} \quad k=1, 2, 3, \dots$$

Geometric distribution.

$$\lambda_0 = n\lambda$$

Proof: Forward equations $i=j=1$

$$P'_{1,1}(t) = P_{1,0}(t)\lambda_0 - P_{1,1}(t)\lambda_1 \quad (\lambda_0 = 0)$$

$$P'_{1,1}(t) = -\lambda P_{1,1}(t) \quad \lambda_1 = \lambda$$

$$\frac{d}{dt} \ln P_{1,1}(t) = -\lambda$$

$$P_{1,1}(t) = \text{constant} \cdot e^{-\lambda t} \quad \text{Boundary condition } P_{1,1}(0) = 1$$

$$\Rightarrow \text{constant} = 1$$

$$\text{So } P(B(t)=1) = P_{1,1}(t) = e^{-\lambda t}$$

For $j > 1$

$$P'_{1,j}(t) = P_{1,j-1}(t)\lambda_{j-1} - P_{1,j}(t)\lambda_j$$

$$P'_{1,j}(t) = P_{1,j-1}(t)(j-1)\lambda - P_{1,j}(t) \cdot j\lambda$$

$$P_{1,j}(t)e^{\lambda t} + P_{1,j}(t) \cdot j\lambda \cdot e^{\lambda t} = P_{1,j-1}(t)(j-1)\lambda \cdot e^{\lambda t}$$

$$\text{Define } r_j(t) = P_{1,j}(t)e^{\lambda t}$$

$$\text{Then } r_j(t) = P_{1,j}(t)e^{\lambda t} = 1 \quad \text{for all } t$$

$$r'_j(t) = (j-1)\lambda r_{j-1}(t)e^{\lambda t}$$

This can be solved recursively.

$$r_2'(t) = \lambda r_1(t) e^{\lambda t} = \lambda e^{\lambda t}$$

$$r_2(t) = e^{\lambda t} + \text{constant}$$

Since $P_{12}(0)=0$ we have $r_2(0) = P_{12}(0) e^{2\lambda 0} = 0$ for all t

$$\text{So } 0 = e^{2\lambda 0} + \text{constant} \Rightarrow \text{constant} = 0 - 1 = -1$$

Therefore $r_2(t) = e^{\lambda t} - 1$ Therefore $P_{12}(t) = e^{-2\lambda t} (e^{\lambda t} - 1)$

$$r_3'(t) = 2\lambda r_2(t) e^{\lambda t} = 2\lambda (e^{\lambda t} - 1) e^{\lambda t} = e^{-\lambda t} (1 - e^{-\lambda t})$$

and so on. By mathematical induction, the general formula

$$P_{ij}(t) = e^{-\lambda t} (1 - e^{-\lambda t})^{j-1} \quad j=1, 2, \dots \quad \text{can be proved.}$$

$$G(s, t) = \sum_{j=1}^n P_{ij}(t) s^j \rightarrow \text{partial differential equations}$$

STAT 833 Oct 21, 2008

Assigned problems 6.8 1, 2, 3, 4, 5. These problems cumulatively will be on test 1 including all related sections.

6.9 Continuous time Markov chains

*. result has regularity assumptions to make it true which are not stated

Def. $X(t)$ is said to be a continuous time markov chain if it satisfies the Markov Property

$$P(X(t_n) = i | X(t_1) = i_1, X(t_2) = i_2, \dots, X(t_{n-1}) = i_{n-1})$$

$$= P(X(t_n) = i | X(t_{n-1}) = i_{n-1}) \quad \text{for all } i_1, \dots, i_{n-1} \in S$$

(countable), and any $t_1 < t_2 < \dots < t_n$

We define the transition probability

$$P_{ij}(s, t) = P(X(t) = i | X(s) = i) \quad \text{for all } 0 \leq s \leq t. \text{ We say that}$$

X is (time) homogeneous if $P_{ij}(s, t) = P_{ij}(0, t-s)$ for all $i, j \in S$

and all $0 \leq s \leq t$. Henceforth, we assume homogeneity. we write

Define a matrix P by $P_{ij}(t) = P(X(t) = i | X(0) = i)$

$$P_t = (P_{ij}(t)) \quad \text{for all } t \geq 0$$

The family of matrices $\{P_t, t \geq 0\}$ is called the transition semigroup of X

Theorem: $\{P_t\}$ is a stochastic semigroup with identity. That is

(a) P_t is a stochastic matrix for all $t \geq 0$
i.e., $P_{ij}(t) \geq 0$ and $\sum_{j \in S} P_{ij}(t) = 1$

(b) P_0 is identity I. matrix.

(c) $P_{st} = P_s \cdot P_t$ for all $s, t \geq 0$

These are the Chapman-Kolmogorov equations.

Definition: $\{P_t\}$ is called standard if $P_{ij}(t)$ is a continuous function of t .

It suffices to check continuity at $t=0$.

So $\{P_t\}$ is standard if and only if

$$\lim_{t \rightarrow 0} P_t = P_0 = I, \text{ that is } \lim_{t \rightarrow 0} P_{ij}(t) = \delta_{ij} \text{ for all } i, j$$

We can show for a standard MC,

$$P_{ij}(h) = \begin{cases} g_{ij}, & i \neq j \\ h + g_{ii}h + o(h), & i = j \end{cases}$$

as $h \downarrow 0$: Define

$$G = (g_{ij}) = \begin{pmatrix} - & + & + \\ + & - & + \\ + & + & - \end{pmatrix} \quad \text{regularity needed}$$

$$\begin{aligned} \text{Now } 1 &= \sum_j P_{ij}(h) = P_{ii}(h) + \sum_{j \neq i} P_{ij}(h) = (h + g_{ii}h) + \sum_{j \neq i} g_{ij}(h) + o(h) \\ &= h + \sum_j g_{ij} \cdot h + o(h) \end{aligned}$$

Let $h \downarrow 0$, $\boxed{\sum_j g_{ij}h = 0}$ \star "rows of G sum to zero"

Example: (7) Birth process

$$P_{ij}(h) = \begin{cases} 1 - \lambda_i h + o(h), & i = j \\ \lambda_i h + o(h), & j = i+1 \\ 0, & j > i+1 \end{cases}$$

$G = (g_{ij})$, where $g_{ii} = -\lambda_i$ & $g_{i,i+1} = \lambda_i$

$g_{ij} = 0$ for all other cases.

Forward Equations for a general MC. Backward Equations for a general MC

$$P_t' = P_t G \quad \star \text{ regularity} \quad P_t' = G P_t \quad \star$$

Boundary condition $P_0 = I$

$$P_t' = (P_{ij}(t))$$

★ There exists a unique solution to each equation.

These are solved by $P_t = \exp(tG)$ where ★

$$\begin{aligned} y' &= ky \\ y' &= yk \end{aligned}$$

$$\exp(A) = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots \text{ for any square matrix } A$$

Next, suppose $X(s) = i$. Define

$$U = \inf \{t \geq 0 : X(s+t) \neq i\}$$

★ Proposition $U \sim \text{Exp}(-g_{ii})$

"Proof": We show U has the memoryless property. Set $s=0$ for simplicity.

$$\begin{aligned} P(U > x+y \mid U > x) &= P(X(t)=i, 0 \leq t \leq x+y \mid X(t)=i, 0 \leq t \leq x) \\ &= P(X(t)=i, x \leq t \leq x+y \mid X(t)=i, 0 \leq t \leq x) \end{aligned}$$

Markov Property

$$= P(X(t)=i, x \leq t \leq x+y \mid X(x)=i)$$

Homogeneity

$$= P(X(t)=i, 0 \leq t \leq y \mid X(0)=i) = P(U > y)$$

"So" $U \sim \text{Exponential}(\lambda)$ for some λ .

$$P(U > h) = 1 - g_{ii}h + o(h) = e^{-g_{ii}h} + o(h) \text{ so } \lambda = -g_{ii}$$

★ Proposition: The probability that immediately upon departure from i , X moves to j is $\frac{-g_{ij}}{g_{ii}}$

Proof: Suppose $X(0)=i$, and $x < U \leq x+h$ for small h .

with high probability, there is exactly one jump in $(x, x+h]$ so

$$\begin{aligned} P(X \text{ jumps to } j \mid X \text{ jumps}) &= \frac{P_{ij}(h)}{1 - P_{ii}(h)} = \frac{g_{ij}h + o(h)}{1 - [1 + g_{ii}(h) + o(h)]} = \frac{g_{ij}h + o(h)}{-g_{ii}h + o(h)} \\ &= \frac{g_{ij}}{-g_{ii}}. \end{aligned}$$

STAT 833 2008 Oct 23

To find a stationary dist. of X we solve

$$\pi P_t = \pi \text{ for all } t$$

Hilroy

Let us use the Chapman-Kolmogorov equation. we differentiate wrt. t.

$$\Pi P_t' = 0$$

$$\therefore \Pi G P_t = 0 \quad \text{for all } t$$

$$\text{or } \Pi P_t G = 0$$

$$\text{Set } t=0, \text{ then } P_0 = I \quad \text{so } \boxed{\Pi G = 0}$$

Solve for Π to find the stationary distribution of X .

Theorem: If X is an irreducible MC. with semigroup $\{P_t, t \geq 0\}$ then

(a) if a stationary distribution Π exists, then it is unique, and

$$\lim_{t \rightarrow \infty} P_{ij}(t) = \pi_{ij} \quad \text{nonnull persistent}$$

(b) if no stationary distribution exists, then

$$\lim_{t \rightarrow \infty} P_{ij}(t) = 0 \quad \text{for all } i, j \quad \text{transient, null persistent}$$

State i is persistent if

$$P(\text{amount of time spent in state } i \text{ is infinite} | X(0)=i) = 1$$

State i is transient if

$$P(\text{amount of time spent in state } i \text{ is infinite} | X(0)=i) = 0$$



Jump chain:

$$P = \begin{pmatrix} 0 & -g_{11} \\ g_{01} & \frac{-g_{11}}{g_{11}} \\ -g_{11} & \ddots \end{pmatrix}$$

Have, the rows sum to one

$$\sum_{j:j \neq i} \frac{-g_{ij}}{g_{ii}} = \frac{-\sum_{j:j \neq i} g_{ij}}{g_{ii}} = \frac{-\sum_j g_{ij} + g_{ii}}{g_{ii}} = 1$$

so P is the transition matrix for a discrete time MC.

This is called the jump chain.

Continuous time MC
 Jump chain \downarrow \uparrow Embedded chain
 Discrete time MC

Stationary distribution of continuous time MC
 \dagger

Stationary distribution of its jump chain. Has jumps down weight sojourn time

6.11. Birth and death processes

$$X(t) = \{0, 1, 2, \dots\}$$

$$P(X(t+h) = n+m | X(t) = n) = \begin{cases} \lambda_n h + o(h) & m=1 \\ \mu_n h + o(h) & m=-1 \\ 1 - (\lambda_n + \mu_n)h + o(h) & m=0 \end{cases}$$

where λ_n, μ_n are the birth rates and
 $\lambda_n \geq 0$ for all $n \geq 0$ and

μ_n, μ_0 are the death/mortality rates and
 $\mu_n \geq 0$ for all n . Also $\mu_0 = 0$

so

$$G = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \dots \\ \mu_1 - (\lambda_1 + \mu_1) & \lambda_1 & 0 & \dots \\ \mu_2 - (\lambda_2 + \mu_2) & \lambda_2 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

The stationary distribution π on $S = \{0, 1, 2, \dots\}$ is found by solving $\pi G = 0$

$$(\pi_0, \pi_1, \pi_2, \dots) \begin{pmatrix} -\lambda_0 & \lambda_0 \\ \mu_1 - (\lambda_1 + \mu_1) & \lambda_1 \\ \mu_2 - (\lambda_2 + \mu_2) & \lambda_2 \end{pmatrix} = 0$$

$$\textcircled{1} \quad -\lambda_0 \pi_0 + \mu_1 \pi_1 = 0$$

$$\textcircled{2} \quad \lambda_0 \pi_0 - (\lambda_1 + \mu_1) \pi_1 + \mu_2 \pi_2 = 0$$

\textcircled{3} \quad :

Solve \textcircled{1} $\pi_1 = \frac{\lambda_0}{\mu_1} \pi_0$

Solving \textcircled{2} by plugging this into \textcircled{2} $\lambda_0 \pi_0 - \frac{\lambda_0 \lambda_1}{\mu_1} \pi_0 - \lambda_0 \pi_0 + \mu_2 \pi_2 = 0$

$$\pi_2 = \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} \pi_0$$

by mathematical induction, we can show that

$$\pi_n = \left(\frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} \right) \pi_0$$

Now we need $\sum_{n=0}^{\infty} \pi_n = 1$ this requires that

$$\sum_{n=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} < \infty$$

If this is true, then

$$\pi_0 = \left(1 + \sum_{n=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} \right)^{-1}$$

Example (6) : simple death with immigration.

$$\lambda_n = \lambda, n \geq 0$$

$$\mu_n = n\mu, n \geq 1$$

$$\text{so } \pi_n = \frac{\lambda^n}{n! \mu^n} \pi_0$$

$$\text{Let } P = \lambda/\mu \text{ Then}$$

$$\pi_n = P^n \pi_0 / n! \quad n=0,1,2,\dots$$

Since $\sum_{n=0}^{\infty} \pi_n = 1$, we must have $\pi_0 = e^{-P}$, so

$$\pi_n = \frac{P^n e^{-P}}{n!} \quad n=0,1,2,\dots$$

So $X(t) \rightarrow \text{Poisson}(P)$

Another important special case is the simple birth and death process

$$\begin{cases} \mu_n = \mu & n \geq 1 \\ \lambda_n = \lambda & n \geq 0 \end{cases}$$

STAT 833 Oct 30, 2008

Problems 6.9: 1, 2, 3, 8, 9

6.11: 1, 2, 3 6.15: 1, 3

Chapter 8 Random Processes

A random process is an indexed family of random variables, $X_t, t \in T$

We also write $X(t) t \in T$ for the same thing.

Typically, $T = \{0, 1, 2, \dots\}$ for discrete time or $T = [0, t_0]$ for continuous time.

Since a random variable is itself a function $X: \Omega \rightarrow \mathbb{R}$, so a stochastic process is a function

$$X: T \times \Omega \rightarrow \mathbb{R}.$$

$$x: (t, \omega) \rightarrow X_t(\omega)$$

for given $\omega \in \Omega$, $X_t(\omega)$ is a function of t , called the sample path.
represent full evolution

For given $t_1, t_2, \dots, t_n \in T$, the vector

$(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ has a joint distribution with joint distribution function $F_t(x)$, where $t = (t_1, t_2, \dots, t_n)$ and $x = (x_1, \dots, x_n)$

Here n is any positive integer.

The collection of all joint distributions

$\{F_t(x) : t, x \text{ vary, } n \text{ varies}\}$ is called the set of finite dimension distributions of X .

Finite dimensional distributions" fdd. for short.

8.2 Stationary processes

Definition: X is said to be **strongly stationary** if $(X_{t_1}, X_{t_2}, \dots, X_{t_n}) \sim (X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h})$ for all $n, t_1, t_2, \dots, t_n, h > 0$.

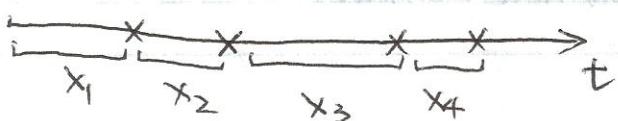
Definition: X is said to be **weakly stationary** or second order stationary if $E(X_{t_1}) = E(X_{t_2})$ for all $t_1, t_2 \in T$

& $\text{Cov}(X_{t_1}, X_{t_2}) = \text{Cov}(X_{t_1+h}, X_{t_2+h})$ for all $t_1, t_2, h > 0$

8.3 Renewal processes

These are generalization of the Poisson process to the case where the times between events are not exponential

Definition: A renewal process $N(t)$, $t \geq 0$ is a process such that
 $N(t) = \max \{n : T_n \leq t\}$ where T_n is the time of the n^{th} event
 with $T_0 = 0$ and $T_n = X_1 + \dots + X_n$, $n \geq 1$ where X_1, X_2, X_3, \dots are iid. non-negative random variables.



RE

$$T_n = \inf \{t : N(t) = n\}$$

$$X_n = T_n - T_{n-1} \quad \text{for all } n \geq 1$$

Special case: $X_i \sim \text{Exp}(\lambda)$, Then $N(t)$ is a Poisson process with rate parameter λ

Theorem: (a) $\lim_{t \rightarrow \infty} \frac{\mathbb{E}[N(t)]}{t} = \frac{1}{\mathbb{E}(X_1)}$

Renewal Theorem (b) Under an assumption

$$\lim_{t \rightarrow \infty} \mathbb{E}[N(t+h) - N(t)] = \frac{h}{\mathbb{E}(X_1)}$$

$$\textcircled{b} \Rightarrow \textcircled{a} \quad \mathbb{E}[N(k)] = \sum_{i=1}^k [N(i) - N(i-1)]$$

$$\text{As } k \rightarrow \infty \quad \textcircled{b} \quad \sum_{i=1}^k \frac{1}{\mathbb{E}(X_1)} = \frac{k}{\mathbb{E}(X_1)} \quad \begin{matrix} t = \alpha \\ h = 1 \end{matrix}$$

$$\text{So } \frac{\mathbb{E}[N(k)]}{k} \rightarrow \frac{1}{\mathbb{E}(X_1)} \text{ which is } \textcircled{a}$$

8.4 Queues

$Q(t) = \# \text{ of customers in line (queue) at time } t$

including the customer being served.

- customers enter as a renewal process
- They join a line to be served
- Service times are iid. random variables
- Can have several servers

8.5 Wiener process \rightarrow diffusions

STAT 833 Oct Nov 4 2008

Assigned problems (after test 1)

6.9: 1, 2, 3, 8, 9 6.11: 1, 2, 3 6.15: 1, 3

8.2: 1 8.3: 2 8.4: 1 8.5: 3

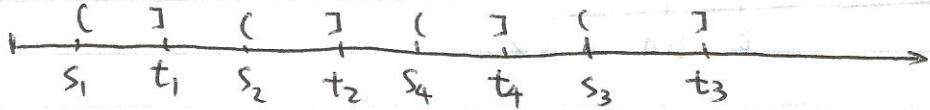
Chapter 13. Diffusions

Recall that the finite dimensional distributions (fdd) of a stochastic process $X(t)$, $t \geq 0$ are the joint distributions of any $X_{t_1}, X_{t_2}, \dots, X_{t_n}$ where

$0 \leq t_1 \leq \dots \leq t_n < \infty$ and where $n \geq 1$ is an integer.

Def'n. A stochastic process X_t , $t \geq 0$ is said to be Gaussian if every finite dimensional distribution is multivariate normal.

Def'n : $X_t, t \geq 0$ is said to have independent increments if $X_{t_i} - X_{t_j}$, where $s_i < t_i$ and $i=1, \dots, n$ are independent for all i whenever intervals $[s_i, t_i]$ are disjoint.



Def'n : $X_t, t \geq 0$ is said to have continuous sample paths if $X_t(w)$ is a continuous function of t for all $w \in \mathcal{W}$.

Slight Relaxation : $P(X_t \text{ continuous at } t) = 1$

Def'n : A Wiener process $W(t), t \geq 0$ is a Gaussian process such that

(a) W has independent increments

(b) $W(s+t) - W(s) \sim N(0, \sigma^2 t)$ for all $s, t \geq 0$ where $\sigma^2 > 0$ is a constant.

(c) W has continuous sample path.

We also usually suppose that $W(0) = w$ where w is fixed. Often $w = 0$

$$\mathbb{E}W(t) = \mathbb{E}W(0) + \mathbb{E}[W(t) - W(0)]$$

$$= w + 0 = w$$

Suppose $s \leq t$,

$$\begin{aligned} \text{Cov}[W(s), W(t)] &= \text{Cov}[W(s), W(s)] + \text{Cov}[W(s), W(t) - W(s)] \\ &= \text{Var}[W(s)] + \text{Cov}[W(s) - W(0), W(t) - W(s)] \\ &= \text{Var}[W(s)] = \sigma^2 s \end{aligned}$$

In general, $\text{Cov}[W(s), W(t)] = \sigma^2 \min(s, t)$

Def'n. W is called a standard Wiener process if $\sigma^2 = 1$ and $W = W(0) = 0$

Let us consider the transition probabilities of a standard Wiener process.

Let $F(t, y | s, x) = P[W(t) \leq y | W(s) = x]$ ($s \leq t$)

with density

$$f(t, y | s, x) = \frac{\partial}{\partial y} F(t, y | s, x) \quad (s \leq t)$$

Since $W(t) - W(s) \sim N(0, t-s)$ for $s \leq t$.

$$\text{so } f(t, y | s, x) = \frac{1}{\sqrt{2\pi(t-s)}} \exp\left\{-\frac{(y-x)^2}{2(t-s)}\right\}$$

It can be checked that this satisfies the forward equations.

$$\frac{\partial F(t)}{\partial t} = \frac{1}{2} \frac{\partial^2 F}{\partial y^2}$$

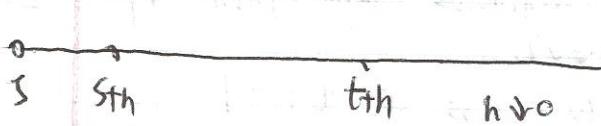
and the backwards equations.

$$\frac{\partial f}{\partial s} = -\frac{1}{2} \frac{\partial^2 f}{\partial x^2}$$

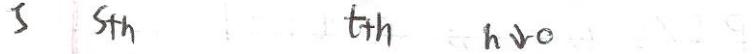
These equations come from an infinitesimal argument



Forward equations



Backwards equations



The infinitesimal properties as $h \rightarrow 0$ of a general Wiener process are

- ① $P(|W(t+h) - W(t)| > \varepsilon | W(t) = x) = o(h) \text{ as } h \rightarrow 0$
- ② $E[W(t+h) - W(t) | W(t) = x] = 0$
- ③ $E[(W(t+h) - W(t))^2 | W(t) = x] = \sigma^2 h + o(h)$

A diffusion is a stochastic process that looks "locally" over a small time interval like a Wiener process with drift.

Def'n: $D(t)$, $t \geq 0$ is said to be a diffusion process if

- ① $P(|D(t+h) - D(t)| > \varepsilon | D(t) = x) = o(h) \text{ as } h \rightarrow 0$
- ② $E[D(t+h) - D(t) | D(t) = x] = a(t, x)h + o(h)$
- ③ $E[(D(t+h) - D(t))^2 | D(t) = x] = b(t, x) \cdot h + o(h)$

Also, D has continuous sample paths with probability one.

Def'n. $F(t, y | s, x) = P(D(t) \leq y | D(s) = x)$

$$f(t, y | s, x) = \frac{\partial F(t, y | s, x)}{\partial y}$$

Forward equations

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial y} [a(t, y)f] + \frac{1}{2} \frac{\partial^2 f}{\partial y^2} [b(t, y)f]$$

Backward equations

$$\frac{\partial f}{\partial s} = -a(s, x) \frac{\partial f}{\partial x} - \frac{1}{2} b(s, x) \frac{\partial^2 f}{\partial x^2}$$

where $f = f(t, y | s, x)$

Standard Wiener process

$$a(t, x) = 0$$

$$b(t, x) = 1$$

General Wiener process

$$\alpha(t, x) = 0$$

Wiener process with drift

$$\alpha(t, x) = \mu$$

STAT 833 Nov 5 2008

13.7 Stochastic Calculus

W is differentiable nowhere with prob. 1.

A diffusion behaves locally in space and time like a Wiener process with drift

$$\Delta D_t \approx \mu \Delta t + \sigma \Delta W_t$$

$$\text{where } \Delta D_t = D_{t+\Delta t} - D_t$$

$$\Delta W_t = W_{t+\Delta t} - W_t$$

and where W is a standard Wiener process. Also

$$\mu = \mu(t, D_t) \quad \sigma = \sigma(t, D_t)$$

As $\Delta t \rightarrow 0$, it is natural to write

$$dD_t = \mu(t, D_t) dt + \sigma(t, D_t) \cdot dW_t$$

We can try to make sense of this by integrating.

$$D_t - D_0 = \int_0^t \mu(s, D_s) ds + \int_0^t \sigma(s, D_s) dW_s$$

Example: A certain diffusion satisfies

$$dX_t = b X_t \cdot dW_t$$

so that $\mu(t, X_t) = 0$ $\sigma(t, X_t) = b X_t$

$$X_t - X_0 = b \int_0^t X_s dW_s \quad \text{Geometric Wiener process}$$

Consider a diffusion built by a smooth transformation of W_t

$$X_t = f(W_t)$$

What are $\mu(t, X_t)$ and $\sigma(t, X_t)$?

By a Taylor expansion,

$$f(W_{t+\Delta t}) = f(W_t) + f'(W_t) \cdot \Delta W_t + \frac{1}{2} f''(W_t) (\Delta W_t)^2 + \dots$$

or

$$X_{t+\Delta t} - X_t = f'(W_t) \Delta W_t + \frac{1}{2} f''(W_t) (\Delta W_t)^2 + \dots$$

$$\text{or } \Delta X_t = f'(W_t) \Delta W_t + \frac{1}{2} f''(W_t) (\Delta W_t)^2 + \dots$$

If the standard theory of calculus worked here (it doesn't!!!) we would have $dX_t = f'(W_t) dW_t$ (Chain Rule)

This can't be right. We should get a drift term $\mu_C \cdot dt$

We threw it away when we threw away $\frac{1}{2} f''(w_t)(dW)^2$

$$E(dw_t)^2 = dt \Rightarrow (dw_t)^2 = dt.$$

"So" $dX_t = f'(w_t)dw_t + \frac{1}{2}f''(w_t)dt$

So X_t has local drift coefficient $\frac{1}{2}f''(w_t)$ and local diffusion coefficient $f'(w_t)$

Example: why standard calculus does not apply

Standard $\int_0^t w_s dw_s$ by parts $= w_s w_s \Big|_0^t - \int_0^t w_s \cdot dw_s$

$$\Rightarrow \int_0^t w_s dw_s = \frac{w_t^2 - w_0^2}{2} = \frac{w_t^2}{2} \quad \boxed{\times}$$

But this is wrong!

To see this is wrong, let $f(w_t) = w_t^2$

then $f'(w_t) = 2w_t$ and $f''(w_t) = 2$ So

$$d(w_t^2) = d[f(w_t)] = f'(w_t)dw_t + \boxed{\frac{1}{2}f''(w_t)dt}$$

Integrating both sides

$$d(w_t^2) = 2w_t \cdot dw_t + dt$$

$$w_t^2 = 2 \int_0^t w_s dw_s + t \quad \text{so}$$

Stratonovich integral
Itô integral

$$\int_0^t w_s dw_s = \frac{1}{2} w_t^2 - \frac{t}{2}$$

In this case, the chain rule is wrong, and the Riemann-Stieltjes definition of the integral cannot be correct.

The fixed version of the chain-rule is called "Itô's lemma".

Extra term Taylor $(dW)^2 = dt$

So how do we define

$$\int_0^t \psi_s dw_s ?$$

where ψ_s can be random also with infinite total variation.

$$\int_0^t \psi_s dw_s = \lim_{n \rightarrow \infty} \sum_{j=0}^n \psi_{\theta_j} [w_{j+1}\Delta - w_{j\Delta}] \quad \begin{array}{l} \text{Does not converge} \\ \text{with prob. 1.} \end{array}$$

where $\Delta = t/n$ and θ_j is chosen from the interval $[j\Delta, (j+1)\Delta]$

unlike the usual Riemann-Stieltjes integral, the value of the limit depends on how the point θ_j is chosen from the interval $[j\Delta, (j+1)\Delta]$.

$$\theta_j = (j + \frac{1}{2})\Delta \quad \text{Stratonovich integral} \quad (\text{ordinary chain rule})$$

$$\theta_j = j\Delta \quad (\text{left end point}) \quad \text{Itô integral} \quad (\text{Itô's lemma})$$

Chain rule with extra term

We only have

$$\sum_{j=0}^n \psi_{j\Delta} [W_{(j+1)\Delta} - W_{j\Delta}] \xrightarrow[\text{with prob 1}]{P, L_2} \int_0^t \psi_s dW_s \quad \text{Itô integral}$$

STAT 833 Nov 11 2008 prof. Small, Christopher.

Problems assigned after Test 1

6.9: 1, 2, 3, 8, 9

6.11: 1, 2, 3

8.2: 1

8.3: 2

8.4: 1

8.5: 3

New assigned problems

13.3: 1, 7

13.12: 1

13.1-3, 7, 8

Read there

Chapter 10 Renewal Processes

(1) Definition: Let X_1, X_2, \dots be an infinite sequence of i.i.d. nonnegative random variables. Let $T_n = X_1 + \dots + X_n$ for $n \geq 1$. Define the renewal process $N(t)$, $t \geq 0$ by

$$N(t) = \max\{n : T_n \leq t\}$$

(2) Theorem: $N(t)$ is honest iff. $E(X_1) > 0$. In other words,

$$P(N(t) < \infty) = 1 \quad \text{if and only if} \quad E(X_1) > 0$$

Note: $E(X_1) > 0$ iff. $P(X_1 = 0) \neq 1$

We shall assume that $N(t)$ is honest.

That is; $P(X_1 = 0) \neq 1$. A stronger useful condition is

$$P(X_1 = 0) = 0 \quad \text{which we now assume.}$$

Let F be the distribution function of X_1 . also of (X_2, X_3, \dots)

Let F_k be the distribution function of $T_k = X_1 + X_2 + \dots + X_k$

Then $X_1 = T_1$, so $F_1 = F$. Also,

$$T_{k+1} = T_k + X_{k+1}$$

$$F_{k+1}(x) = \int_0^x F_k(x-y) dF(y)$$

The probability function of $N(t)$ is given by

$$(5) \text{ Lemma: } P(N(t)=k) = F_k(t) - F_{k+1}(t)$$

$$\text{proof: } P(N(t)=k) = P(N(t) \geq k) - P(N(t) \geq k+1)$$

$$\begin{aligned} &= P(\max\{n \in \mathbb{N} : T_n \leq t\} \geq k) \\ &\quad - P(\max\{n \in \mathbb{N} : T_n \leq t\} \geq k+1) \\ &= P(T_k \leq t) - P(T_{k+1} \leq t) \\ &= F_k(t) - F_{k+1}(t) \end{aligned}$$

(6) Definition we define the renewal function $m(t)$ by

$$m(t) = \mathbb{E}[N(t)]$$

(7) Lemma

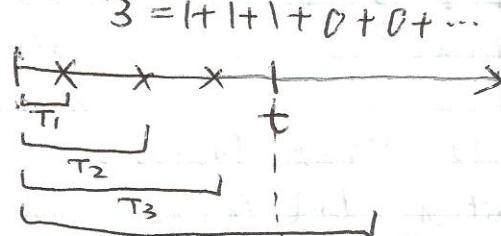
$$m(t) = \sum_{k=1}^{\infty} F_k(t)$$

$$\text{proof: } N(t) = \sum_{k=1}^{\infty} \mathbf{1}_{(T_k \leq t)}$$

$$\text{so } \mathbb{E}[N(t)] = \mathbb{E}\left(\sum_{k=1}^{\infty} \mathbf{1}_{(T_k \leq t)}\right)$$

$$= \sum_{k=1}^{\infty} \mathbb{E} \mathbf{1}_{(T_k \leq t)}$$

$$= \sum_{k=1}^{\infty} P(T_k \leq t) = \sum_{k=1}^{\infty} F_k(t) \quad \text{QED.}$$



(monotone convergence theorem)

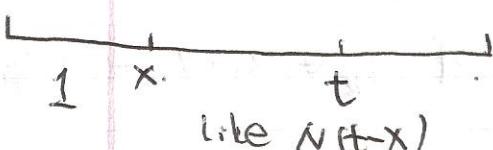
(8) Lemma

$$m(t) = F(t) + \int_0^t m(t-x) dF(x)$$

$$\text{proof: } m(t) = \mathbb{E}[N(t)]$$

$$= \mathbb{E}[\mathbb{E}[N(t) | X_1]]$$

$$\text{Now: } \mathbb{E}[N(t) | X_1=x] = \begin{cases} 0 & t < x \\ 1 + \mathbb{E}[N(t-x)] & t \geq x \end{cases}$$



$$\text{so } m(t) = \mathbb{E}[\mathbb{E}[N(t) | X_1]]$$

$$= \int_0^{\infty} \mathbb{E}(N(t) | X_1=x) dF(x)$$

$$\begin{aligned}
 &= \int_0^t [1 + \mathbb{E}[N(t-x)]] dF(x) \\
 &= \int_0^t [1 + m(t-x)] dF(x) \\
 &= F(t) + \int_0^t m(t-x) dF(x)
 \end{aligned}$$

QED.

Example: $x_1, x_2, \dots \sim \text{Exp}(\lambda)$

$F_k(x)$ is the distribution function for $T_k = x_1 + \dots + x_k$
 $\sim \text{Gamma}(k, \lambda)$

$$P(N(t)=k) = F_k(t) - F_{k+1}(t)$$

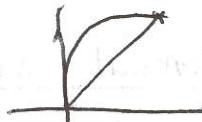
$$\begin{aligned}
 &= \int_0^t \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!} dx - \int_0^t \frac{\lambda^{k+1} x^k e^{-\lambda x}}{k!} dx \\
 &= \frac{\lambda^k e^{-\lambda t}}{k!} \quad (\text{prove this by part})
 \end{aligned}$$

$$\begin{aligned}
 m(t) &= \sum_{k=1}^{\infty} F_k(t) = \sum_{k=1}^{\infty} \int_0^t \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!} dx \\
 &= \int_0^t \lambda e^{-\lambda x} \sum_{h=1}^{\infty} \frac{(\lambda x)^{h-1}}{(h-1)!} dx \\
 &= \int_0^t \lambda e^{-\lambda x} \cdot e^{\lambda x} \cdot dx = \lambda t
 \end{aligned}$$

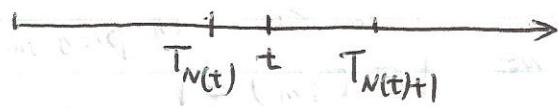
9 questions:

$$\text{Newmark} = 100 \sqrt{\frac{\text{mark}}{90}}$$

STAT 833 Nov 13th, B-Day 2008.



Sec 10.2 Limit theorems for renewal processes

We now consider the behaviour of $N(t)$ as $t \rightarrow \infty$ Let $M = \mathbb{E}(X_1)$ (1) Theorem: $\frac{N(t)}{t} \rightarrow \frac{1}{M}$ as $t \rightarrow \infty$, with probability oneProof: since $N(t) = \max\{n, T_n \leq t\}$ 

$$T_K = \sum_{i=1}^K X_i$$

We have $T_{N(t)} \leq t < T_{N(t)+1}$. Suppose $N(t) > 0$, then

$$\frac{T_{N(t)}}{N(t)} \leq \frac{t}{N(t)} < \frac{T_{N(t)+1}}{N(t)}$$

or.

$$\frac{T_{N(t)}}{N(t)} \leq \frac{t}{N(t)} < \frac{T_{N(t)+1}}{N(t)+1} \left[1 + \frac{1}{N(t)} \right]$$

Let $t \rightarrow \infty$ Then $N(t) \rightarrow \infty$ with prob. 1. because the X_k 's are $< \infty$ with prob. one.

$$\frac{T_{N(t)}}{N(t)} \rightarrow \lim_{k \rightarrow \infty} \frac{T_k}{k} \quad \left(= \lim_{k \rightarrow \infty} \frac{\sum_{i=1}^k X_i}{k} \text{ strong law of large numbers} \right)$$

SLLN.
iid. finite Expectation.

so $\lim_{T \rightarrow \infty} \frac{T_{N(t)}}{N(t)} = \mu$ with prob. 1.

also $\lim_{T \rightarrow \infty} \frac{T_{N(t)+1}}{N(t)+1} = \mu$ with prob. one

$$\lim_{T \rightarrow \infty} \left(1 + \frac{1}{N(t)} \right) = 1 \text{ with prob. one}$$

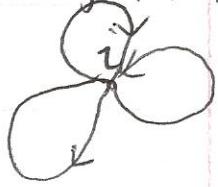
Therefore

$$\mu \leq \lim_{t \rightarrow \infty} \frac{t}{N(t)} \leq \mu$$

with prob. one.

Thus, $\lim_{t \rightarrow \infty} \frac{N(t)}{t} = \frac{1}{\mu}$ with prob. one

Ergodic MC in discrete time



$N(n)$ = # of returns to state i up to time n .

Excursions from state i are iid. with finite expectation

$$\frac{N(n)}{n} \rightarrow \pi_i \quad \text{Renewal idea } \frac{N(n)}{n} \rightarrow \frac{1}{\mu_i}$$

Recall that $m(t) = E(N(t))$

c) Elementary Renewal Theorem

$$\frac{m(t)}{t} \rightarrow \frac{1}{\mu} \text{ as } t \rightarrow \infty$$

Note: This "looks" obvious. because

$$\frac{N(t)}{t} \rightarrow \frac{1}{\mu} \text{ with prob. one. so}$$

$$E\left[\frac{N(t)}{t}\right] \rightarrow \frac{1}{\mu} \text{ or } \frac{m(t)}{t} \rightarrow \frac{1}{\mu}$$

In general

RV. $Y_m \rightarrow C$ with prob. 1



$$E(Y_m) \rightarrow C$$

even if $E(Y_m)$ exists

$$\text{eg. } Y_m = \begin{cases} 0 & 1 - 1/m \\ m^2 & \text{with prob. } 1/m \end{cases}$$

$Y_m \rightarrow 0$ with prob one

$$E(Y_m) = 1$$

To prove the elementary renewal thm. we need a definition and a result.

Def'n. Let X_1, X_2, \dots be a sequence of random variables.

A positive integer valued random variable M is called a stopping time if the event $\{M \leq m\}$ can be determined by the values of X_1, \dots, X_m .

For example

$M = \min \{ m : X_m \geq a \}$ is a stopping time. For example.
 $M \leq 2$ iff. $X_1 > a$ or $X_2 > a$.

However,

$M' = \min \{ m : X_{m+1} > a \}$ is NOT a stopping time.

$M' \leq 2$ iff. $X_1 > a, X_2 > a, X_3 > a$

Now consider a renewal process $N(t)$ with inter-arrival times

X_1, X_2, \dots . Define

$$M = N(t) + 1$$

Then M is a stopping time because

$$M \leq k \text{ iff } N(t) \leq k-1$$

iff $T_k > t$ iff.

$$\sum_{i=1}^k X_i > t$$

(*) Lemma (Wald's Equation)

Let M be a stopping time for a sequence $X_n, n \geq 1$ of iid. random variables, then

$$\mathbb{E} \left[\sum_{i=1}^M X_i \right] = \mathbb{E}[X_1] \cdot \mathbb{E}[M]$$

$$\text{Proof: } \sum_{i=1}^M X_i = \sum_{i=1}^{\infty} X_i \mathbf{1}_{(M \geq i)}$$

$$\begin{aligned} \text{So } \mathbb{E} \left(\sum_{i=1}^M X_i \right) &= \mathbb{E} \left(\sum_{i=1}^{\infty} X_i \mathbf{1}_{(M \geq i)} \right) \\ &= \sum_{i=1}^{\infty} \mathbb{E}(X_i \mathbf{1}_{(M \geq i)}) \end{aligned}$$

prob. 5.6.2. in textbook

dominating convergence

$$\sum \mathbb{E} = \mathbb{E} \sum$$

$$\text{Now } \mathbf{1}_{(M \geq i)} = \mathbf{1} - \mathbf{1}_{(M < i)}$$

$$= 1 - \mathbf{1}_{(M \leq i-1)}, M \text{ stopping time}$$

Then event $M \leq i-1$ is determined by X_1, \dots, X_{i-1} which are independent of X_i , so

$\mathbf{1}_{(M \geq i)}$ is independent of X_i .

$$\text{So } \sum_{i=1}^{\infty} \mathbb{E}(X_i \mathbf{1}_{(M \geq i)}) = \sum_{i=1}^{\infty} \mathbb{E}(X_i) \mathbb{E} \mathbf{1}_{(M \geq i)}$$

$$= \sum_{i=1}^{\infty} \mathbb{E}(X_i) \cdot P(M \geq i) = \sum_{i=1}^{\infty} \mathbb{E}(X_i) P(M \geq i)$$

$$= E(X_1) \cdot \sum_{i=1}^{\infty} P(M \geq i) = E(X_1) \cdot E(M)$$

QED

$E[E(\sum_i X_i | M)]$ Does not work here.
function of X_i 's. of Towel's Law

Sketch of proof of Elementary Renewal Theorem.

Use Wald's equation, with $M = N(t) + 1$. Then

$$E[\sum_{i=1}^{N(t)+1} X_i] = E(X_1) E[N(t)+1]$$

$$\text{or } E[T_{N(t)+1}] = u[m(t)+1]$$

We saw earlier that $t < T_{N(t)+1}$. So

$$t < E[T_{N(t)+1}] \text{ Thus}$$

$$t < u[m(t)+1]$$

This becomes

$$\frac{1}{u} < \frac{m(t)}{t} + \frac{1}{t} \quad \text{or} \quad \frac{m(t)}{t} > \frac{1}{u} - \frac{1}{t} \text{ for all } t.$$

so for all $\epsilon > 0$

$$\frac{m(t)}{t} > \frac{1}{u} - \epsilon \text{ for sufficiently large } t$$

Also it can be shown that

$$\frac{m(t)}{t} < \frac{1}{u} + \epsilon \text{ for sufficiently large } t. \text{ See page 420}$$

so

$$\lim_{t \rightarrow \infty} \frac{m(t)}{t} = \frac{1}{u}. \quad \text{QED.}$$

STAT 833 Nov 18, 2008.

Assigned problems 10.1: 1, 3. 10.2: 1, 3. 10.6: 1, 4

No class Thursday Lecture Thursday next week
(Dec 2) Ch 11.

Queuing theory.

i. 1 single server queues

To build a single server queue, we construct two sequences $X_n, n \geq 1$

and $S_n, n \geq 1$ of independent positive random variables.

The $X_n, n \geq 1$ are iid. with common distribution function F_x . These are the interarrival times of customers. The $S_n, n \geq 1$ are iid. with distribution function F_s , these are the service times of customers.

So the n^{th} customer arrives at time $T_n = \sum_{i=1}^n X_i$ and joins the line. When each customer gets to the front of the line, he/she is served with service time S_n .

$Q(t) = \# \text{ of customers in the line at time } t$, including the customer being served.

Typically, $Q(0) = 0$.

Basic Questions.

- ① When is $Q(t)$ a MC?
- ② When does $Q(t)$ have a limit distribution as $t \rightarrow \infty$? How can we modify $Q(t)$ to make this distribution "Optimal".
- ③ When does $Q(t) \rightarrow \infty$ as $t \rightarrow \infty$?

I.e. server cannot cope.

Queues are classified using the notation $A/B/S$

- A describes the interarrival time distribution F_x .

- B describes the service time distribution F_s

- S = # of servers. Here $S=1$

M: (Markovian) or exponential $\text{Exp}(\lambda)$

λ is the rate parameter.

D: (Deterministic) time is a constant, say d.

G: general, fixed, Unspecified distribution.

① M/M/1 $(\text{M}(\lambda)/\text{M}(\mu)/1)$

customers arrive as a Poisson process with intensity λ , and each is served with service time that is $\text{exp}(\mu)$. $Q(t)$ is a continuous time MC.

② M/D/1 $(\text{M}(\lambda)/\text{D}(d)/1)$

Customers arrive as a Poisson process with intensity λ , Each is served in a time of length served d, say.

Definition: We define the traffic intensity of $Q(t)$ to be

$$\rho = E(S)/E(X)$$

II.2 M/M/1 Queue

here, $Q(t)$ is a continuous time MC, a birth death process with
 $\lambda_n = \lambda$ for all n .

for $n \geq 1$ $\mu_n = \mu$, also $\mu_0 = 0$

$$P = E(s)/E(x) = \frac{\lambda_0}{\lambda_0} = \lambda/\mu \quad Q(0) = 0.$$

$$P_n(t) = P(Q(t)=n)$$

The solution to the forward equations for $P_n(t)$ is given on page 442.

Recall the stationary distribution:

$$\pi_n = \lim_{t \rightarrow \infty} P(Q(t)=n)$$

$$\pi_n = \frac{\lambda_0 \lambda_1 \dots \lambda_{n-1}}{\mu_1 \dots \mu_n} \cdot \pi_0 \quad (n \geq 1) \quad (\text{birth-death formula})$$

$$= \frac{\lambda^n}{\mu^n} \cdot \pi_0 \quad (\text{for M/M/1 Queue})$$

$$= p^n \cdot \pi_0$$

$$\pi_0 = \left(\sum_{n=0}^{\infty} p^n \right)^{-1} = \left(\frac{1}{1-p} \right)^{-1} = 1/(1-p)$$

so $\pi_n = p^n/(1-p)$ $n \geq 0$ geometric distribution. if $0 < p < 1$

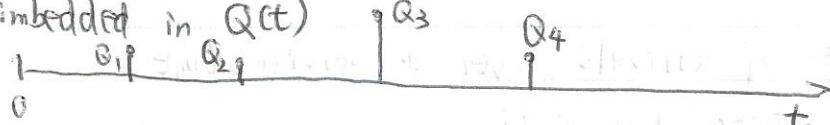
if $p \geq 1$ there is no stationary distribution.

11.3 M/G/1 Queue Not a MC Nov 27, 2008

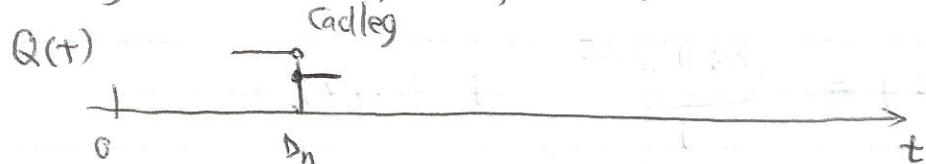
This queue is not a MC, unless $G = M$.
 Customers line up as a Poisson process of intensity λ and are served with service times having i.i.d. distributions, which are general in nature.

$Q(t) = \# \text{ of people in line at time } t$.

To use MC theory, we find a discrete time M.C. Q_1, Q_2, Q_3 which is imbedded in $Q(t)$



Let D_n be the time of departure of the n -th customer from the line.
 Immediately after the finish of service)



Let $Q(D_n)$ be the number of customers in line after the n -th customer has left. Define

$$Q(D) = \{Q(D_1), Q(D_2), \dots\}$$

Then $Q(D)$ is a discrete time stochastic process.

We shall see that $Q(D)$ is a M.C.

Suppose $Q(D_n) > 0$, then

$$Q(D_{n+1}) = (\# \text{ of people in line at time } D_n)$$

$$+ (\# \text{ of people who arrive between } D_n \text{ and } D_{n+1})$$

$$- (\text{a count of 1: } (n+1)\text{st customer who leaves})$$

$$= Q(D_n) + U_n - 1, \text{ where}$$

$$U_n = \# \text{ of people arriving between } D_n \text{ and } D_{n+1},$$

when $Q(D_n) = 0$:

$$Q(D_{n+1}) = (\# \text{ of people who arrive between } D_n \text{ and } D_{n+1}) - 1$$

$$= U'_n - 1$$

In this case, we put a prime on U_n because $D_{n+1} - D_n$ is not "service time". So $U_n \neq U'_n$

$U'_n = (\text{Count 1 for arrival of customer } n+1) + (\# \text{ of customers between arrival of customer } n+1 \text{ and departure of customer } n+1)$

$$\sim 1 + U_n$$

so $Q(D_n) = c$ then $Q(D_{n+1}) \sim (1 + U_n) - 1 = U_n$

Let us summarise

$$Q(D_{n+1}) = \begin{cases} Q(D_n) + U_n - 1 & \text{if } Q(D_n) > 0 \\ U_n & \text{if } Q(D_n) = 0 \end{cases}$$

in a sense of distribution

where U_n is the # of arrivals over a service time.

Let's calculate the distribution of U_n .

Let S be any service time. Condition on $S=s$, the # of arrivals is Poisson (λs)

$$P(U=j | S=s) = \frac{(\lambda s)^j e^{-\lambda s}}{j!} \quad j=0, 1, 2, \dots$$

$$\text{or } P(U=j | \underline{S=s}) = \frac{(\lambda s)^j e^{-\lambda s}}{j!} \quad j=0, 1, 2, \dots$$

so for $j=0, 1, 2, \dots$

$$\begin{aligned} P(U=j) &= E[P(U=j | S)] \\ &= E\left[\frac{(\lambda s)^j e^{-\lambda s}}{j!}\right] \end{aligned}$$

where S has a known general distribution ($M/G/1$)

(4) Theorem: The sequence $Q(D_n)$ $n \geq 1$ is a discrete time MC with transition matrix

$$P = \begin{pmatrix} \delta_0 & \delta_1 & \delta_2 & \delta_3 & \dots \\ \delta_0 & \delta_1 & \delta_2 & \delta_3 & \dots \\ \delta_0 & \delta_1 & \delta_2 & \delta_3 & \dots \\ 0 & \delta_0 & \delta_1 & \delta_2 & \dots \\ 0 & 0 & \delta_0 & \delta_1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad \text{where } \delta_j = E\left[\frac{(\lambda s)^j e^{-\lambda s}}{j!}\right]$$

and S is a typical service time.

proof: That $Q(D_n)$ is a MC follows from the formula for

$Q(D_{n+1})$ in terms of $Q(D_n)$

$$P_{0j} = P(Q(D_{n+1})=j | Q(D_n)=0)$$

$$= P(U_n=j) = \delta_j$$

$$\begin{aligned}
 P_{ij} &= P(Q(D_{n+1})=j | Q(D_n)=i) \\
 &= P(U_n + Q(D_n) - 1 = j | Q(D_n)=i) \\
 &= P(U_n = j | Q(D_n)=i) \\
 &= P(U_n = j | Q(D_n)=i) = \underbrace{P(U_n = j)}_{\text{independent}} = \delta_{ij}
 \end{aligned}$$

(13) $P_{2j} = P(Q(D_{n+1})=j | Q(D_n)=2)$

$$\begin{aligned}
 &= P(U_n + 2 - 1 = j | Q(D_n)=2) = P(U_n = j-1) = \delta_{j-1,0} \\
 &\quad \dots \text{etc.} \quad \text{QED.}
 \end{aligned}$$

(5) The traffic intensity is $P = \frac{\mathbb{E}(S)}{\mathbb{E}(X)} = \lambda \mathbb{E}(S)$

Theorem: (a) If $P < 1$, then $Q(D)$ is ergodic with a unique stationary distribution $\pi = (\pi_0, \pi_1, \pi_2, \dots)$ with generating function

$$G(s) = \sum_{j=0}^{\infty} \pi_j s^j = (1-P)(s-1) \frac{M_S[\lambda(s-1)]}{s - M_S[\lambda(s-1)]}$$

where M_S is the moment generating function of S .
That is, $M_S[\lambda(s-1)] = \mathbb{E}[e^{\lambda(s-1)S}]$.

(b) If $P > 1$, then $Q(D)$ is transient.

(c) If $P=1$, then $Q(D)$ is null persistent/recurrent.

If $P < 1$,

$$\lim_{t \rightarrow \infty} P(Q(t)=n) = \pi_n$$

Why? Because D_n is a typical large time when n is large.



End of Lectures .