updated: 30 March 2009

Applied Probability 2

• Unit organiser: Dr. Stanislav Volkov **(office SM 3.13)**

• Lectured by: Prof. John McNamara and Dr. Stanislav Volkov

• Timetable:

Monday	**10^{00} am** ***and*** **11^{10} am (problems class)**
Tuesday	**10^{00} am**
Wednesday	**10^{00} am**

• Prerequisites: First year Core Mathematics calculus, analysis, linear algebra and Probability.

• Applied probability 2 is a prerequisite for Queueing Networks, Probability 3, and also Financial Mathematics, and is relevant to other Level 3 probabilistic units.

Assessment Methods

- **10% from the weekly homework**, **assessed on all** (approximately) 10 weekly homework marks. Each homework will be marked out of 10. The homework must be handed in by the specified time; otherwise it will be given a mark of 0. Medical or other special grounds for late or non-submission of homework must be discussed with the Unit Organiser.

- **90% from the examination**: 2 ½ hours in May/June. FIVE questions, a candidate's FOUR best answers will be used for assessment

Texts

- Taylor, H.M. & Karlin, S. ***An Introduction to Stochastic Modelling*** (3rd Ed.) (Academic Press) - **Required text**

- Grimmett, G.R. & Stirzaker, D.R. ***Probability and Random Processes***. (OUP)

0. Review of probability

Random experiment

ω - elementary outcome, or <u>sample point</u>
Ω - set of all elementary outcomes (<u>sample space</u>)

Let $\Im$ be a collection of subsets of Ω. We call it a *sigma-field*, if

1) $\varnothing \in \Im$ and $\Omega \in \Im$.
2) If $A_1, A_2, \ldots \in \Im$ then $\cup_i A_i \in \Im$.
3) If $A \in \Im$ then $A^c \in \Im$.

Examples: (1) $\Im = \{\varnothing, \Omega\}$, (2) $\Im$=all subsets of Ω (usually when Ω is finite)

A subset $A \subseteq \Omega$ of elementary outcomes is called an event, if $A \in \Im$. Generally, sigma-fields are beyond the scope of the course, and you can assume that any $A \subseteq \Omega$ is an event.

Interpretation – say event A occurs if outcome $\omega \in A$

Two events A, B:

$A \cup B$ occurs $\Leftrightarrow$ $\omega \in A$ or $\omega \in B$ $\Leftrightarrow$ A occurs or B occurs
$A \cap B$ occurs $\Leftrightarrow$ $\omega \in A$ and $\omega \in B$ $\Leftrightarrow$ A occurs and B occurs

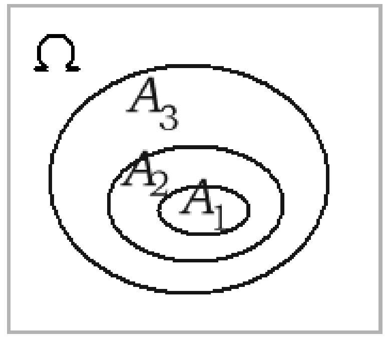
To each event A assign $P(A)$, which is probability if (axioms):

1) $0 \le P(A) \le 1$
2) $P(\Omega)=1$
3) For any (possibly infinite) sequence of events $A_1, A_2, \ldots$ such that $A_i \cap A_j = \varnothing$ whenever $i \neq j$ we have $P(\cup A_i) = \Sigma_i P(A_i)$

$(\Omega, \Im, P)$ is called *a probability space*.

Increasing and decreasing sequences of events

Let A_1 , A_2 ,…be a sequence of events such that $A_n \subseteq A_{n+1}$ for $n=1,2,3\ldots$ Then this sequence is said to be increasing



Note $A_n \subseteq A_{n+1} \quad \Leftrightarrow \quad \omega \in A_n \Rightarrow \omega \in A_{n+1}$
$\Leftrightarrow \quad A_n$ occurs $\Rightarrow A_{n+1}$ occurs

Example (a coin is tossed infinitely many times):
Let A_n= "from the n^{th} toss onwards all tosses give a head"
Then $\omega \in A_n$ = tosses n, n+1, n+2,… all heads $\Rightarrow$ tosses n+1, n+2,… all heads = $\omega \in A_{n+1}$
So $A_n \subseteq A_{n+1}$

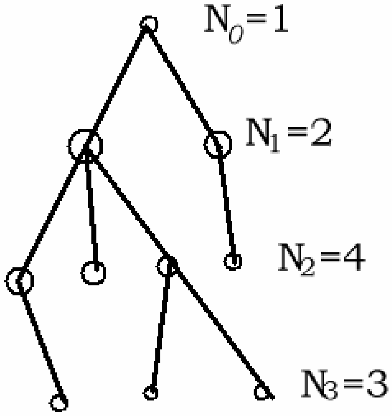
Interpretation in terms of a sample space:

Sample point $\omega=(x_1 ,x_2 ,x_3 ,\ldots.)$ where $x_i=1$ if the i-th toss is a head and $x_i=0$ otherwise.

$\Omega=\{\omega=(x_1 ,x_2 ,x_3 ,\ldots)$ where each $x_i=0$ *or* $1\}$

Events:
$A_1=\{(1,1,1,\ldots)\}$ – one sample point
$A_2=\{(0,1,1,\ldots) ; (1,1,1,\ldots)\}$ – two sample points
$A_3=\{(0,1,1,\ldots) ; (1,1,1,\ldots); (0,0,1,\ldots) ; (1,0,1,\ldots) \}$ – four sample points
(can see indeed $A_n \subseteq A_{n+1}$)
Let $A = \cup_{i=1}^{\infty} A_i$ be the event "at least one A_i occurred" i.e. from some time onward all heads

Example [population growth]



N_0 initial population size; N_k size after k generations.
Let $A_k=\{N_k=0\}$ – population extinct by generation k.
Since $N_k=0$ implies $N_{k+1}=0$ again $A_k \subseteq A_{k+1}$
Let $A = \cup_{i=1}^{\infty} A_i$ be the event that the population eventually becomes extinct.

Theorem [continuity of probability]
Suppose A_1 , A_2 ,…is a sequence of increasing events and let

$$A = \cup_{i=1}^{\infty} A_i$$

Then $P(A)= \lim_{n\to\infty} P(A_n)$

(Note that $A_n = \cup_{i=1}^{n} A_i$ and we can think of A as a limit of A_n)

Proof: Let $D_1=A_1$ and let $D_n=A_n \setminus A_{n-1}$
Then $D_i \cap D_j=\varnothing$ if $i\neq j$ and also $A_n=\cup_{i=1}^{n} D_i$ and $A=\cup_{i=1}^{\infty} D_i$
By axioms of probability for disjoint events D_j we have

$$P(A_n)=\Sigma_{i=1}^{n} P(D_i) \quad \text{and} \quad P(A)=\Sigma_{i=1}^{\infty} P(D_i)$$

Thus $P(A)=\Sigma_{i=1}^{\infty} P(D_i)= \lim_{n\to\infty} \Sigma_{i=1}^{n} P(D_i) = \lim_{n\to\infty} P(A_n)$
by definition of an infinite sum. *QED*

Application [population growth]
P(pop. becomes extinct)=$\lim_{n\to\infty}$ P(extinct by generation n)

Definition:

A sequence B_1 , B_2 ,…of events is said to be decreasing

if $B_{n+1} \subseteq B_n$ for $n=1,2,3\ldots$

Theorem [continuity of probability]

Suppose B_1 , B_2 ,…is a sequence of decreasing events and let

$$B = \cap_{i=1}^{\infty} B_i$$

Then $P(B)= \lim_{n\to\infty} P(B_n)$

Proof: Set $A_n=(B_n)^c$. Then A_j's are an increasing sequence. Also

$$A = \cup_{i=1}^{\infty} A_i = \cup_{i=1}^{\infty} (B_i)^c = (\cap_{i=1}^{\infty} B_i)^c = B^c$$

Thus by previous theorem

$$P(A)=P(B^c) = \lim P(A_n) = \lim P(B_n^c)$$

But $P(B^c)=1-P(B)$ and $P(B_n^c)=1-P(B_n)$, so

$$1-P(B)=1- \lim P(B_n)$$

whence this Theorem follows.

QED

Example [urn]
Initially there is one white and one red ball.
A ball is chosen at random and returned to the urn alongside with (always) extra red ball.

Thus when the n^{th} ball is chosen, there are n red and 1 white ball, hence

$$P(n^{th} \text{ chosen ball is red}) = \frac{n}{n+1}$$

Question: What is the probability a white ball is never chosen? (i.e. all chosen balls are red)

Answer: Let B_n be the event that the first n balls are red.
Let $B = \cap_{i=1}^{\infty} B_i$ be the event of interest (= none white).
Note that "B_{n+1} occurs" implies "B_n occurs", that is

$$\omega \in B_{n+1} \Rightarrow \omega \in B_n$$

Hence B_n's form a decreasing sequence and by continuity of probability

$$P(\text{all red}) = P(B) = \lim_{n\to\infty} P(B_n)$$

Meanwhile $P(B_n) = \frac{1}{2}\frac{2}{3}\frac{3}{4}\cdots\frac{n}{n+1} = \frac{1}{n+1}$

Consequently $P(\text{all red}) = \lim_{n\to\infty} 1/(n+1) = 0$ that is with probability 1 at least one white ball is chosen (note: can show that actually white is chosen infinitely many times!)

Random variables

A random variable X is a *measurable* function
from Ω onto $R=(-\infty,+\infty)$, $X: \Omega \to R$

meaning that for any numbers a and b such that $a<b$
$\{a<X\leq b\}\in\Im$, that is, $\{a<X\leq b\}$ is an event (in fact, then we can show that $\{a<X<b\}$, $\{X\leq b\}$, etc. are all events) – beyond the scope of this course.

Note: $\{a<X\leq b\}$ is a shortcut for $\{\omega\in\Omega: a<X(\omega)\leq b\}$

The (cumulative) distribution function (c.d.f.) of a random variable X is

$$F_X(b)=P(X\leq b)= P(\omega\in\Omega: X(\omega)\leq b)$$

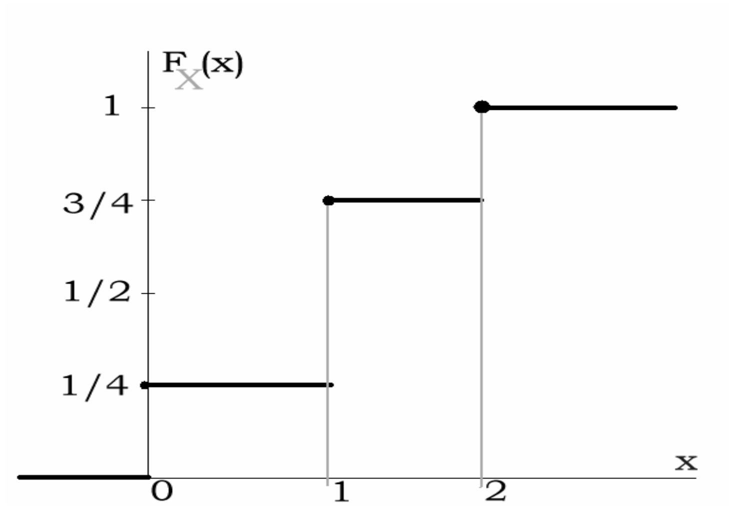
<u>Example</u> [fair coin tossed 2 times]

$\omega=(x_1,x_2)$ where each x_j is 0 if i^{th} toss is a tail and $=1$ if head.

$\Omega=\{(x_1,x_2): x_j=0 \text{ or } 1, j=1,2\}$ – 4 sample points altogether.

Let $X(\omega)=x_1+x_2$ – the total number of heads, in fact has Bin(2, ½) distribution

$$F(x)=\begin{cases} 0, & x<0 \\ 1/4, & 0\leq x<1 \\ 3/4, & 1\leq x<2 \\ 1, & 2\leq x \end{cases}$$



Lemma

Let X be a random variable. Let $F(b)=F_X(b)$ be the cdf of X.
Then for any $b\in R$

- $\lim_{n\to\infty} F(b-1/n)=P(X<b)$
- $\lim_{n\to\infty} F(b+1/n)=P(X\le b)=F(b)$

(that is F is right-continuous with left limits)

Proof of (a): Let $A_n=\{X\le b-1/n\}$ so that $P(A_n)= F(b-1/n)$.
The sequence A_n is an *increasing* sequence of events:

$\omega\in A_n \Leftrightarrow X(\omega)\le b-1/n \quad \Rightarrow X(\omega)\le b-1/(n+1) \Leftrightarrow \omega\in A_{n+1}$

Let $A = \cup_{i=1}^{\infty} A_i$

Note that $\omega\in A \Leftrightarrow X(\omega) \le b-1/n$ for *some* n which is the same as $X(\omega)<b$!

Thus $A=\{X<b\}$

By continuity of probability

$$P(X<b)=P(A)=\lim_{n\to\infty} P(A_n)= \lim_{n\to\infty} F(b-1/n)$$

Proof of (b): Let $B_n=\{X\le b+1/n\}$ … etc.

Further examples of random variables:

A fair coin is tossed 100 times

$\omega=(x_1,x_2,\ldots,x_{100})$
where each x_j is 0 if i^{th} toss is a tail and $=1$ if head.

Ω has 2^{100} equally likely sample points

Set $X_j(\omega)=x_j$ = the outcome of the j-th toss.

The event

$\{X_3=1\}=\{\omega\in\Omega:\ \omega=(x_1,x_2,1,x_4,\ldots,x_{100})$ each $x_j=0$ or $1\}$

has 2^{99} sample points

$P(X_3=1)= \dfrac{2^{99}}{2^{100}}=\dfrac{1}{2}$ that is $X_3\sim\text{Bernoulli}(\tfrac{1}{2})$

Similarly can show that $X_1,X_2,\ldots,X_{100}$ are **independent identically distributed** (*iid*) Bernoulli(½) random variables.

remember this acronym!

Stochastic process

For any set $\Delta \subseteq R$ collection $\{X_t, t \in \Delta\}$ of random variables is called a stochastic process.
Think of X_t as state of some system at time t. *In our course*, this will often be non-negative integers, i.e. $\Delta = \{0,1,2,\ldots\}$ (or some subset of this set) – discrete time stochastic process.

<u>Example</u> [fair coin is tossed 100 times]

Let $X_t(\omega)$ be the outcome of the t^{th} toss, $t=1,2,\ldots 100$
Then $\{X_t, t=1,\ldots, 100\}$ is a stochastic process.

Now let $Y_0=0$ and for $t=1,2,\ldots,100$ let $Y_t=X_1+X_2+\ldots+X_t$
$\{Y_t, t=0,1,\ldots, 100\}$ is also a stochastic process

Note that Y_t is the total number of heads obtained by time t.

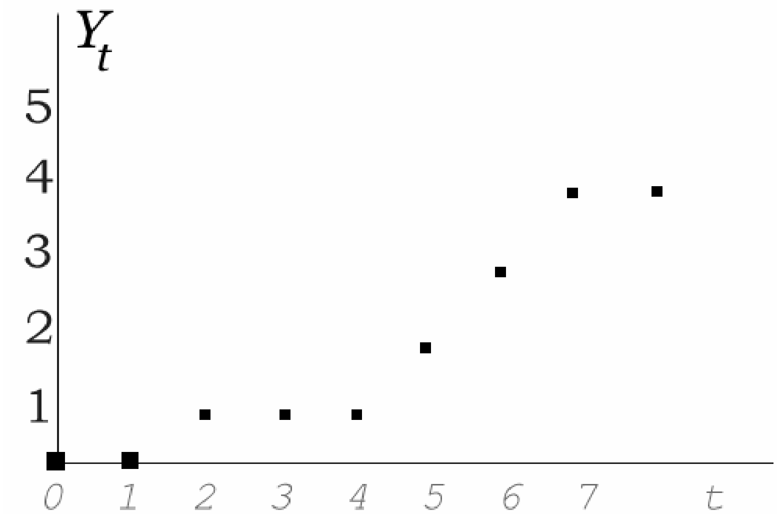
Two views of the stochastic process:

$X: \Delta \times \Omega \to R$

- for a fixed t, it is a random variable $Y_t(\bullet)$. In our case, Y_t has a Binomial(t, ½) distribution, that is for $k=0,1,\ldots, t$

$$P(Y_t = k) = \binom{t}{k}\left(\frac{1}{2}\right)^k \left(1-\frac{1}{2}\right)^{t-k}$$

- for a fixed ω, it is a function of t, $Y_\bullet(\omega)$ (called *sample path*, or *realization of the process*). For example, suppose $\omega=(0,1,0,0,1,1,1,0,1,\ldots)$. Then:



Definition: A stochastic process $\{X_t, t=0,1,\ldots, \}$ is called *Markov* process, if for every times t and s such that $s>t$ and any collection of states $(x_1,x_2,\ldots)$

$$P(X_s=x_s \mid X_t=x_t, X_{t-1}=x_{t-1},\ldots, X_0=x_0)= P(X_s=x_s \mid X_t=x_t)$$

Examples:

Both X_t and Y_t in the examples above.

Counterexample:

Let X_0 and X_1 be two independent random variables. For $t=2,3,\ldots$ let $X_t=X_{t-2}$, so that

$X_1= X_3= X_5= X_7=\ldots$ and $X_2= X_4= X_6= X_8=\ldots$

Then the distribution of X_{t+1} given all the past, is different from the distribution of X_{t+1} given X_t (of the latter it is actually independent!)

The strong law of large numbers

First look at coin tossing. Y_{100} is total number of heads in 100 tosses. Using binomial, get $P(40 \leq Y_{100} \leq 60) \approx 0.95$, that is of 2^{100} sample points of Ω about 95% are such that Y_{100} is between 40 and 60.

Now look at infinite number of tosses of a fair coin.
$\omega = (x_1, x_2, \ldots, x_n, \ldots)$
where each x_j is 0 if j^{th} toss is a tail and =1 if head.

Ω has uncountable many sample points

Set $X_j(\omega) = x_j$ = the outcome of the j^{th} toss (1 if head, 0 if tail)

Then there exists (*essentially*) unique prob. measure on Ω such that $X_1, X_2, \ldots$ are iid Bernoulli(½) ← Kolmogorov's existence theorem (beyond…)

Again, set $Y_0 = 0$ and for $t = 1, 2, \ldots$ set $Y_t = X_1 + X_2 + \ldots + X_t$

Look at the event $A = \{\omega \in \Omega : Y_t(\omega)/t \to ½ \text{ as } t \to \infty\}$

Can show that $P(A) = 1$, that is $P(\lim_{t\to\infty} Y_t/t = ½) = 1$.

Note that "$\lim_{t\to\infty} Y_t/t = ½$" in general is *not* true, e.g. take $\omega = (0,0,0,\ldots)$, but the *probability* of these "bad" sample points is zero.

We write this as $\frac{Y_t}{t} \xrightarrow{a.s.} \frac{1}{2}$ ("almost surely")

Convergence almost surely ("a. s.")

Let $X_1, X_2, \ldots$ be a sequence of random variables and X be another random variable.

Let $A = \{\omega \in \Omega: X_t(\omega) \to X(\omega) \text{ as } t \to \infty\}$

We say that $X_t \to X$ a.s. if $P(A)=1$

<u>Example</u> [strong law of large numbers]:

<u>Theorem</u>: Let $X_1, X_2, \ldots$ be a sequence of *iid* random variables with finite mean $E(X_i)=\mu$. Then

$$\frac{X_1 + X_2 + \cdots + X_N}{N} \xrightarrow{a.s.} \mu$$

when $N \to \infty$

Proof is beyond the scope (see Grimmett & Stirzaker p. 329 or sign in for *Probability 3*)

Example [construction of uniform [0,1] random variable]

Let $X_1, X_2, \ldots$ be a sequence of iid random variables with

$P(X_j=0)= P(X_j=1)=\ldots P(X_j=9)=\frac{1}{10}$

Then the sample space consists of $\omega=(x_1,x_2,\ldots,x_n,\ldots)$ where each x_j is $0,1,2,\ldots$ or 9 (think of digits of a number between 0 and 1, e.g. $\underline{.14159265\ldots}$ - *what is it?*)

For $n=1,2,\ldots$ let $Y_n=10^{-1} X_1+10^{-2} X_2+\ldots+10^{-n} X_n$

and look at stochastic process $\{Y_n, n=1,2,\ldots\}$

Observe that
Y_1 is equally likely to take any value $\in \{0, 0.1, 0.2, \ldots, 0.9\}$
Y_2 is equally likely … $\{0, 0.01, 0.02, \ldots, 0.99\}$
Y_3 is equally likely … $\{0, 0.001, 0.002, \ldots, 0.998, 0.999\}$
etc.
Y_n is equally likely $\{0, 1\times10^{-n}, 2\times10^{-n}, \ldots, 0.99\ldots99=1-10^{-n}\}$

Looks like Y_n converges to a uniform [0,1] random variable!

Let us just prove that Y_n indeed converges.

This is true, since for every $\omega\in\Omega$

$Y_n(\omega)$ is non-decreasing;
for any n, $Y_n(\omega)\le 9\times10^{-1}+9\times10^{-2}+\ldots+9\times10^{-n}=0.99\ldots9<1$

Denote the limit (which in fact always exist, not just a.s.) as $Y(\omega)$.

Can show that $Y\sim U[0,1]$

Probability Generating Functions

X r.v. taking values in $\{0,1,2,3,\ldots\}$

Definition: The p.g.f. of X is the function $G_X:[-1,1]\to R$ given by

$$G_X(s) = E\,(s^X) \equiv \Sigma_{k=0}^{\infty} P(X=k)\, s^k$$

Property: $G_X(1)=\Sigma_{k=0}^{\infty} P(X=k)=1$

Relation to MGF (Moment Generating Function):

$$M_X(t) = E\,[e^{tX}] = E\,[\,(e^t)^X\,] = G_x(s) = G_x(e^t)$$

where we set $s=e^t$ ($s>0$)

Thus $G_X(s)=M_X(\log s)$

Example [$X \sim \text{Poisson}(\lambda)$]

$$G_X(s) = \Sigma_{k=0}^{\infty} e^{-\lambda} \lambda^k/k! \times s^k = e^{-\lambda} \times \Sigma_{k=0}^{\infty} (\lambda s)^k/k!$$

$$= e^{-\lambda} \times e^{\lambda s} = e^{\lambda(s-1)}$$

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

$$G_X(s) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \times s^k = \sum_{k=0}^{n} \binom{n}{k} (ps)^k q^{n-k} = (ps+q)^n$$

PGF is a useful encoding of the distribution of X, alternative to p.m.f.

Theorem [uniqueness]: Let X and Y be two non-negative integer random variables with $G_X(s)= G_Y(s)$ for all $s\in[-1,1]$. Then the distribution of X and Y are the same, i.e. $P(X=k)=P(Y=k)$ for $k=0,1,2,\ldots$

Decoding: how do we get p.m.f. (prob. mass func.) from G(s)?

Answer: simply expand G(s) as a power series of s. The coefficient on s^k is $p_k = P(X=k)$

Example:

Suppose we are given that $G_X(s) = \dfrac{ps}{1-(1-p)s}$

where $0<p<1$. What is the distribution of X?

A: Set $q=1-p$. Then $G(s) = ps(1-qs)^{-1}=ps(1+qs+q^2s^2+q^3s^3+\ldots)$
$= 0\times s^0 + p\times s^1+pq\times s^2+ pq^2\times s^3+ pq^3\times s^4+\ldots$
so that

$P(X=0) = 0$
$P(X=1) = p$
$P(X=2) = pq^1$
$P(X=3) = pq^2$
…
$P(X=k) = pq^{k-1}$ $\qquad k>0$

So that $X\sim \text{Geom}(p)$

Lemma:[mean and variance] Let X have pgf $G(s)=G_X(s)$. Then

$$E\,X = G'(1)$$
$$\mathrm{Var}\,(X) = G''(1) + G'(1) - [G'(1)]^2$$

Proof:

(a) $G(s)=\Sigma_{k=0}^{\infty} P(X=k)\, s^k$, so that $G'(s)=\Sigma_{k=0}^{\infty} P(X=k) \times k\, s^{k-1}$

Hence $G'(1) =\Sigma_{k=1}^{\infty} P(X=k) \times k = E\,X$

(b) $G''(s)=\Sigma_{k=0}^{\infty} P(X=k) \times k\,(k-1)\, s^{k-2}$

Hence $G''(1) =\Sigma_{k=2}^{\infty} P(X=k) \times k(k-1) = E\,[X(X-1)] = EX^2 - EX$

and $\mathrm{Var}(X) = EX^2 - (EX)^2 = G''(1) + EX - (EX)^2$
$= G''(1) + G'(1) - [G'(1)]^2$

QED

Example [mean and variance of Poisson]
$X \sim \mathrm{Pois}(\lambda)$, then $G(s)=e^{\lambda(s-1)}$

$\Rightarrow G'(s) = \lambda e^{\lambda(s-1)}$, $\quad G''(s) = \lambda^2 e^{\lambda(s-1)}$
$\Rightarrow G'(1) = \lambda$, $\quad G''(1) = \lambda^2$

Hence $EX=\lambda$ and $\mathrm{Var}(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$

Theorem [sum of independent r.v.]: Suppose X and Y are independent random variables with pgf's $G_X(s)$ and $G_Y(s)$ respectively. Then if $Z=X+Y$, ($|s|\le 1$)
$G_Z(s) = G_X(s) \times G_Y(s)$

Proof: let $s\in[-1,1]$. Since X and Y are independent,
so are s^X and s^Y.

by independence

Thus $G_Z(s) = E\,s^Z = E\,s^{X+Y} = E\,(s^X \times s^Y)$
$= (E\,s^X) \times (E\,s^Y) = G_X(s) \times G_Y(s)$

QED

Corollary: Let $X_1, X_2, \ldots X_n$ be independent r.v. and set
$Z = X_1 + X_2 + \ldots + X_n$
Then

$$G_Z(s) = \prod_{k=1}^{n} G_{X_k}(s)$$

Application [sum of Poisson r.v.]

Recall that if $X \sim \text{Poisson}(\lambda)$ then $G_X(s) = e^{\lambda(s-1)}$
Let X_k be independent with $X_k \sim \text{Poisson}(\lambda_k)$
and let $Z = X_1 + X_2 + \ldots + X_n$. Then

$$G_Z(s) = \prod_{k=1}^{n} G_{X_k}(s) = \prod_{k=1}^{n} e^{\lambda_k (s-1)} = e^{\left(\sum_{k=1}^{n} \lambda_k\right)(s-1)}$$

$\Rightarrow Z \sim \text{Pois}(\Sigma\, \lambda_k)$ by uniqueness of pgf

Compound distribution

Let $X_1, X_2, \ldots$ be an infinite sequence of non-negative integer-valued i.i.d. r.v. and let N be a non-negative random variable independent of them all.

Set $T = X_1 + X_2 + \ldots + X_N$ with $T=0$ whenever $N=0$.

(e.g. N is number of insurance claims per year, X_k is the value of the claim, T=total sum payable)

<u>Theorem</u>: denote pgf of X_k's as $G_X(s)$. Then

$$G_T(s) = G_N(\, G_X(s) \,), \qquad |s| \le 1$$

<u>Proof</u>: $G_T(s) = E\, s^T = \Sigma_{n=0}^{\infty} E(s^T \mid N=n)\, P(N=n)$

Yet

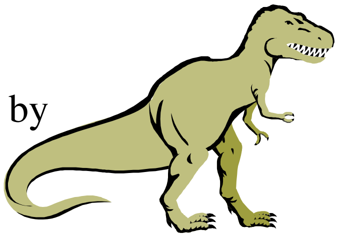
$$E\left(s^T \mid N=n\right) = E\left(s^{X_1+X_2+\ldots+X_N} \mid N=n\right) = E\left(s^{X_1+X_2+\ldots+X_n}\right)$$
$$= G_{X_1+X_2+\ldots+X_n}(s) = \left[G_X(s)\right]^n$$

since N is independent of X_k's and X_k's are independent.

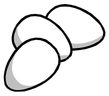
Thus $G_T(s) = \Sigma_{n=0}^{\infty} [G_X(s)]^n\, P(N=n) = E\, [G_X(s)]^N = G_N(\, G_X(s)\,)$

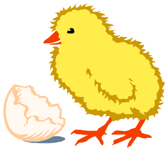
In conditional expectation language, $E(s^T|N) = [G_X(s)]^N$ and $E\, s^T = E\, [G_X(s)]^N = G_N(\, G_X(s)\,)$

QED

Example: Suppose the number N of eggs laid by 

has a Poisson distribution with mean λ.

Suppose each  hatches with probability p.

independently of other eggs and their quantity. Find the distribution of the total number of eggs  that hatch.

Solution:

Let N be the total number of eggs laid. Let $X_k=1$ if the egg number k hatches and $=0$ otherwise.

Then the total number of eggs hatching is

$$T = \begin{cases} 0 & N=0 \\ X_1+X_2+\ldots+X_N & N\geq 1 \end{cases}$$

Now

$t:=G_X(s)=(1-p)+ps$ /for Bernoulli/

$G_N(t)=e^{\lambda(t-1)}$ /for Poisson/

So $G_T(y)= G_N(\ G_X(s)\ )= G_N(\ (1\text{-}p)+ps\ )= e^{\lambda[(1-p)+ps-1]} = e^{\lambda p(s-1)}$

$\Rightarrow$ $T \sim \text{Poisson}(\lambda p)$ by uniqueness theorem.

Chapter 0: total 5 lectures, 4 slides per lecture

1. Branching processes

Model of population growth

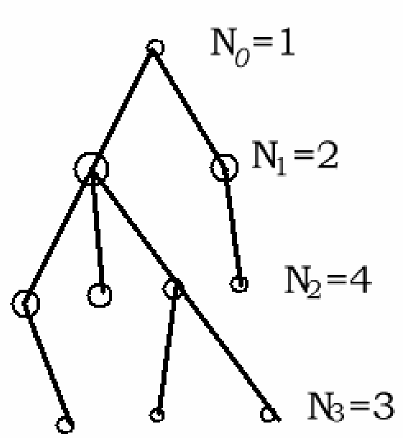
Each individual in generation j produces a random number of off-springs with common p.m.f. $\{p_k\}$ independently of the others

$j+1^{st}$ generation are the off-springs of generation j

Examples:
Male off-springs of a man (surnames)
Neutron strikes a nucleus producing a number of further neutrons out of the nucleus (chain reaction)
Growth of the number of a mutant gene in a population

Let N_j be the number of individuals in generation j, and suppose $N_0=1$.

Example of the number of off-springs Bin(3, ½)
Then $p_0= 1/8$, $p_1=3/8$, $p_2=3/8$, $p_3=1/8$



Questions:

What is the distribution of N_j, $j \geq 1$?
What about $E(N_j)$, $Var\ N_j$?

What is the behaviour of N_j as $j \to \infty$? In particular, what is the probability that the population dies out? If it does not die out, how can we describe its asymptotic growth rate?

Representation of N_j

Members of generation j are the off-springs of generation $j-1$

For $k=1,2\ldots N_{j-1}$ let Z_k be the number of children of member k in generation $(j-1)$, these r.v.'s are independent and have the same distribution as N_1

Then

$$N_j = Z_1 + Z_2 + \ldots + Z_{N_{j-1}} \tag{1.1}$$

Members of generation j are descended from members of generation 1

For $k=1,2\ldots N_1$ let W_k be the number of members of generation j who descended from member k in generation 1. These r.v.'s are independent and have the same distribution as N_{j-1}

Then

$$N_j = W_1 + W_2 + \ldots + W_{N_1} \tag{1.2}$$

Mean population size

Let $\mu=\Sigma_{k=0}^{\infty} kp_k =$ be the mean number of off-springs produced by one individual

Note $E\ N_1=\mu$ since $N_0=1$

$E\ N_j=\Sigma_{k=0}^{\infty} E(N_j \mid N_{j-1}=k) \times P(N_{j-1}=k)$

By (1.1)

$$E(N_j \mid N_{j-1}=k)=E(Z_1+Z_2+\ldots+Z_{N_{j-1}} \mid N_{j-1}=k)$$
$$=E(Z_1+Z_2+\ldots+Z_k \mid N_{j-1}=k)=E(Z_1+Z_2+\ldots+Z_k)$$
$$=kEZ_i=k\mu$$

since Z_j's are iid with mean μ. Consequently

$E\ N_j=\Sigma_{k=0}^{\infty} \mu k \times P(N_{j-1}=k)= \mu\ \Sigma_{k=0}^{\infty} k\ P(N_{j-1}=k) = \mu\ E\ N_{j-1}$

By induction, iterating this formula we get

$E\ N_j = \mu\ E\ N_{j-1} = \mu^2\ E\ N_{j-2} = \mu^3\ E\ N_{j-3} = \ldots = \mu^{j-1}\ E\ N_1=\mu^j$

<u>Exercise</u>: obtain this from (1.2), not (1.1)!

Probability generating function of N_j

Let $G(s)=\Sigma_{k=0}^{\infty}\, p_k\, s^k$ be the p.g.f. of the offspring distribution of 1 individual, $s\in[-1,1]$

Denote $\quad G_j(s)=G_{Nj}(s)=E\, s^{Nj} \quad$ the pgf of N_j

Note that $N_0=1$ and hence N_1 has the same distribution as the number of off-spring of 1 individual, so that $G_1(s)\equiv G(s)$

For $j\geq 2$ look at (1.1)

$$N_j = Z_1 + Z_2 + \ldots + Z_{N_{j-1}}$$

Then by theorem (about p.g.f. of compound distr.)

$$G_{N_j}(s) = G_{N_j-1}\left(G_Z(s)\right) = G_{j-1}\left(G(s)\right)$$

since $\quad G_{Nj}\equiv G_j\,, \quad G_{Nj-1}\equiv G_{j-1}\,, \quad G_Z\equiv G$

So

$$\boxed{G_j(s) = G_{j-1}\left(G(s)\right)}$$

which can be expanded as

$$\boxed{G_j(s) = G(G(G(\ldots(G(s)))))}$$

j times

Hence also $G_j(s)=G\left(G_{j-1}(s)\right)$ which can be obtained directly from (1.2):

$$N_j = W_1 + W_2 + \ldots + W_{N_1}$$

$\Rightarrow$

$$G_{N_j}(s) = G_{N_1}\left(G_W(s)\right) = G\left(G_{N_{j-1}}(s)\right) = G\left(G_{j-1}(s)\right)$$

<u>Example</u> [off-spring distribution Bin(3, ½)]

$$p_0 = \frac{1}{8} \quad p_1 = \frac{3}{8} \quad p_2 = \frac{3}{8} \quad p_3 = \frac{1}{8}$$

$$G(s) = \left(\frac{1}{2} + \frac{1}{2}s\right)^3 = \frac{(1+s)^3}{8} = G_1(s)$$

$$G_2(s) = G(G(s)) = \left(\frac{1}{2} + \frac{1}{2}G\;\;(s)\right)^3 = \frac{\left[1+\left(\frac{1+s}{8}\right)^3\right]^3}{8}$$

$$G_3(s) = G(G(G(s))) = \left(\frac{1}{2} + \frac{1}{2}G_2(s)\right)^3 = \frac{\left[1+\left(\frac{1+\left(\frac{1+s}{8}\right)^3}{8}\right)^3\right]^3}{8}$$

Nevertheless – useful!

Lemma: Let $\mu = E\,N_1$ and let $\sigma^2 = \mathrm{Var}\,N_1$. Then:

(a) $EN_j = \mu^j$

(b) $$Var(N_j) = \begin{cases} \dfrac{\sigma^2 \mu^{j-1}(\mu^j - 1)}{\mu - 1} & \mu \neq 1 \\ j\sigma^2 & \mu = 1 \end{cases}$$

Proof:

Using generating functions.

Know: $E\,N_j = G_j'(1)$, $E\,N_1 = G_1'(1) = G'(1) = \mu$

Since $G_j(s) = G_{j-1}(G(s))$

$\Rightarrow [G_j(s)]' = [G_{j-1}(G(s))]' = G'_{j-1}(G(s)) \times G'(s)$ *(chain rule)*

Now $E\,N_j = G_j'(1) = G'_{j-1}(G(1)) \times G'(1)$
($G(1)=1$, $EX=G_X'(1)$)

$= G'_{j-1}(1) \times \mu$
$= (E\,N_{j-1}) \times \mu$

and by iterations $E\,N_j = \mu^j$ as before

(b) $\mathrm{Var}\,N_j = G_j''(1) + G_j'(1) - [G_j'(1)]^2 = G_j''(1) + \mu^j - \mu^{2j}$

while

$$G_j''(s) = [G_j'(s)]' = [\,G'_{j-1}(G(s)) \times G'(s)\,]'$$
$$= G''_{j-1}(G(s)) \times G'(s)^2 + G'_{j-1}(G(s)) \times G''(s)$$

Now plug in $s=1$ …

QED

Extinction Probabilities

Let A_j be the event $\{N_j=0\}$ - i.e. population extinct by generation j.

Let $e_j=P(A_j)=P(N_j=0)$ be the probability of extinction by generation j

Let $A = \cup_{i=1}^{\infty} A_i$ be the event that the population is extinct *eventually*.

Let $e=P(A)$ be the probability of eventual extinction

Lemma:

a) $e_n \leq e_{n+1} \leq e_{n+2} \leq \ldots$

b) $e = \lim_{n\to\infty} e_n$

Proof: Note that $A_n \subseteq A_{n+1}$ for all n, i.e. $N_n=0 \Rightarrow N_{n+1}=0$

thus $e_j=P(A_j)\leq P(A_{j+1})=e_{j+1}$

by continuity of probability

$P(A)= P(\cup_{i=1}^{\infty} A_i) = \lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} e_n$

QED

Lemma:

For $n \geq 1$, $e_n = G(e_{n-1})$

$e = G(e)$

Furthermore, e is the smallest non-negative root of the equation $s = G(s)$

Note that $s = 1$ is always a solution of this equation!

Proof:

By definition $G_j(s) = \Sigma_{k=0}^{\infty} P(N_j = k)\, s^k$

Hence $e_j = P(N_j = 0) = G_j(0)$ for all j.

By recursion, $G_j(s) = G(G_{j-1}(s))$ so $e_j = G_j(0) = G(G_{j-1}(0)) = G(e_{j-1})$

By previous lemma,

$$e = \lim_{n\to\infty} e_n = \lim_{n\to\infty} G(e_{n-1}) = G(\lim_{n\to\infty} e_{n-1}) = G(e)$$

since $G(s)$ is continuous.

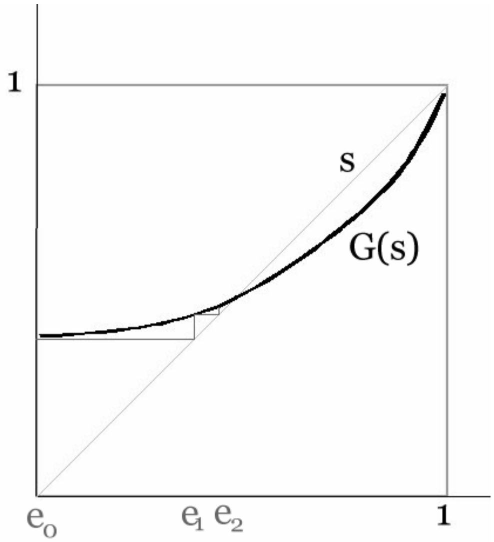
Note that $G(s)$ is an increasing function of s when $s \geq 0$
Suppose that $s \geq 0$ satisfies equation $s = G(s)$.

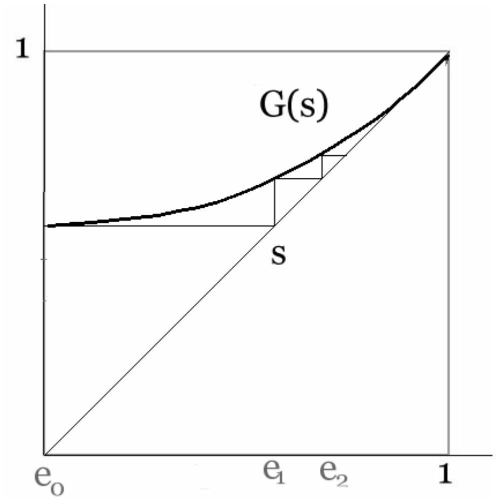
Since $e_0 = 0$ we have $e_0 \leq s$
If $e_n \leq s$ then $e_{n+1} = G(e_n) \leq G(s) = s \;\Rightarrow e_{n+1} \leq s$

Consequently by induction for all n, $e_n \leq s$,
whence $e = \lim_{n\to\infty} e_n \leq s$

Therefore $0 \leq e = G(e)$ and also $e \leq$ any positive solution of $s = G(s)$. This can happen only if e is the smallest positive solution.

QED





<u>Example</u>: [N_1~Bin(3, ½)]

$G(s)=(1+s)^3/8$

$e_0=0$
$e_1=G(e_0)= (1+0)^3/8=0.125$
$e_2=G(e_1)= (1+0.125)^3/8=0.178\ldots$
$e_3=G(e_2)= (1+0.178\ldots)^3/8=0.2043\ldots$
$e_4=0.2183$
$e_5=0.2261$
$e_6=0.2304$
$\ldots\ e = \lim_{n\to\infty} e_n$

Equation $s=G(s)=(1+s)^3/8$

$\Rightarrow \quad 8s=1+3s+3s^2+s^3 \Rightarrow 1-5s+3s^2+s^3=0$
$(s-1)(s^2+4s-1)=0$

$s=1$ or $s=(-4+[16\pm4]^{0.5})/2$

$s\in[-4.23\ldots, \quad 0.23607\ldots, \quad 1]$

The smallest non-negative solution is $e=0.23607\ldots=-2+\sqrt{5}$

Alternative (heuristic) motivation of the equation $G(e)=e$

Suppose these are k members of generation 1. Then the possibility the entire population becomes extinct is e^k . Thus

$$e = \Sigma_{k=0}^{\infty} P(\text{extinction} \mid N_1=k)\, P(N_1=k)$$
$$= \Sigma_{k=0}^{\infty} e^k\, P(N_1=k) = G(e)$$

Thus $e=G(e)$.

Sub- and super critical processes

Let $\mu=\Sigma_{k=0}^{\infty} kp_k =$ be the mean number of off-springs produced by one individual

Definition:

(a)	if $\mu=1$	the process is a	*critical*	branching process
(b)	if $\mu<1$		*sub-critical*	
(c)	if $\mu>1$		*super-critical*	

Lemma:

Suppose $p_1\neq 1$. Then:

if $\mu \leq 1$ then $e=1$

if $\mu > 1$ then $e<1$

Proof: Look at the properties of $G(s)$ for $0\leq s\leq 1$.

$G(0)=p_0$

$G(1)=1$

$G'(s) = \Sigma_{k=1}^{\infty} ks^{k-1}p_k \geq 0$

$G''(s) = \Sigma_{k=2}^{\infty} k(k-1)s^{k-2}p_k \geq 0$

Now $G'(1) = \Sigma_{k=1}^{\infty} kp_k=\mu$ and since $G''(s) \geq 0$ we have that $G'(s)$ is increasing

$G'(s) \le G'(1)$ for $0 \le s \le 1$.
$G'(s) \le \mu$ for $0 \le s \le 1$

Case A $\mu<1$

Then $G(1)=1$ and $G'(s) \le \mu < 1$. Hence

$$1-G(a)=\int_a^1 G'(s)ds < \int_a^1 1ds = 1-a$$

And so $G(s)>s$ for $0\le s<1$. *Thus* there is no solution for $s\in[0,1)$ of the equation $s=G(s)$ and therefore the smallest non-negative solution is $s=1$. So $e=1$.

Case B $\mu=1$

Since $p_1\neq 1$ we must have $p_k>0$ for some $k\ge 2$ and hence $G''(s)>0$ for $s>0 \Rightarrow G'(s)$ is strictly increasing for $s>0 \Rightarrow G'(s)<G'(1)=1$ for $0\le s<1$ and by the same argument $G(s)>s$ for $0\le s<1$. *Thus* …

Case C $\mu>1$

Then $G'(1) = \mu > 1$ and therefore by Taylor expansion
for $s = 1- \varepsilon < 1$ (i.e. $\varepsilon>0$), we have

$$G(s) = G(1-\varepsilon)=G(1) - G'(1)\varepsilon+o(\varepsilon)=1-\mu\,\varepsilon + o(\varepsilon) < 1-\varepsilon = s$$

when $\varepsilon>0$ is sufficiently small

Therefore, for s sufficiently close to 1,
$G(s)-s<0$, and at the same time $G(0)-0\ge 0$.
By continuity of $G(s)-s$, there must be an $s\in[0,1-\varepsilon)$ such that $G(s)=s$ and therefore the smallest non-negative root $e<1$.

QED

<u>Summary</u>

If $\mu<1$ then
$E\,N_j \to 0$
$Var\,N_j \to 0$
$e=1$ i.e. $N_j \to 0$ a.s.

Sub-critical

If $\mu=1$ then
$E\,N_j=1$
$Var\,N_j = j\sigma^2 \to \infty$
$e=1 \Rightarrow N_j \to 0$ a.s.

Critical

If $\mu>1$ then
$E\,N_j \to \infty$
$Var\,N_j \to \infty$
$e<1 \Rightarrow P(N_j \to 0)<1$

Super-critical

Chapter 1: total 3 lectures, 4 slides per lecture

2. The Poisson Process

Revision of the Poisson distribution

Recall: a random variable X has a Poisson(λ) distribution if it has a p.m.f. (k=0,1,2,…)

$$P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}$$

Properties: E X $=\lambda$
Var X$=\lambda$

Poisson distribution can be viewed as a limit of Binomial:

Let X~Bin (n,p) where p is small and n is large

Set λ=np; then (approximately) X~Poisson(λ)

To be more precise, fix k. Now let $n\to\infty$ and $p\to 0$ in such a way that np = constant $=\lambda$

Then

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

Yet, when $n\to\infty$

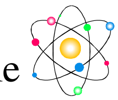
(1) $$\frac{n\ !}{(n-k)!n^k}=\frac{n(n-1)(n-2)\ldots(n-k+1)}{n^k}=1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\to 1$$

(2) $$\left(1-\frac{\lambda}{n}\right)^{-k}\to 1$$ (3) $$\left(1-\frac{\lambda}{n}\right)^n\to e^{-\lambda}$$

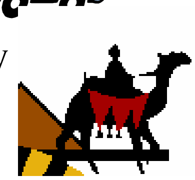
Therefore

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}\to\frac{\lambda^k e^{-\lambda}}{k!}\qquad as\quad n\to\infty$$

<u>Examples of Poisson rv's</u>

(a) Number of decays of radioactive source in 1minute time 

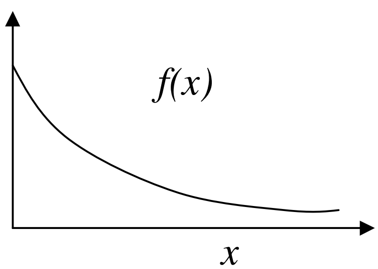
(b) Number of admissions for acute appendicitis to BRI 

(c) Number of people who fell from a camel in Sahara in 1 day 

Revision of exponential distribution

A random variable $X \sim \exp(\lambda)$ if it has a continuous distribution with density

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



Thus for x>0

$$F_X(x) = \int_{-\infty}^{x} f_X(u)\,du = \int_{0}^{x} \lambda e^{-\lambda u}\,du = 1 - e^{-\lambda x}$$

and $P(X>h) = e^{-\lambda h}$

Lack of memory property

Let $x>0$, $h>0$. Then $P(X>x+h \mid X>x) = P(X>h) = e^{-\lambda h}$

Motivation for exponential distribution as waiting time

Take $h>0$ small, and suppose $x \ge 0$

Assume a random variable X has the property

$P(X\in (x,x+h] \mid X>x)=\lambda h+o(h)$
$\Rightarrow P(X>x+h \mid X>x)=1-\lambda h+o(h)$

Look at the distribution of X:
Let $g(x)=P(X>x)$
Then $g(x+h)=P(X>x+h)$

$=P(X>x+h \mid X>x)\, P(X>x) + P(X>x+h \mid X\le x)\, P(X\le x)$

$=(1 -\lambda h+o(h))\, g(x)=(1 -\lambda h)\, g(x) + o(h)$

$g(x+h) - g(x)= -\lambda h\, g(x)+o(h)$

$$\Rightarrow \quad \frac{g(x+h)-g(x)}{h} = -\lambda g(x) + \overset{=}{o}(1)$$

whence $g'(x)= -\lambda g(x)$. Initial condition: $g(0)=P(X>0)=1$.

We have

$$\frac{g'(x)}{g(x)} = -\lambda \;\Rightarrow\; \frac{d}{dx}\log\big(g(x)\big) = -\lambda \;\Rightarrow\; \log g(x) = -\lambda x + C$$

Taking into account that $g(0)=1$, we have $C=0$ and thus

$g(x) = P(X>x) = e^{-\lambda x}$

Examples of exponential rv's:

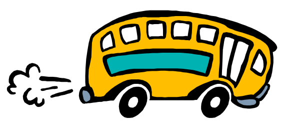
(a) Time to wait for an atom to decay
(b) Time to wait for an appendicitis case to arrive to the BRI
(c) Time to wait for the next person to fall from a camel

Minimum of independent random variables

Let $X \sim exp(\lambda)$ and $Y \sim exp(\mu)$ be independent. ($\lambda, \mu > 0$)

Example:

X time to wait for No 41 bus
Y time to wait for No 42 bus



Both go on the same route. How long do you have to wait for the first of them to come?

Time to wait $W = min(X,Y)$

Let $t \geq 0$.

$P(W>t) = P(\min(X,Y) > t) = P(X>t \text{ and } Y>t)$
(independent) $= P(X>t) \times P(Y>t) = e^{-\lambda t} \times e^{-\mu t} = e^{-(\lambda+\mu)t}$

$W \sim exp(\lambda+\mu)$

<u>Alternative (heuristic) approach</u>:

Suppose no bus arrived by time t.

$P(W\in [t,t+h] \mid W>t) = 1- P(W>t+h \mid W>t)$
$= 1- P(X>t+h \mid W>t) \times P(Y>t+h \mid W>t)$
$= 1- (1-\lambda h+o(h)) \times (1-\mu h+o(h))=1-[1-(\lambda+\mu)h+o(h)]$
$= (\lambda+\mu)\, h + o(h)$

so $W \sim \exp(\lambda+\mu)$

Also suppose no bus by time t –take it as given.

Let A_{41} = { No 41 arrives diring $(t,t+h)$ }
A_{42} = { No 42 arrives diring $(t,t+h)$ }
A_F = { any of them arrives diring $(t,t+h)$ } = $A_{41}\cup A_{42}$

Then

$P(A_{41})=\lambda h+o(h)$
$P(A_{42})=\mu h+o(h)$
$P(A_F)= P(A_{41}) + P(A_{42}) - P(A_{41} \cap A_{42})$
$= P(A_{41}) + P(A_{42}) - P(A_{41})\, P(A_{42}) = (\lambda+\mu)h+o(h)$

Hence

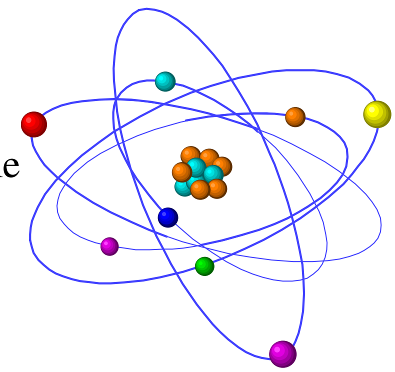
$$P(A_{41} \mid A_F) = \frac{P(A_F \mid A_{41}) P(A_{41})}{P(A_F)} = \frac{1\times(\lambda h + \overline{\overline{o}}(h))}{(\lambda+\mu)h + \overline{\overline{o}}(h)}$$

$$= \frac{\lambda}{\lambda+\mu} + \overline{\overline{o}}(h) \to \frac{\lambda}{\lambda+\mu}$$

So if we let $h\to 0$, the probability that the first bus to arrive is No 41 converges to $\lambda/(\lambda+\mu)$ (*and to $\mu/(\mu+\lambda)$ for bus No 42*)

In fact, this is independent of the waiting time W!

Further example [Lump of radioactive material]



n atoms

Each atom decays after $exp(\lambda_1)$ waiting time independently of the others.
Let T be the time to 1st decay. Then

$T \sim exp(\lambda)$ where $\lambda = n\lambda_1$

Suppose there was no decay by time t.

Look what happens during time $(t,t+h)$, where $h>0$ is small:

$P(\text{no decay}) = (1-\lambda_1 h + o(h))^n = 1 - n\lambda_1 h + o(h) = 1 - \lambda h + o(h)$

$P(\text{exactly one decay}) = n \times (\lambda_1 h + o(h)) \times (1 - \lambda_1 h + o(h))^{n-1}$
$= n\lambda_1 h + o(h) = \lambda h + o(h)$

$P(\text{more than one decay}) = 1 - P(\text{exactly one decay}) - P(\text{no decay})$
$= 1 - [\lambda h + o(h)] - [1 - \lambda h + o(h)] = o(h).$

After 1st decay, we have $(n-1)$ atoms left.
Thus the next decay occurs in $exp(\lambda')$ – distributed time where

$$\lambda' = (n-1)\lambda_1 = \frac{n-1}{n}\lambda \approx \lambda, \quad \text{provided } n \text{ is large.}$$

So we can ignore the fact that $\lambda' \neq \lambda$, and suppose that the next decay again occurs in a random time distributed $\sim exp(\lambda)$, etc.

Definition of Poisson Process

Continuous-time stochastic process $[N(t), t \geq 0]$

Recall: $\Delta \subseteq R$, $\{X_t, t \in \Delta\}$, *here we have* $\Delta = R_+ = [0, +\infty)$

$N(t)$ counts the number of events (e.g. arrivals) to have occurred by time t. Rate of occurring of the events is $\lambda = \text{const} > 0$ and is independent of anything

<u>Example</u>: radioactive decay of a large chunk of U^{238} by time t
<u>Example</u>: cars arriving at petrol station

<u>2.1 Introduction</u>

<u>Definition</u>: Let $[N(t), t \geq 0]$ be a stochastic process such that

(i) $N(t)$ is a non-negative integer $\forall\ t \geq 0$
(ii) $N(0) = 0$
(iii) The process has stationary and independent increments
(iv) $P(N(t+h)-N(t) = 0) = 1-\lambda h+o(h)$ as $h \downarrow 0$
(v) $P(N(t+h)-N(t) = 1) = \lambda h+o(h)$ as $h \downarrow 0$

$$\Rightarrow \quad P(N(t+h)-N(t) \geq 2) = o(h) \quad \text{as } h \downarrow 0$$

Then $[N(t); t \geq 0]$ is a Poisson process of rate λ

Explanation of (iii): for any $n \geq 2$ and any collection of times

$$0 \leq t_0 < t_1 < t_2 < \ldots < t_n$$

random variables

$$N(t_1)-N(t_0), \quad N(t_2)-N(t_1), \quad \ldots, \quad N(t_n)-N(t_{n-1})$$

are independent;

$N(t+s)-N(s)$ has the same distribution for all $s > 0$ (which depends on t only).

Theorem: for $t\geq 0$, $N(t)\sim\text{Poisson}(\lambda t)$, i.e.

$$P(N(t)=k)=e^{-\lambda t}\frac{(\lambda t)^k}{k!}$$

Note: hence $E\ N(t) = \lambda t = \text{Var}\ N(t)$, so the expected number of events to arrive by time t is λt.

Also, since $N(0)=0$, and the increments are stationary,

$$N(t+h) - N(t) \sim N(h) - N(0) \sim N(h) \sim \text{Poisson}(\lambda h)$$

i.e. number of events to arrive on any time interval $[t,t+h]$ is Poisson with average $= \lambda h$.

Proof: Let $p_n(t)=P(N(t)=n)$, $n=0,1,2,\ldots$

Now $p_0(t+h)= P(N(t+h)=0) = P(N(t)=0$ and $N(t+h)-N(t)=0)$

(independence of the increments)
$= P(N(t)=0) \times P(N(t+h)-N(t)=0)$
(axiom iv)
$= p_0(t) \times (1-\lambda h+o(h))$ as $h\to 0$.

Hence $$\frac{p_0(t+h)-p_0(t)}{h}=-\lambda p_0(t)+\bar{\bar{o}}(1)$$

$$\Rightarrow \quad p_0'(t)=-\lambda p_0(t) \qquad (\clubsuit)$$

Now let $n\geq 1$.

$p_n(t+h)= P(N(t+h)=n) = P(N(t)=n$ and $N(t+h)-N(t)=0)$
$+ P(N(t)=n-1$ and $N(t+h)-N(t)=1)$
$+ P(N(t)=n-2$ and $N(t+h)-N(t)=2)$
$+\ldots$
Because of (iv) and (v), all terms except the first two, are $o(h)$

and hence by independence of the increments

$p_n(t+h) = \quad P(N(t)=n) \quad \times \quad P(N(t+h)-N(t)=0)$
$+ \quad P(N(t)=n-1) \times \quad P(N(t+h)-N(t)=1) + o(h)$

$= p_n(t)\ (1-\lambda h+o(h)) + p_{n-1}(t)\ (\lambda h+o(h)) + o(h)$

$= p_n(t)\ (1-\lambda h)+ \lambda h\ p_{n-1}(t) +o(h)$

Consequently,

$$\frac{p_n(t+h)-p_n(t)}{h} = -\lambda p_n(t) + \lambda p_{n-1}(t) + \overset{=}{o}(1)$$

and

$$p_n'(t) + \lambda p_n(t) = \lambda p_{n-1}(t) \qquad (\clubsuit\clubsuit)$$

We want to solve (♣) and (♣♣) subject to

$p_0(0)=1 \qquad (\clubsuit\clubsuit\clubsuit)$
$p_n(0)=0, \quad n \geq 1.$

Observe that (♣) gives $p_0(t)=Ae^{-\lambda t}$,
and taking into account (♣♣♣) we get $A=1$ so $p_0(t)=e^{-\lambda t}$.

At the same time (♣♣) gives

$$\frac{d}{dt}\left(e^{\lambda t} p_n(t)\right) = \lambda \times e^{\lambda t} p_{n-1}(t)$$

If we denote

$$\boxed{q_n(t) := e^{\lambda t} p_n(t)}$$

the differential equation becomes

$$q_n'(t) = \lambda q_{n-1}(t)$$

…whence

$$q_n(t) = \lambda \int_0^t q_{n-1}(s)ds + const_n$$

We have $q_0(t) = e^{\lambda t} p_0(t) \equiv 1$, and also $q_n(0)=0 \;\Rightarrow\; const_n = 0$

Therefore

$$q_1(t) = \lambda \int_0^t q_0(s)ds = \lambda \int_0^t 1ds = \lambda t$$

$$q_2(t) = \lambda \int_0^t q_1(s)ds = \lambda \int_0^t \lambda s ds = \frac{\lambda^2 t^2}{2!}$$

…

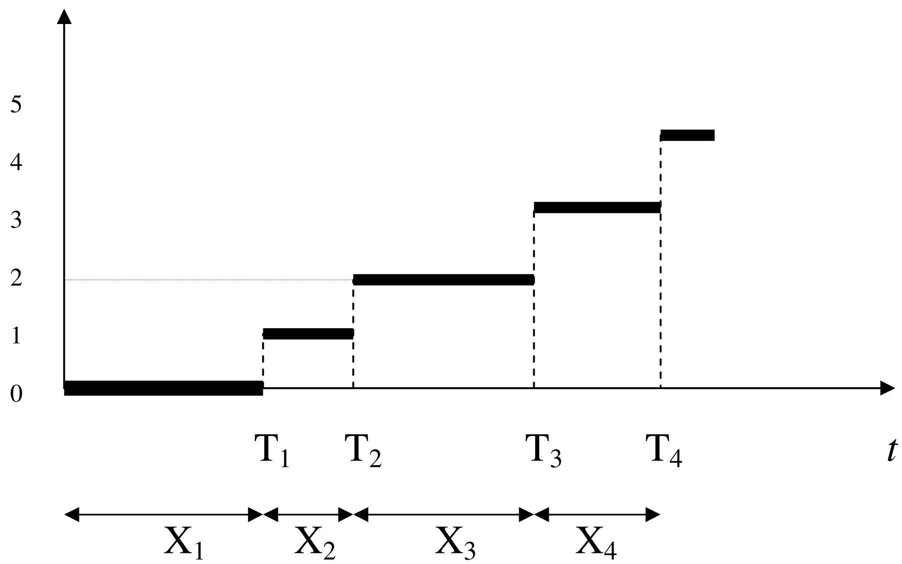
$$q_n(t) = \lambda \int_0^t q_{n-1}(s)ds = \lambda \int_0^t \frac{\lambda^{n-1} s^{n-1}}{(n-1)!} ds = \frac{\lambda^n t^n}{n!}$$

Hence

$$P(N(t) = n) = p_n(t) = e^{-\lambda t} \times q_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$$

QED

2.2. Inter arrival times



T_1 time to 1st event, $T_1=\inf\{ t \geq 0: N(t) > 0 \}$

Then $\{T_1>t\} = \{N(t)=0\}$

Hence $P(T_1>t)=P(N(t)=0)=e^{-\lambda t}$ since $N(t) \sim \text{Poisson}(\lambda t)$.

Thus $T_1 \sim \text{Exp}(\lambda)$, $E\, T_1=1/\lambda$

Let T_n be the time of the n^{th} event.
$T_n = \inf \{ t \geq 0 : N(t)=n \}$

Let $X_1=T_1$, $X_2= T_2-T_1$, $X_3= T_3-T_2$, $X_n= T_n-T_{n-1}$

be the sequence of the inter-arrival times.

Look at X_n.

$P(X_n>x \mid T_{n-1}=t) = P(N(t+x)-N(t)=0 \mid T_{n-1}=t)$
$= P(N(t+x)-N(t)=0)$

by independence of the increments.
But $N(t+x)-N(t) \sim \text{Poisson}(\lambda x)$ and thus $P(N(t+x)-N(t)=0)=e^{-\lambda x}$

Hence $P(X_n>x \mid T_{n-1}=t) = e^{-\lambda x}$ and this is independent of t and

Therefore X_n is independent of T_{n-1}. Similarly, can show that

$P(X_n>x \mid T_{n-1}=t_{n-1}, T_{n-2}=t_{n-2}, \ldots, T_1=t_1) = e^{-\lambda x}$

and X_n is independent of $T_1, T_2, \ldots, T_{n-1}$.
$\Rightarrow$ X_n is independent of $X_1, X_2, \ldots, X_{n-1}$

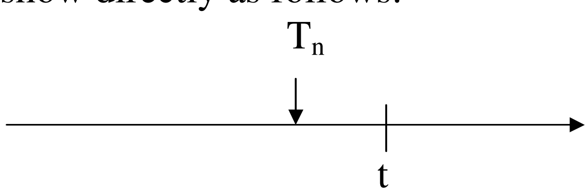
$\Rightarrow$ $X_1, X_2, \ldots,$ are iid $\text{Exp}(\lambda)$ random variables.

$T_n=X_1+X_2+\ldots X_n \sim \text{Gamma}(n,\lambda)$

Recall: its density is

$$f_{T_n}(t) = p_{n,\lambda}(t) = \begin{cases} \dfrac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!}, & t>0 \\ 0 & t \le 0 \end{cases}$$

Can show directly as follows:



$\{ T_n \le t \} = \{ N(t) \ge n \}$

$\Rightarrow P(T_n \le t) = P(N(t) \ge n) = \Sigma^{\infty}_{j=n} e^{-\lambda t}(\lambda t)^j / j!$

Hence the density

$$f_{T_n}(t) = \frac{d}{dt} P(T_n \le t) = \sum_{j=n}^{\infty} \left[e^{-\lambda t} \frac{\lambda^j j t^{j-1}}{j!} - \lambda e^{-\lambda t} \frac{\lambda^j t^j}{j!} \right]$$

$$= \lambda e^{-\lambda t} \sum_{j=n}^{\infty} \left[\frac{(\lambda t)^{j-1}}{(j-1)!} - \frac{\lambda^j t^j}{j!} \right] = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} = \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!}, \quad t > 0$$

Summary

$N(t+h) - N(t)$ = number of events in interval $[t, t+h]$
$\sim$ Poisson (λh)

Inter-arrival times $X_1, X_2, \ldots$, are
i.i.d. *Exp(λ)* random variables

Time of n^{th} arrival $T_n \sim$ Gamma (n, λ)

Example of Poisson process – decay example or death example

Conditional distribution

Suppose N(t)=1
Thus $T_1 \le t$, $T_2 > t$.

What can we say about the distribution of T on [0,t]?

Let $0 \le s \le t$

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

$$= \frac{P(N(s)=1, N(t)-N(s)=0)}{P(N(t)=1)} = \frac{P(N(s)=1)P(N(t)-N(s)=0)}{P(N(t)=1)}$$

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

So the conditional on N(t)=1 distribution $T_1 \sim U[0,t]$.

Now suppose $N(t) = n \ge 1$. Let $0 \le s \le t$.

Then $0 \le N(s) \le n$. What is the <u>*conditional*</u> distribution of N(s)?

Fix k, $\quad 0 \le k \le n$

$$P(N(s)=k \mid N(t)=n) = \frac{P(N(s)=k, N(t)=n)}{P(N(t)=n)}$$

$$= \frac{P(N(s)=k, N(t)-N(s)=n-k)}{P(N(t)=n)} = \frac{P(N(s)=k) \times P(N(t)-N(s)=n-k)}{P(N(t)=n)}$$

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

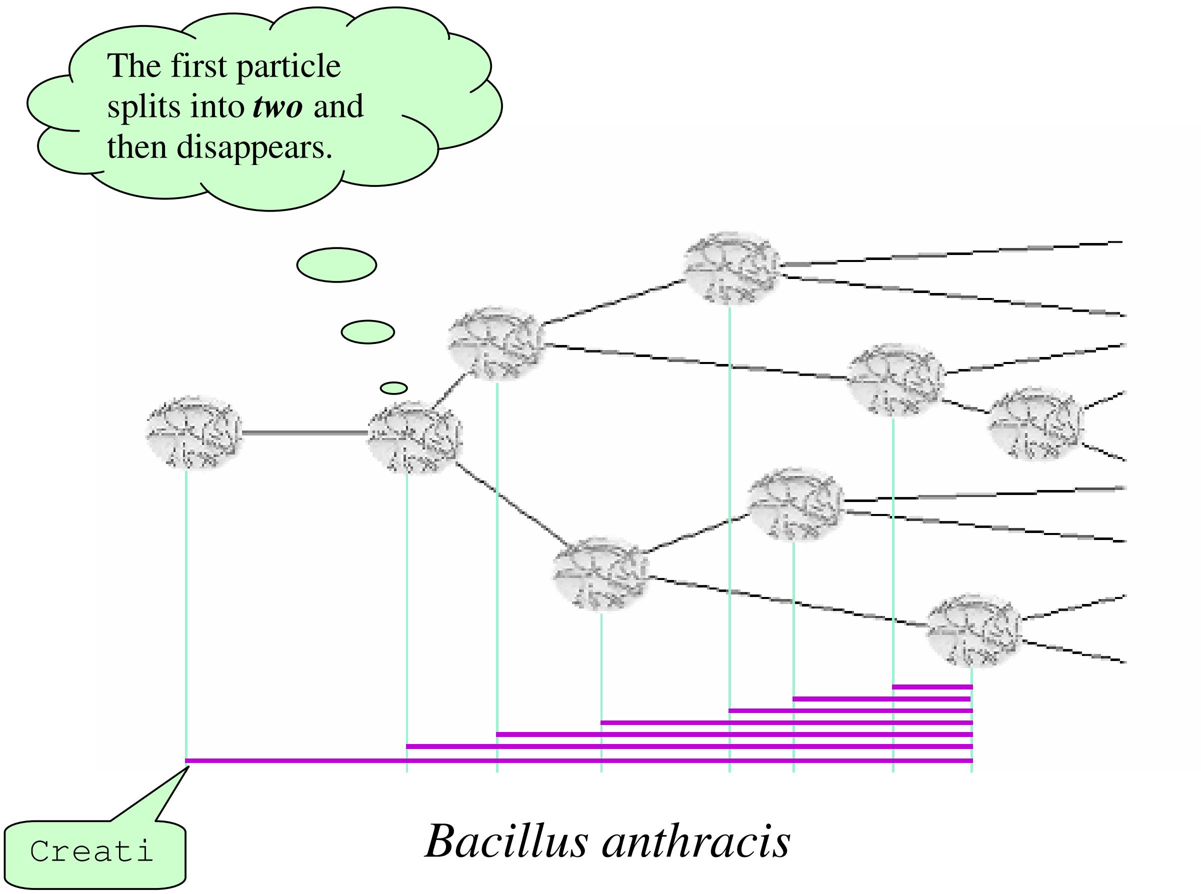
$\Rightarrow$ given N(t)=n, $\quad N(s) \sim \text{Bin}(n, s/t)$

3. Birth and Death Processes

Linear birth process (Yule process)

- Population of individuals
- An individual present at t splits into 2 during $(t,t+h)$ with probability $\lambda h+o(h)$

All individuals behave independently



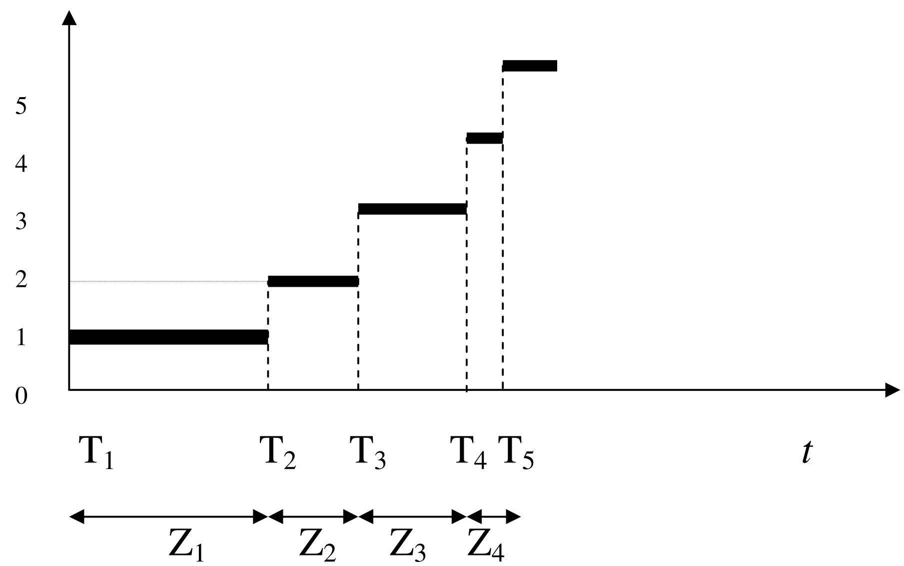
Bacillus anthracis

$N(t)$ population size at time t

$[N(t);\ t\geq 0]$ – linear birth process (Yule's process)

Inter-birth times

Population size



Suppose $N(0)=1$

For $n\geq 1$ let $T_n = \min\{t: N(t) \geq n\}$ – time till the population size reaches n

For $j\geq 1$ set $Z_j=T_{j+1}-T_j$ – time to grow from size j to $j+1$

What is the distribution of Z_j ?

Suppose population size is j. Let $X_1, X_2, \ldots, X_j$ be further times till the i^{th} member of population splits. Then $X_1, X_2, \ldots,$ are i.i.d. $Exp(\lambda)$ random variables

$Z_j = \min\{ X_1, X_2, \ldots, X_j \} \Rightarrow Z_j \sim Exp(j\lambda)$

Thus $EZ_j=1/(j\lambda)$

Now $T_n=Z_1+Z_2+\ldots+Z_{n-1}$

$\Rightarrow$

$$ET_n = E(Z_1) + E(Z_2) + \ldots + E(Z_{n-1}) = \frac{1}{\lambda}\left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n-1}\right) \cong \frac{\log n}{\lambda}$$

for large n.

<u>Distribution of N(t)</u>:

As before, assume $N(0)=1$

Let $p_n(t)=P(N(t)=n)$

Then $p_n(t+h) = P(N(t)=n$ and no births during $[t,t+h])$
$+ P(N(t) = n-1$ and 1 birth during $[t,t+h])$
$+ o(h)$

Now

$P(N(t)=n$ & no births in $[t,t+h])$
$= P(N(t)=n) \times P($no births in $[t,t+h] \mid N(t)=n)$
$= p_n(t)\times(1-\lambda h+o(h))^n = p_n(t)\times(1-n\lambda h) \quad +o(h)$

$P(N(t)=n-1$ & one birth in $[t,t+h])$
$= P(N(t) = n-1) \times P(1$ birth in $[t,t+h] \mid N(t)=n-1)$
$= p_{n-1}(t) \times (n-1)\, (\lambda h+o(h))^1\, (1-\lambda h+o(h))^{n-2}$
$= p_{n-1}(t) \times (n-1)\lambda h \quad + o(h)$

Thus
$p_n(t+h) = (1-n\lambda h)\, p_n(t) + \lambda h(n-1)\, p_{n-1}(t) +o(h)$

Consequently,

$$\frac{p_n(t+h) - p_n(t)}{h} = -\lambda n p_n(t) + \lambda(n-1)p_{n-1}(t) + \overline{\overline{o}}(1)$$

and hence

$$p_n'(t) = -\lambda n p_n(t) + \lambda(n-1) p_{n-1}(t)$$

Also for n=1 $\quad p_1'(t) = -\lambda p_1(t)$

<u>Initially</u>: $p_1(0)=1$ and $p_n(0)=0$, $n \geq 2$.

Rather than solving this, we can verify that

$$p_n(t) = P(N(t) = n) = e^{-\lambda t}\left(1 - e^{-\lambda t}\right)^{n-1}, \quad n=1,2,\ldots$$

is the solution of this system, hence $N(t) \sim \text{Geom}\,(e^{-\lambda t})$
and hence $E\,N(t)=e^{\lambda t}$

(Not surprising – given $E\,T_n \approx (\log n)/\lambda$).

Note: since there is *only one* solution to the system of equations of this type with given initial conditions, we found *<u>the</u> <u>solution</u>*.

Linear birth and death process

Look at individual present at time t. Suppose during [t,t+h] it either:

- ♥ splits into two (birth) with probability $\lambda h + o(h)$
- ♣ dies out with probability $\mu h + o(h)$
- ♦ neither with probability $1-(\lambda+\mu)h + o(h)$

Note: probability that either birth or death occurs is $(\lambda+\mu)h+o(h)$, and given *something* has happened, the probability of birth is $\lambda/(\lambda+\mu)$ and that of death $\mu/(\lambda+\mu)$

Suppose population size is n at time t. In [t,t+h] we have:

- ♥ 1 birth with probability $n\lambda h + o(h)$
- ♣ 1 death with probability $n\mu h + o(h)$
- ♦ nothing, with probability $1- n(\lambda+\mu)h + o(h)$

Let $N(0)=n_0$ and set as usual $p_n(t) = P(N(t)=n)$

Then for $n\geq 1$

$p_n(t+h) = [1- n(\lambda+\mu)h] \times p_n(t) + \lambda(n-1)h\, p_{n-1}(t)$
$+ \mu(n+1)h \times p_{n+1}(t) + o(h)$

Consequently,

$$\frac{p_n(t+h)-p_n(t)}{h} = -(\lambda+\mu)np_n(t)+\lambda(n-1)p_{n-1}(t)+\mu(n+1)p_{n+1}(t)+\bar{\bar{o}}(1)$$

and

$$p_n'(t) = -(\lambda+\mu)np_n(t)+\lambda(n-1)p_{n-1}(t)+\mu(n+1)p_{n+1}(t)$$

Generalized birth and death process

Population size is n at time t.
During time [t,t+h] either:

♥	one birth	with probability	$\lambda_n h \quad + o(h)$
♣	one death	with probability	$\mu_n h \quad + o(h)$
♦	neither	with probability	$1-(\lambda_n+\mu_n)h + o(h)$

If $p_n(t)=P(N(t)=n)$ then we can obtain (as before) that

$$p_n(t+h) = \quad [1-(\lambda_n+\mu_n)h]p_n(t)+(\lambda_{n-1}h)p_{n-1}(t)+(\mu_{n+1}h)p_{n+1}(t)+\bar{\bar{o}}(h)$$

yielding

$$\boxed{p_n'(t) = -(\lambda_n+\mu_n)p_n(t)+\lambda_{n-1}p_{n-1}(t)+\mu_{n+1}p_{n+1}(t)}$$

Example 1: $\lambda_n \equiv \lambda = \text{constant} > 0$, $\mu_n \equiv 0$

Pure immigration ⇒ Poisson process (λ)

Example 2: $\lambda_n = \lambda n$ where $\lambda > 0$, $\mu_n \equiv 0$
⇒ Yule process

Example 3: $\lambda_n \equiv \lambda > 0$ (constant immigration)
$\mu_n \equiv n\mu$ (death rate μ per individual)

For this example (3), using general formula we get for $n \geq 1$

$$p_n'(t) = -(\lambda + n\mu)p_n(t) + \lambda p_{n-1}(t) + \mu(n+1)p_{n+1}(t)$$
$$p_0'(t) = -\lambda p_0(t) + \mu p_1(t)$$

Without proving it here, assume that as $t \to \infty$

$$\begin{cases} p_n(t) \to \hat{p}_n \\ p_n'(t) \to 0 \end{cases}$$

Then

$$\begin{cases} 0 = -(\lambda + n\mu)\hat{p}_n + \lambda \hat{p}_{n-1} + \mu(n+1)\hat{p}_{n+1} \\ 0 = -\lambda \hat{p}_0 + \mu \hat{p}_1 \end{cases}$$

and it follows for any $n \geq 1$

$$a_n := [\mu(n+1)\hat{p}_{n+1} - \lambda \hat{p}_n] = [n\mu \hat{p}_n - \lambda \hat{p}_{n-1}] =: a_{n-1}$$

and also $a_0 = 0$ $\Rightarrow$ $a_n \equiv 0$ for all n $\Rightarrow$

$$\hat{p}_n = \frac{\lambda}{\mu n}\hat{p}_{n-1} \quad \Rightarrow \quad \hat{p}_n = \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n \hat{p}_0$$

On the other hand, we know

$$\sum_{n=0}^{\infty} \hat{p}_n \equiv 1 \quad \Rightarrow \quad \hat{p}_0 \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n = 1 \quad \Rightarrow \quad \hat{p}_0 e^{\frac{\lambda}{\mu}} = 1$$

$$\Rightarrow \quad \hat{p}_n = \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n e^{-\lambda/\mu} \qquad n = 0,1,2,\ldots$$

Thus the stationary distribution is Poisson (λ/μ)
Chapter 2 and 3: total 4 lectures, 5 1/4 slides per lecture

4. Random walks (version 4 March 08)

Gambler's ruin

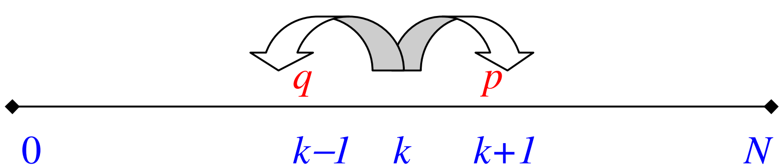
Gambler has £ k
Opponent has £ $N-k$
Total =£ N between them
Bet £1 at a time

Gambler increases capital to £$k+1$ with probability p
Gambler decreases capital to £$k-1$ with probability $q=1-p$

Successive bets are independent

Repeat procedure until one player is out of money (… we show later that the probability of this is 1) so that the gambler has either £0 or £N

Want to know: P(gambler ruins)=P(…ends up with £0).



Let Y_n be the gambler's capital in ££ after n bets. Then the process $[Y_k: k=0,1,2,\ldots]$ is a Markov process with

$$P(Y_{n+1} = m+1 \mid Y_n = m) = p \qquad \text{for } 1 \le m \le N-1$$
$$P(Y_{n+1} = m-1 \mid Y_n = m) = q$$

and additionally

$$P(Y_{n+1}=0 \mid Y_n=0) = 1 \qquad \Leftarrow \text{absorbing barriers at } 0 \text{ and } N$$
$$P(Y_{n+1}=N \mid Y_n=N) = 1$$

Gambler is ruined if $Y_n=0$ for some $n\geq 0$.

Let $p_k=P(\text{ruin} \mid Y_0=k) \leftarrow$ defined for $k=0,1,2,\ldots,N-1,N$

Then

$$p_0 = 1 \qquad \textbf{(4.1)}$$

and

$$p_N = 0 \qquad \textbf{(4.2)}$$

Suppose that he starts with £k.

Let A be the event "gambler ruined"
B be the event "gambler wins the 1^{st} bet"

Then

$$P(A) = P(A|B)\,P(B)+ P(A|B^c)\,P(B^c)$$

(partition theorem)

$$\Rightarrow \quad p_k=p_{k+1}p + p_{k-1}q \qquad k=1,2,\ldots,N-1 \qquad (4.3)$$

Want to solve (4.3) subject to the boundary condition (4.1) and (4.2)

Try to solve it, assuming $p_k=\theta^k$

$$\theta^k = p\theta^{k+1} + q\theta^{k-1}$$

$$p\theta^2 - \theta + q = 0$$

$$(p\theta - q)(\theta - 1)=0$$

$$\theta_1=q/p \qquad \text{and} \qquad \theta_2=1$$

We can verify, that if <u>$p\neq q$</u> then (4.3) has general solution

$$p_k = A\theta_1^{\,k} + B\theta_2^{\,k} = A\,(q/p)^k + B, \quad k=0,1,2,\ldots,N$$

From (4.1) we have

$1 = A + B \quad \Rightarrow \quad p_k = A[(q/p)^k - 1] + 1$

From (4.2) we have

$p_N = A[(q/p)^N - 1] + 1 = 0 \quad \Rightarrow \quad A = 1 / (1-(q/p)^N)$.

$\Rightarrow$

$$p_k = \frac{\left(\frac{q}{p}\right)^k - 1}{1 - \left(\frac{q}{p}\right)^N} + 1 = \frac{\left(\frac{q}{p}\right)^k - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N}, \qquad k = 0,1,\ldots N \quad (4.4)$$

If p=q=½:

Then $\theta_1 = \theta_2 = 1$ and general solution of (4.3) is

$p_k = (A+Bk)\theta^k = A + Bk$

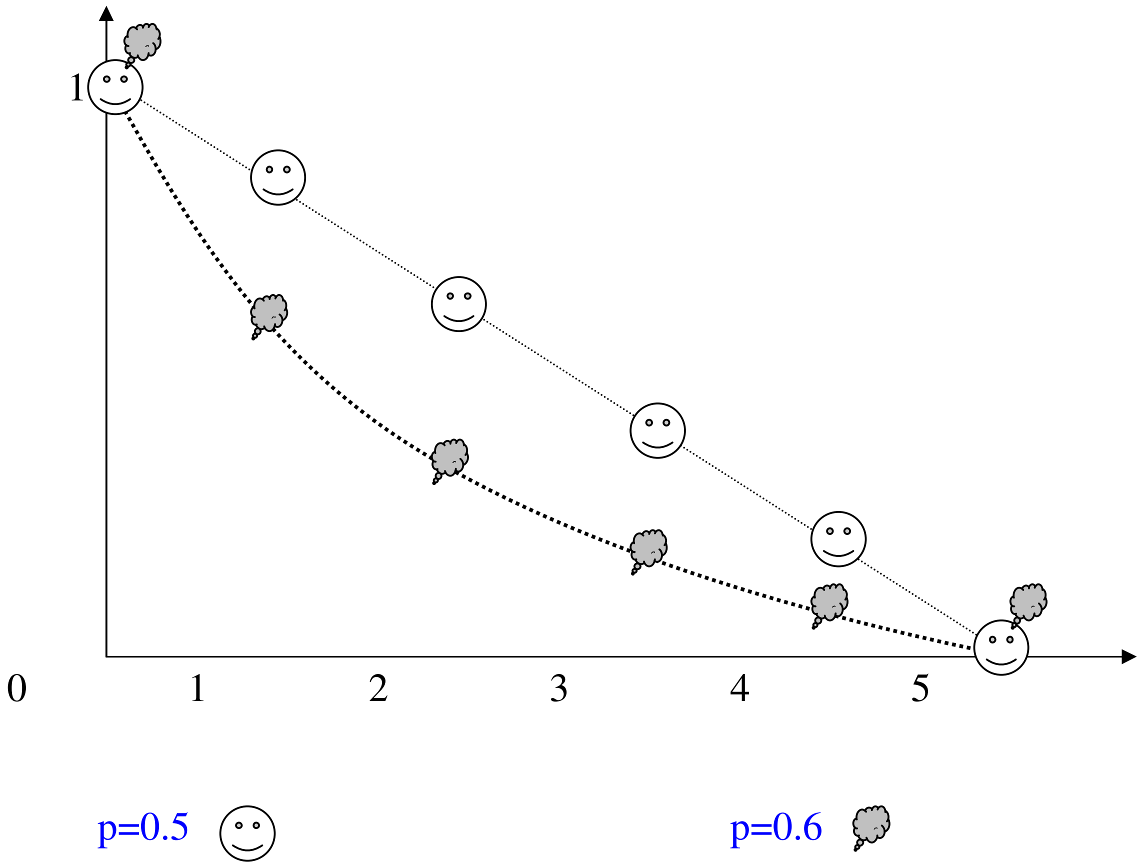
$p_o = 1 \quad \Rightarrow \quad A = 1 \quad \Rightarrow \quad p_k = 1 + Bk$

$p_N = 0 \quad \Rightarrow \quad 1 + BN = 0 \quad \Rightarrow \quad B = -1/N$

$\Rightarrow$

$$p_k = 1 - \frac{k}{N}, \qquad k = 0,1,\ldots N \quad (4.5)$$

Example [N=5]



	k =	0	1	2	3	4	5
p= 0.5	p_k =	1	0.8	0.6	0.4	0.2	0
p= 0.6	p_k =	1	0.616	0.360	0.190	0.076	0

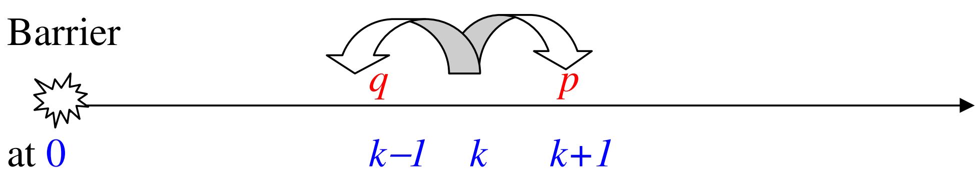
Absorbing barrier at 0 only

An insurance company starts with assets k.
Every day the assets either

$\Uparrow$ increase with probability p
$\Downarrow$ decrease with probability $q=1-p$

Successive days are independent.

What is the probability the company will go bankrupt?



Let Y_n be the company's assets after n days

$[Y_n: n=0,1,2,\ldots]$ is a Markov process, called a random walk (RW) with absorbing barrier at 0. No upper barrier!

Let A be the event of bankruptcy.

$A = \{Y_n = 0$ for some $n \geq 0\}$

<u>Lemma</u>: Let $k \geq 1$ be fixed and assume $Y_0=k$. Let p_k be the probability of bankruptcy for the company.
Let $p_k^{(N)}$ be the probability of ruin starting at £k for the gambler ruin problem with upper barrier at N.
Then $\lim_{N\to\infty} p_k^{(N)} = p_k$

Proof. Fix $k\geq 1$. Assume $Y_0=k$. Let

$A_N=\{\omega: \exists\, n\geq 1$ (depending on ω) such that $Y_k(\omega)\leq N-1$ for $k=0,1,2,\ldots,n-1$ and $Y_n(\omega)=0\}$

i.e. the event that the process absorbed at 0, never really reaches level N. So

$P(A_N)=p_k^{(N)}$

Note: $A_N\subseteq A_{N+1}$

/indeed, if $\omega\in A_N$

$\Rightarrow\quad \exists\, n$ such that $Y_n(\omega)=0$ and $Y_k(\omega)\leq N-1$ for $0\leq k\leq n-1$

$\Rightarrow\quad \exists\, n$ such that $Y_n(\omega)=0$ and $Y_k(\omega)\leq N$ for $0\leq k\leq n-1$

$\Rightarrow\quad \omega\in A_{N+1}$

Let $$A=\bigcup_{N=1}^{\infty}A_N$$

Then $\omega\in A \quad\Rightarrow\quad \omega\in A_N$ for some N

$\Rightarrow\quad Y_n(\omega)=0$ for some n

Conversely, if $Y_n(\omega)=0$ for some n

$Y_n(\omega)=0$ and $Y_k(\omega)\leq N-1$ for $k=0,1,2,\ldots,n-1$

where $N:=\max(Y_0,Y_1,\ldots,Y_n)+1$

$\Rightarrow\quad \omega\in A_N \quad\Rightarrow\quad \omega\in A$

Therefore $A\equiv\{\, Y_n=0$ for some $n\,\}$

By continuity of probability we have

$p_k=P(A)=\lim_{N\to\infty}P(A_N)=\lim_{N\to\infty}p_k^{(N)}$ *Q.E.D.*

Lemma: For the random walk with absorbing barrier at 0 but no upper barrier

$$\text{P(absorbed at } 0 \mid \text{start at } k\text{)} = \begin{cases} \left(\dfrac{q}{p}\right)^k & \text{if} \quad q < p \\[2ex] 1 & \text{if} \quad q \geq p \end{cases} \tag{4.6}$$

Proof:

For $q \neq p$ we have by equation (4.4)

$$p_k^{(N)} = \frac{\left(\dfrac{q}{p}\right)^k - \left(\dfrac{q}{p}\right)^N}{1 - \left(\dfrac{q}{p}\right)^N} \xrightarrow{N \to \infty} \begin{cases} \left(\dfrac{q}{p}\right)^k & \text{if} \quad q < p \\[2ex] 1 & \text{if} \quad q > p \end{cases}$$

For p=q=½ by equation (4.5)

$$p_k^{(N)} = 1 - \frac{k}{N} \xrightarrow{N \to \infty} 1$$

Thus P(absorbed at 0)=1 for this case p=q=1/2.

Q.E.D.

Alternative interpretation of this formula (heuristic)

Let $\theta = p_1 = P(\text{absorbed at } 0 \mid Y_0 = 1)$
Then

$p_k = P(\text{absorbed at } 0 \mid Y_0 = k) = P(\text{reach 1 from k})P(\text{reach 0 from 1})$
$= p_{k-1}\theta$

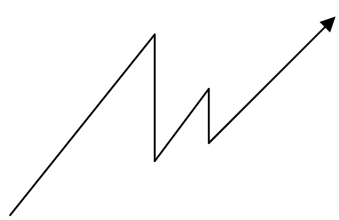
So $p_k = \theta^k$ for *some* θ.

Now
$p_k = p_{k+1}p + p_{k-1}q$

$\Rightarrow \quad \theta^k = p\theta^{k+1} + q\theta^{k-1}$
$p\theta^2 - \theta + q = 0$
$\theta = q/p \qquad \text{or} \qquad \theta = 1$

In fact, θ is the smallest root (can generalize to other skip-free processes)

Insurance company assets:



not skip-free, nevertheless, can get the formula for the asymptotic probability of ruin

<u>Example</u>: X_n are asset at day n. Cost is £1 per day;

income is $\begin{cases} 0 & with & probability & 0.2 \\ 1 & with & probability & 0.4 \\ 2 & with & probability & 0.3 \\ 3 & with & probability & 0.1 \end{cases}$

Start at £k. Find the probability of ruin.

We have:

$p_k = 0.2\, p_{k-1} + 0.4\, p_k + 0.3\, p_{k+1} + 0.1\, p_{k+2}$

Try $p_k = \theta^k$. Then

$\theta^k = 0.2\, \theta^{k-1} + 0.4\, \theta^k + 0.3\, \theta^{k+1} + 0.1\, \theta^{k+2}$

$\theta = 0.2 \;\; + 0.4\, \theta + 0.3\, \theta^2 + 0.1\, \theta^3$

$\theta^3 + 3\theta^2 - 6\theta + 2 = 0$

$(\theta-1)(\theta^2 + 4\theta - 2) = 0$

Roots: $\theta=1$; $\quad$ $\theta = -2+\sqrt{6}$; $\quad$ $\theta = -2-\sqrt{6}$;

Smallest positive root $\theta = \sqrt{6} - 2 = 0.45\ldots$

So $p_k \approx (0.45)^k$

Unrestricted random walk

Let $X_1, X_2, \ldots$ be i.i.d. r.v. with

$P(X_i=1) \quad = p,$
$P(X_i=-1) = q = 1-p$

Let k be an integer

Set $Y_0=k$ and for $n\geq 1$

$Y_n = k + \Sigma^n_{i=1} X_i$

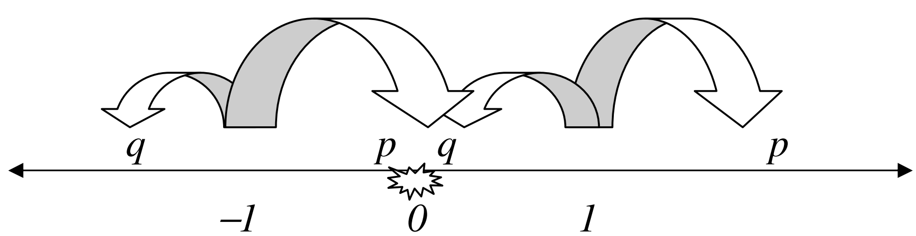
Then the process $[Y_i\text{: } i=0,1,2,\ldots]$ is a simple (unrestricted) random walk

Returns to the origin – Suppose $k=0$ so that $Y_0=0$

Interested in

- whether walk ever returns to 0
- time taken to return

Case $p>q$



By the results for random walk with absorbing barrier at 0

$P(Y_n=0 \text{ for some } n\geq 1 \mid Y_1=1) \quad =q/p$
$P(Y_n=0 \text{ for some } n\geq 1 \mid Y_1=-1) \quad =1$

Thus for a walk starting at 0

$r = P(Y_n=0 \text{ for } n\geq 1)=P(Y_n=0 \text{ for } n\geq 1 \mid Y_1=1)\, P(Y_1=1)$
$+ P(Y_n=0 \text{ for } n\geq 1 \mid Y_1=-1)\, P(Y_1=-1)$
$=(q/p)\times p+1\times q = 2q$

Let R be the total number of returns to 0.

Shown: $P(R\geq 1)=r$.

Easy to show using Markov property that

$P(R=0) = 1-r$
$P(R=1) = r(1-r) \qquad \Rightarrow \qquad P(R=m)=r^m(1-r), \qquad m\geq 0$
$P(R=2) = r^2(1-r)$ etc. $\qquad$ *(shifted geometric)*

Also by the strong law of large numbers

$$\frac{Y_n}{n}=\frac{1}{n}\left(\sum_{i=1}^{n} X_i\right) \xrightarrow{a.s.} E(X_i)=p-q>0 \quad as \quad n\to\infty$$

In summary $Y_n \to \infty$ as $n \to \infty$ and before doing so it makes a geometric number of returns to the origin.

Similarly, if $q>p$ the walk returns to zero at least once with probability $r=2p$.
It makes a geometric number of returns, $Y_n\to -\infty$ as $n\to\infty$.

Case $p=q=\frac{1}{2}$

$P(Y_n=0$ for $n\geq 1 \mid Y_1= \ \ 1)=1$
$P(Y_n=0$ for $n\geq 1 \mid Y_1= -1)=1$

Thus $r=P(Y_n=0 \text{ for } n\geq 1)=1\times p+1\times q=1$

Walk is certain to return at least once to 0.
Easy to show using Markov property that
$P(\text{returns at least } m \text{ times}) = r^m = 1$

Let $B_m=$ {returns at least m times} and
$B=\cap^{\infty}_{m=1}B_m =$ {returns ∞ many times}

Since $B_{m+1}\subseteq B_m$ by continuity of probability

$P(\text{returns } \infty \text{ many times})= P(B)=\lim_{m\to\infty} P(B_m)=1$

In general: (when $p=q= \frac{1}{2}$)

Let j be any integer

Then we know by results for RW with barrier at 0 that

$P(Y_n=j$ for some $n\geq 1 \mid Y_0=0)=1$

So RW hits level j with probability 1 after hitting the origin 0; hence RW visits j infinitely many times (*e.g. use Borel-Cantelli lemma*)

Digression (gambling)

Suppose $p=q=\frac{1}{2}$. Then X_i can be interpreted as the amount one on the i^{th} bet gambling on a fair bet

Let $Y_0=0$ and $Y_n=\Sigma^n_{i=1} X_i$, $n\geq 1$, be the profits after n bets.

Let M be some sum of money, e .g. M=£1,000,000

Then $P(Y_n=M \text{ for some } n\geq 1 \mid Y_0=0)=1$

Thus if we eploy the following strategy:

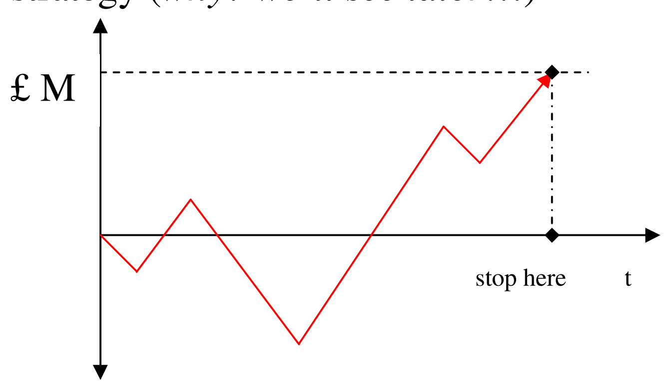
```
Continue until the net profit is M; then
stop
```

we are sure to earn M pounds – this strategy always makes profit

However, you would need:

∞ amount of money
∞ time

to carry out this strategy (*why? we'll see later…*)



Return times

From now on assume $[S_n: n\geq 0]$ is an unrestricted simple random walk with $S_0=0$.

Definitions:
Set $p_{oo}(0)=1$, and for $n \geq 1$, $p_{oo}(n)=P(S_n=0 \mid S_o=0)$.
Thus $p_{oo}(n)$ is the probability to be at the origin after n steps

For $n \geq 1$ set
$f_{oo}(n)=P(S_n=0 \text{ but } S_1\neq 0, S_2\neq 0,\ldots, S_{n-1}\neq 0, \mid S_o=0)$
i.e. the probability to come back in **exactly** n steps.

Note: let B_m be the event {1st return to origin is after m steps}, so that $P(B_m)=f_{oo}(m)$. Then

{Return to origin at some time}$= \cup^{\infty}_{m=1}B_m$
and since these events are disjoint,

$P(\text{return to origin})=\Sigma^{\infty}_{m=1}P(B_m) =\Sigma^{\infty}_{m=1} f_{oo}(m)$

Thus
if $\Sigma^{\infty}_{m=1} f_{oo}(m)=1$ return is certain $\Rightarrow$ RW is *persistent* (or *recurrent*)

if $\Sigma^{\infty}_{m=1} f_{oo}(m)<1$ may not return $\Rightarrow$ RW is *transient*

Since $P(S_n=0)= \Sigma^{\infty}_{m=1}P(S_n=0 \mid B_m)P(B_m)$,
and $P(S_n =0 \mid B_m) \;= p_{oo}(n-m),\quad$ if $1\leq m\leq n$
$\qquad\qquad\qquad\;\;= 0,\quad$ if $m>n$

$$\boxed{p_{oo}(n) = P(S_n=0) = \Sigma^{n}_{m=1} p_{oo}(n-m)f_{oo}(m)} \quad \text{for } n\geq 1$$

Definition. Set

$$P(s) = \sum_{n=0}^{\infty} p_{00}(n) s^n \quad \text{and} \quad F(s) = \sum_{n=1}^{\infty} f_{00}(n) s^n \quad \text{for } |s|<1$$

Theorem. For $|s|<1$

(i) $P(s) = 1+P(s)F(s)$

(ii) $P(s) = (1-4pqs^2)^{-1/2}$

(iii) $F(s) = 1-(1-4pqs^2)^{1/2}$

(the latter two are specific to $q-p$ walk)

Proof.

(i)

$$\sum_{n=0}^{\infty} p_{00}(n)s^n = p_{00}(0) + \sum_{n=1}^{\infty} p_{00}(n)s^n = 1 + \sum_{n=1}^{\infty} p_{00}(n)s^n$$

$$= 1 + \sum_{n=1}^{\infty}\left(\sum_{k=1}^{n} p_{00}(n-k) f_{00}(k)\right)s^n = 1 + \sum_{k=1}^{\infty}\left(\sum_{n=k}^{\infty} p_{00}(n-k) f_{00}(k) s^n\right)$$

$$= 1 + \sum_{k=1}^{\infty} f_{00}(k)s^k\left(\sum_{n=k}^{\infty} p_{00}(n-k)s^{n-k}\right) = 1 + \sum_{k=1}^{\infty} f_{00}(k)s^k\left(\sum_{m=0}^{\infty} p_{00}(m)s^m\right)$$

$$= 1 + \sum_{k=1}^{\infty} f_{00}(k)s^k P(s) = 1 + P(s)\sum_{k=0}^{\infty} f_{00}(k)s^k = 1 + P(s)F(s)$$

(ii) Note that $S_n \neq 0$ if n is odd. Now let $n=2m$.
If $S_n=0$, then there were exactly m of $+1$'s and m of -1's. And number of $+1$'s in $2m$ trials is $\text{Bin}(2m,p)$, hence

$$\text{P}(S_{2m}=0) = \binom{2m}{m} p^m q^m = p_{oo}(2m)$$

So

$$P(s) = 1 + p_{00}(2)s^2 + p_{00}(4)s^4 + \ldots = \sum_{m=0}^{\infty} \binom{2m}{m} p^m q^m s^{2m} = 1 \Big/ \sqrt{1 - 4pqs^2}$$

Reason: use Taylor expansion:

$$(1-x)^{-\frac{1}{2}} = 1 + \frac{1}{2}x + \frac{1}{2}\frac{3}{2}\frac{x^2}{2!} + \frac{1}{2}\frac{3}{2}\frac{5}{2}\frac{x^3}{3!} + \ldots = \sum_{m=0}^{\infty} \frac{1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2m-1)}{2^m}\frac{x^m}{m!}$$

$$= \sum_{m=0}^{\infty} \frac{1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2m-1)}{2^m}\frac{2 \cdot 4 \cdot 6 \cdot \ldots \cdot (2m)}{2 \cdot 4 \cdot 6 \cdot \ldots \cdot (2m)}\frac{x^m}{m!} = \sum_{n=0}^{\infty} \frac{(2m)!}{2^m}\frac{1}{2^m \cdot 1 \cdot 2 \cdot 3 \cdot \ldots \cdot m}\frac{x^m}{m!}$$

$$= \sum_{m=0}^{\infty} \frac{(2m)!}{m!m!}\frac{x^m}{4^m} = = \sum_{m=0}^{\infty} \binom{2m}{m} \left(x \big/ 4\right)^m$$

Now: $x=4pqs^2$

(iii) Use $F(s) = 1 - 1/P(s) = 1 - (1 - 4pqs^2)^{1/2}$ *Q.E.D*

<u>Corollary</u>.

(i) The probability that the walk returns to 0 is

$$\sum_{n=0}^{\infty} f_{00}(n) = 1 - | p - q |$$

(ii) Suppose $p=q=½$ so that return is certain. Let T_{00} be the time of return. Then

$$\mathrm{E}(T_{00}) = \sum_{n=0}^{\infty} n f_{00}(n) = \infty$$

<u>Proof</u>.

(i)

$$\sum_{n=1}^{\infty} f_{00}(n) = \lim_{s \uparrow 1} \sum_{n=1}^{\infty} f_{00}(n)s^n = \lim_{s \uparrow 1} F(s)$$

$$= \lim_{s \uparrow 1} \left[1 - \left(1 - 4pqs^2\right)^{1/2}\right] = 1 - \sqrt{1 - 4pq} = 1 - \sqrt{(p+q)^2 - 4pq}$$

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

(ii)

Return certain ⇔ $\Sigma\, f_{00}(n)=1$ ⇔ $p=q$ ⇔ $p=q=½$

Thus in this case

$$F(s)=1-(1-s^2)^{1/2}$$

Now

$$E(T_{00}) = \sum_{n=1}^{\infty} n f_{00}(n) = \lim_{s\uparrow 1} \sum_{n=1}^{\infty} n s^{n-1} f_{00}(n)$$

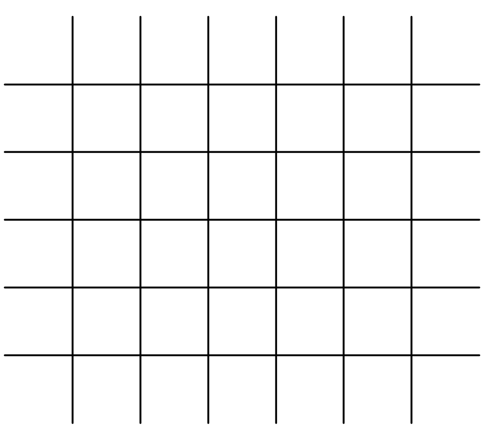
$$= \lim_{s\uparrow 1} F'(s) = \lim_{s\uparrow 1} \frac{s}{\sqrt{1-s^2}} = +\infty$$

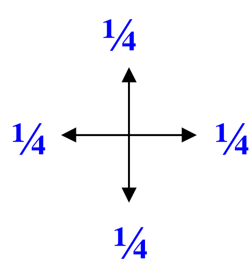
<u>Summary</u>. Simple random walk (unconstrained) starting at the origin:

<u>Case</u> $p\neq q$ walk transient (returns a "geometric" number of times P(return n times)$=r^n(1-r)$ where $r = \Sigma\, f_{00}(n) = \min(2p,2q)$.

<u>Case</u> $p=q=½$ walk recurrent (returns infinitely many times)

<u>Higher dimensions d=2, d ≥ 3 for simple random walk</u>.





2.3 Stopping times and the Wald lemma

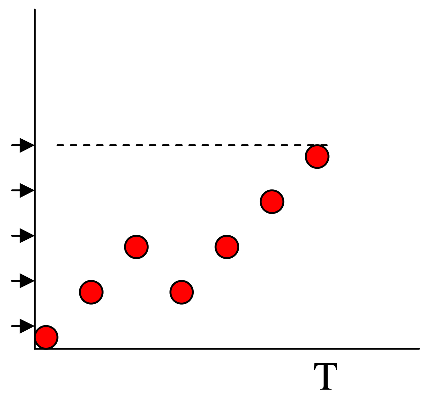
Let [$Y(t)$: t=0,1,2,…] be a stochastic process.
T a random variable taking non-negative integer values. Then T is said to be a stopping time for {$Y(t)$} if for every n one can determine whether [T=n] or not just by looking at $Y(0), Y(1), \ldots, Y(n)$.
Formally, $I_{\{T=n\}}$ is a function of $Y(0), Y(1), \ldots, Y(n)$.

Example. Simple random walk (unconstrained) starting at 0.

$$Y(t) = \sum_{i=1}^{t} X_i$$

where $P(X_i = -1) = 1/3$
$P(X_i = +1) = 2/3$

Let $T = \min\{t: Y(t) = 5\}$



Then $\{T=17\} \Leftrightarrow Y(0), Y(1), \ldots, Y(16) < 5$ and $Y(17) = 5$.

T is a stopping time for {Y(t)}.

Let $S = \max\{t: Y(t) = 5\}$.
Then we cannot unambiguously decide whether {S=17} or not by looking just at $Y(0), Y(1), \ldots, Y(17)$. So S is *not* a stopping time for {$Y(t)$}.

Theorem. (Wald's lemma)
Let $X_1, X_2, \ldots$ be a sequence of i.i.d. r.v. with $E(X_i) < \infty$, $Y(t) = \sum_{i=1}^{t} X_i$ and let T be a stopping time for the process $\{Y(t)\}$, with $E(T) < \infty$. Then

$E(Y(T)) = E(X_i)\ E(T)$, i.e. $E\left(\sum_{i=1}^{T} X_i\right) = E(X_i)E(T)$

Note that $\{T=t\}$ depends on $Y(1), Y(2), \ldots, Y(t) \Leftrightarrow \{T=t\}$ depends on $X_1, X_2, \ldots, X_t$ so would just say that T is stopping time for $\{X_i\}$

Proof. Let $Z_n \;=1$ if $T \geq n$
$=0$ otherwise (i.e. $T < n$)

Then $Y(T) = \sum_{n=1}^{T} X_n = \sum_{n=1}^{\infty} X_n Z_n$, so

$$E(Y(t)) = E\left(\sum_{n=1}^{\infty} X_n Z_n\right) = \sum_{n=1}^{\infty} E(X_n Z_n)$$

But $\{Z_n=0\} = \{T < n\}$ depends only on $X_1, X_2, \ldots, X_{n-1}$ and therefore is independent of X_n. Similarly $\{Z_n=1\} = \{Z_n=0\}^c$ is independent of X_n and hence X_n and Z_n are independent.

Consequently,

$$E(X_n Z_n) = E(X_n)E(Z_n) = E(X_n)P(T \geq n)$$

and

$$\begin{aligned} E(Y(t)) &= \sum_{n=1}^{\infty} E(X_n)P(T \geq n) \\ &= E(X_i)\sum_{n=1}^{\infty} P(T \geq n) = E(X_i)E(T) \end{aligned}$$

Application 1. Simple random walk with

$P(X_i=-1) = 1/3$
$P(X_i=1) = 2/3$

$T=\min\{t: Y(t)=5\}$, then easy to show that $E(T)<\infty$, $E(X_i)=1/3$.

Thus by Wald $E(Y(T))= E(X_i)\, E(T) = 1/3\, E(T)$.
However, $Y(T)=5$ hence $E(T)=15$.

Application 2. (Gambling with fair bet)

Simple random walk with $P(X_i=1) = P(X_i= -1) = ½$.
Thus under *any* strategy with finite expected time the expected profit is $E(Y(T)) = E(T)\, E(X_i) = \text{something} \times 0 = 0$

For example, if gambler starts with k and stops when either hits £0 or £N. Set $Y(0)=0$, so

$$T=\min\{ t: Y(t)= -k \text{ or } Y(t) = N-k \}$$

Can be shown that $E(T) < \infty$. Thus, since $E(Y(T)) = 0$, we have

$$-k\, P(Y(T)= -k) + (N-k)\, P(Y(T)=N-k) = 0$$

Let p_k=probability of ruin = $P(Y(T)= -k)$, then

$$-kp_k+(N-k)(1-p_k)=0 \Rightarrow$$

$$p_k = \frac{N-k}{N}$$

Suppose gambler has ∞ amount of capital and decides to stop when $Y(T)=1$. Then will stop with probability $=1$, and for sure wins £1. But $E(T)=\infty$ then! (Wald does ***not*** hold).

Chapter 4: total 4 lectures, 5 slides per lecture

5. Martingales

Review of conditional expectations

- Let $X_1,\ldots,X_n$ be random variables, and Y be a r.v.

Define $\mathrm{E}\,(Y|X_1,X_2,\ldots,X_n)$ to be the r.v. $\phi(X_1,X_2,\ldots,X_n)$
where $\phi(x_1,x_2,\ldots,x_n) = \mathrm{E}\,(Y \mid X_1=x_1,X_2=x_2,\ldots,X_n=x_n)$

- Let X, Y be random variables (for simplicity, non-negative integer valued)

Fact: $\mathrm{E}\,(Y)= \Sigma_{k=0}^{\infty}\, \mathrm{E}\,(Y \mid X=k)P(X=k)$, so if $\phi(x)= \mathrm{E}\,(Y|X=x)$
then $\mathrm{E}\,(Y)= \Sigma_{k=0}^{\infty}\, \phi(k)P(X=k)= \mathrm{E}\,(\phi(X))$

i.e. $\mathrm{E}\,(Y)= \mathrm{E}\,(\phi(X))$ where $\phi(X)= \mathrm{E}\,(Y|X)$ hence

$$\boxed{\mathrm{E}\,(Y)= \mathrm{E}\,(\mathrm{E}\,(Y|X)) \quad \text{(Law of Iterated Expectations)}}$$

Lemma. Let $X_1,X_2,\ldots,X_n$ be random variables, Z be any random variable, and $Y=f(X_1,X_2,\ldots,X_n)$. Then

$$\boxed{\mathrm{E}\,(YZ \mid X_1,X_2,\ldots,X_n) = Y\, \mathrm{E}(Z|X_1,X_2,\ldots,X_n)}$$

Proof. Let $\phi(x_1,x_2,\ldots,x_n) = \mathrm{E}\,(YZ \mid X_1=x_1,X_2=x_2,\ldots,X_n=x_n)$. Then

$$\begin{aligned}\phi(x_1,x_2,\ldots,x_n) &= \mathrm{E}\,(\, YZ \mid X_1=x_1,X_2=x_2,\ldots,X_n=x_n\,)\\ &= \mathrm{E}\,(\, f(X_1,X_2,\ldots,X_n)Z \mid X_1=x_1,X_2=x_2,\ldots,X_n=x_n)\\ &= \mathrm{E}\,(\, f(x_1,x_2,\ldots,x_n)Z \mid X_1=x_1,X_2=x_2,\ldots,X_n=x_n\,)\\ &= f(x_1,x_2,\ldots,x_n)\, \mathrm{E}\,(Z \mid X_1=x_1,X_2=x_2,\ldots,X_n=x_n)\end{aligned}$$

Therefore

$$\begin{aligned}\mathrm{E}\,(YZ \mid X_1,X_2,\ldots,X_n) &\overset{\text{defn}}{=} \phi(X_1,X_2,\ldots,X_n)\\ &= f(X_1,X_2,\ldots,X_n)\, \mathrm{E}\,(Z|X_1,X_2,\ldots,X_n) = Y\, \mathrm{E}(Z|X_1,X_2,\ldots,X_n)\end{aligned}$$

Corollary: if $Y=f\,(X_1,X_2,\ldots,X_n)$ then $\mathrm{E}\,(Y|X_1,X_2,\ldots,X_n) = Y$

Example: I roll a die. Let X be the score. Then I toss a coin X times, and let Y be the total number of heads obtained. Find $E(XY)$.

Solution: $E(XY) = E[E(XY|X)]$ by L.I.E.

But $E(XY|X) = X\,E(Y|X) = X \times X/2 = X^2/2$.

Hence $E(XY) = E(X^2/2) = \frac{1}{6}(1^2+2^2+3^2+4^2+5^2+6^2)/2$

> "Complimentary" fact: if Y is independent of $(X_1, X_2, \ldots, X_n)$, then $E(Y|X_1, X_2, \ldots, X_n) = E(Y)$

Definition: The stochastic process $[Y_n, n = 0, 1, 2 \ldots]$ is called **martingale** with respect to the sequence of random variables $[X_n, n = 0, 1, 2 \ldots]$, if for all $n \geq 0$,

(a) $E(|Y_n|) < \infty$;

(b) $E(Y_{n+1} \mid X_1, X_2, \ldots, X_n) = Y_n$

(in (b): *super***martingale**. if $\leq$, *sub***martingale** if $\geq$)

In the following example with gambler, the sequence of X's coincides with the sequence of Y's – this is allowed ("martingale with respect to itself").

Note: this implies by law of iterated expectations

$$E(Y_{n+1}) = E\left[E(Y_{n+1} \mid X_1, X_2, \ldots, X_n)\right] = E(Y_n)$$

for all n (replaced by $\leq$ or $\geq$ for super- or sub-martingales resp.)

Martingale examples

Example 1: Symmetric Random Walk (unrestricted)

Let $X_1, X_2, \ldots$ be i.i.d. r.v. with

$$\mathrm{P}(X_i=1) = \mathrm{P}(X_i=-1) = ½$$

Fix k, and let $Y_0=k$ and for $n \geq 1$,

$$Y_n=k+X_1+X_2+\ldots+X_n$$

Then $[Y_n, n=0,1,2,\ldots]$ is the symmetric random walk starting at k.

Note $Y_{n+1}=Y_n+X_{n+1}$.

Thus

$$\mathrm{E}\,(Y_{n+1}|X_1,X_2,\ldots,X_n)$$

$$= \mathrm{E}\,(Y_n|X_1,X_2,\ldots,X_n) + \mathrm{E}\,(X_{n+1}|X_1,X_2,\ldots,X_n)$$

$$=Y_n+ \mathrm{E}\,(X_{n+1})= Y_n$$

By corollary to lemma

Since X_{n+1} is independent of $X_1,\ldots,X_n$

Since $\mathrm{E}\,(X_{n+1})=0$, $\mathrm{E}\,(Y_{n+1}|X_1,X_2,\ldots,X_n) = Y_n$.

Hence Y_n is martingale with respect to $X_1,X_2,\ldots$.

Example 2: Symmetric Random Walk $[Y_n, n = 0,1,2\ldots]$ again

Let $Z_n = Y_n^2 - n, \quad n \geq 0$.

Look at the process $[Z_n, n=0,1,2\ldots]$

$\mathrm{E}\,(Z_{n+1} | X_1, X_2, \ldots, X_n)$

$$= \mathrm{E}\,(Y_{n+1}^2 - (n+1) \mid X_1, X_2, \ldots, X_n)$$

(recall that $Y_{n+1} = Y_n + X_{n+1}$)

$$= \mathrm{E}\,((Y_n + X_{n+1})^2 - (n+1) \mid X_1, X_2, \ldots, X_n)$$

$$= \mathrm{E}\,(Y_n^2 + X_{n+1}^2 + 2\, X_{n+1} Y_n - n - 1 \mid X_1, X_2, \ldots, X_n)$$

$$= Y_n^2 + 2\, Y_n\, \mathrm{E}\,(X_{n+1} | X_1, X_2, \ldots, X_n) + \mathrm{E}\,(X_{n+1}^2 \mid X_1, X_2, \ldots, X_n) - n - 1$$

$$= Y_n^2 + 2\, Y_n\, \mathrm{E}\,(X_{n+1}) + \mathrm{E}\,(X_{n+1}^2) - n - 1 = Y_n^2 + 0 + 1 - n - 1$$

$$= Y_n^2 - n = Z_n$$

i.e. $\quad \mathrm{E}(Z_{n+1} | X_1, X_2, \ldots, X_n) = Z_n$

Example 3: Gambler

Gambler places £1 on an even bet for the 1^{st} game, and henceforth £2^{n-1} on the n^{th} bet. If wins, he stops immediately.

Gambler must eventually win, and when s/he does, the net profit is always £1, since

$$2^n - (1+2+2^2+\ldots+2^{n-1}) = 1$$

Yet, if you let Y_n = the cumulative gain after n^{th} game, then

$$Y_{n+1} = \begin{cases} Y_n - 2^{n-1}, & prob = \frac{1}{2} \\ Y_n + 2^{n-1}, & prob = \frac{1}{2} \end{cases}$$

So that $E(Y_{n+1} \mid Y_1, Y_2, \ldots, Y_n) = Y_n$

(St Petersburg paradox)

Example 4: Branching process

N_j size of the j^{th} generation.
Assume $N_0=1$.

Let $\mu= E(N_1)$ - expected number of children per member.

Recall $E(N_j)=\mu^j$.

Suppose $N_j = n$ then

$$N_{j+1} = N_1^{(1)} + N_1^{(2)} + \ldots + N_1^{(n)}$$

$N_1^{(i)}$ have the same distribution as N_1

So $E(N_{j+1} \mid N_j = n) = \mu\, n$

$\Rightarrow$ by definition $E(N_{j+1} \mid N_j)= \mu N_j$,

hence also $E(N_{j+1} \mid N_j, N_{j-1},\ldots, N_1)= \mu N_j$

since $\{ N_j \}$ is a Markov process.

Define $W_j = N_j / E(N_j) = N_j / \mu^j$ then

$E(W_{j+1} \mid N_1,N_2,\ldots,N_j) = E(N_{j+1} / \mu^{j+1} \mid N_1,N_2,\ldots,N_j) = \mu\, N_j / \mu^{j+1}$

Hence $E(W_{j+1} \mid N_1,N_2,\ldots,N_j) = W_j$ $\Rightarrow$ martingale w.r.t. N_i

Martingale results

Recall: Time T is a stopping time if for all $n \geq 0$ event $\{T=n\}$ depends on $X_1, X_2, \ldots, X_n$ only.

Optional Stopping Theorem.

Let $[Y_n, n = 0, 1, 2 \ldots.]$ be a martingale with respect to the sequence of $[X_n, n = 0, 1, 2 \ldots.]$, and let T be a stopping time for $[X_n, n = 0, 1, 2 \ldots.]$.

Suppose furthermore

(a) $\mathbb{P}(T<\infty) = 1$;

(b) $\mathbb{E}(|Y_T|) < \infty$;

(c) $\mathbb{E}(|Y_n| \times I_{\{T>n\}}) \to 0$ as $n \to$

$$I_{\{T>n\}} = \begin{cases} 1 & if \quad T > n \\ 0 & if \quad T \leq n \end{cases}$$

Then $\mathbb{E}(Y_T) = Y_0$

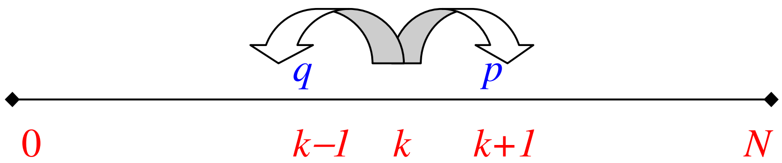
Martingale Convergence Theorem.

Let $[Y_n, n = 0, 1, 2 \ldots.]$ be a martingale (or *sub*martingale) with respect to the sequence of $[X_n, n = 0, 1, 2 \ldots.]$.
Suppose there is $A > 0$ such that $\mathbb{E}(|Y_n|) \leq A$ for all $n \geq 0$.

Then there exists r.v. Y such that $P(\lim_{n\to\infty} Y_n = Y) = 1$

Note: If Z_n is a *non-negative super*martingale with respect to the sequence of $[X_n]$, let $Y_n := -Z_n$. Then Y_n turns out to be a *sub*martingale with respect to the very same sequence, with $\mathbb{E}(|Y_n|) = \mathbb{E}(Z_n) \leq \mathbb{E}(Z_{n-1}) \leq \ \ldots \ \leq \mathbb{E}(Z_1) \leq \mathbb{E}(Z_0) = A$.
Hence Y_n converges a.s. by MCT, hence $Z_n = -Y_n$ also converges.

Application of OST and MCT to Gambler's ruin



Let $X_1, X_2, \ldots$ be i.i.d. r.v.

$\mathbb{P}(X_i = 1) \quad = p$
$\mathbb{P}(X_i = -1) = q = 1 - p$

Let $S_0 = k$, where $1 \le k \le N$ and for $n \ge 1$ set $S_n = k + X_1 + \ldots + X_n$.

Thus $[S_n]$ is an unrestricted random walk starting at k.

Now let

$$T = \begin{cases} \min\{n : S_n = 0 \quad or \quad S_n = N\} & \\ \infty, & if \quad 1 \le S_n \le N-1 \quad for \quad all \quad n \end{cases}$$

Set $Y_n = S_{\min\{n,T\}}$
Then $[Y_n, n = 0, 1, 2 \ldots.]$ is a random walk with absorbing barriers at 0 and N.

In particular, if $1 \le i \le N-1$

$$\mathbb{P}(Y_{n+1} = i-1 \mid Y_n = i) = q$$
$$\mathbb{P}(Y_{n+1} = i+1 \mid Y_n = i) = p$$

Also

$$\mathbb{P}(Y_{n+1} = 0 \mid Y_n = 0) = 1$$
$$\mathbb{P}(Y_{n+1} = N \mid Y_n = N) = 1$$

Lemma. Suppose $p=q=½$ (fair bet). Then

(a) $\mathbb{P}(T<\infty) = 1$,

(b) $\mathbb{P}(Y_T = N) = k / N$,

(c) $\mathbb{E}(T) = k(N-k)$.

Proof.

Since $p=q=½$, we have

$$\mathbb{E}(Y_{n+1} \mid Y_n=i)=i$$

whenever $1 \le i \le N-1$, and also

$$\mathbb{E}(Y_{n+1} \mid Y_n=0)=0$$
$$\mathbb{E}(Y_{n+1} \mid Y_n=N)=N$$

Thus for all i we have $\mathbb{E}(Y_{n+1} \mid Y_n)=Y_n$ and therefore, $\mathbb{E}(Y_{n+1} \mid Y_1,\ldots,Y_n)=Y_n$ since the process is Markov. Consequently, $[Y_n, n = 0, 1, 2 \ldots.]$ is a martingale with respect to itself.

(a) Note $0 \le Y_n \le N$ for all $n \Rightarrow \mathbb{E}(|Y_n|) \le N =: A$ for all n.

So by MCT there is random variable Y such that $Y_n \to Y$ a.s.

Let $A=\{\omega \in \Omega: Y_n(\omega) \to Y(\omega)\}$, then $\mathbb{P}(A)=1$

Suppose $T(\omega)=\infty$; i.e. $1 \le S_n(\omega) \le N-1$ for all n

Then $Y_{n+1}(\omega)=Y_n(\omega)+X_{n+1}(\omega)$ for all n

$\Rightarrow | Y_{n+1}(\omega) - Y_n(\omega)| = 1$

$\Rightarrow$ the sequence $[Y_n(\omega), n = 0,1,\ldots]$ cannot converge to a limit

$\Rightarrow \omega \in A^c$

$\Rightarrow \{T=\infty\} \subseteq A^c \quad \Rightarrow 0 \le \mathbb{P}(T=\infty) \le \mathbb{P}(A^c) = 0$

Therefore, $\mathbb{P}(T=\infty)=0 \Rightarrow \mathbb{P}(T<\infty) = 1$.

(b) Look at conditions of OST. Have shown $\mathbb{P}\,(T<\infty)=1$, also know that $Y_T \in \{0, N\}$ whence $\mathbb{E}|Y_T| < \infty$ as required.

Finally,
$0 \leq \mathbb{E}\,(Y_n I_{\{T>n\}}) \leq N\,\mathbb{E}\,(I_{\{T>n\}}) = N\,\mathbb{P}\,(T > n) \to 0$ as $n \to \infty$,
since $\mathbb{P}\,(T < \infty)=1$ implies $P(T > n) \to 0$ as $n \to \infty$.
Thus the conditions of the OST hold.
Hence $\mathbb{E}\,(Y_T) = Y_0 = k$.

At the same time

$$\mathbb{E}\,(Y_T) = 0 \times \mathbb{P}\,(Y_T=0) + N \times \mathbb{P}\,(Y_T= N) = N\,\mathbb{P}\,(Y_T = N)$$

So $\quad N\,\mathbb{P}\,(Y_T= N)= k \Rightarrow \quad \mathbb{P}\,(Y_T=N) = k / N$

(c) Let $Z_n=S_n^{\;2} - n$. (Recall: $S_n = k+X_1+\ldots+X_n$.)

Then $[Z_n, n = 0, 1, 2 \ldots.]$ is a martingale
with respect to $[X_n, n = 0, 1, 2 \ldots.]$

The conditions of the OST are met (easy to check!)

Then

$$\mathbb{E}\,(Z_T) = Z_0 = k^2 - 0 = k^2$$

Consequently $\quad k^2 = \mathbb{E}\,(Z_T) = \mathbb{E}\,(S_T^{\;2} - T) = \mathbb{E}\,(S_T^{\;2}) - \mathbb{E}\,(T)$

But

$$\mathbb{E}\,(S_T^{\;2}) = 0 \times P(S_T=0) + N^2 \times P(S_T=N) = N^2 \times k / N = Nk$$

Therefore $\quad \mathbb{E}\,(T) = \mathbb{E}\,(S_T^{\;2}) - k^2 = Nk - k^2 = k\,(N - k)$

QED

Lemma. Consider gambler's ruin when $p \neq q$. Let Y_n be the capital of the gambler after n plays.
Set

$$V_n = \left(\frac{q}{p}\right)^{Y_n}$$

Then $[V_n, n = 0, 1, 2 \ldots.]$ is a martingale with respect to $[Y_n, n = 0, 1, 2 \ldots.]$

Proof.

$$\mathrm{E}(V_{n+1} \mid Y_1, Y_2, \ldots, Y_n) = \mathrm{E}\left(\left(\frac{q}{p}\right)^{Y_{n+1}} \mid Y_1, Y_2, \ldots, Y_n\right) = \mathrm{E}\left(\left(\frac{q}{p}\right)^{Y_{n+1}} \mid Y_n\right)$$

since $[Y_n]$ is Markov.

Thus we need to show

$$\mathrm{E}\left(\left(\frac{q}{p}\right)^{Y_{n+1}} \mid Y_n\right) = V_n = \left(\frac{q}{p}\right)^{Y_n}$$

$$\Leftrightarrow$$

$$\mathrm{E}\left(\left(\frac{q}{p}\right)^{Y_{n+1}} \mid Y_n = y\right) = \left(\frac{q}{p}\right)^{y}$$

First, if $Y_n = y = 0$ then $Y_{n+1} = 0$ as well, hence

$$\mathrm{E}\left(\left(\frac{q}{p}\right)^{Y_{n+1}} \mid Y_n = 0\right) = \mathrm{E}\left(\left(\frac{q}{p}\right)^{0} \mid Y_n = 0\right) = \mathrm{E}(1 \mid Y_n = 0) = 1 = \left(\frac{q}{p}\right)^{0} = \left(\frac{q}{p}\right)^{y}$$

Similarly, if $Y_n = y = N$ then $Y_{n+1} = N$, so

$$\mathrm{E}\left(\left(\frac{q}{p}\right)^{Y_{n+1}} \mid Y_n = N\right) = \mathrm{E}\left(\left(\frac{q}{p}\right)^{N} \mid Y_n = N\right) = \left(\frac{q}{p}\right)^{N} = \left(\frac{q}{p}\right)^{y}$$

Now suppose $Y_n = y$ where $1 \le y \le N-1$. Then

$$\mathrm{E}\left(\left(\frac{q}{p}\right)^{Y_{n+1}} \mid Y_n = y\right) = \mathrm{E}\left(\left(\frac{q}{p}\right)^{Y_n + X_{n+1}} \mid Y_n = y\right)$$

$$= \mathrm{E}\left(\left(\frac{q}{p}\right)^{y + X_{n+1}} \mid Y_n = y\right) = \mathrm{E}\left(\left(\frac{q}{p}\right)^{y + X_{n+1}}\right) = \left(\frac{q}{p}\right)^{y} \mathrm{E}\left(\left(\frac{q}{p}\right)^{X_{n+1}}\right)$$

since X_{n+1} is independent of Y_n, and $Y_{n+1} = Y_n + X_{n+1}$.

Now

$$\mathrm{E}\left(\left(\frac{q}{p}\right)^{X_{n+1}}\right) = \left(\frac{q}{p}\right)^{-1} \mathrm{P}(X_{n+1} = -1) + \left(\frac{q}{p}\right)^{+1} \mathrm{P}(X_{n+1} = +1)$$

$$= \frac{p}{q} q + \frac{q}{p} p = p + q = 1$$

Thus

$$\mathrm{E}\left(\left(\frac{q}{p}\right)^{Y_{n+1}} \mid Y_n = y\right) = \left(\frac{q}{p}\right)^{y} \qquad \text{Q.E.D.}$$

Lemma. Consider the gambler ruin problem with $p \neq q$. Let T be the time at which game ends, and $Y_0 = k$.
Then

$$(a) \qquad P(T < \infty) = 1$$

$$(b) \qquad P(Y_T = N) = \frac{1 - \left(q/p \right)^k}{1 - \left(q/p \right)^N}$$

Proof.

Let $V_n = (q/p)^{Y_n}$. Then $[V_n, n = 0, 1, 2 \ldots.]$ is a martingale. Note that V_n lies somewhere between 1 and $(q/p)^N$, thus

$$E(|V_n|) < \max\{1, (q/p)^N\} < \infty.$$

So the conditions of the MCT hold. Hence there is a random variable V such that $V_n \to V$ a.s. as $n \to \infty$.

As before, $\{T = \infty\} \Rightarrow \{V_n(\omega) \text{ does not converge}\}$, whence $P(T = \infty) = 0$.

Conditions of OST hold, so we apply it to V_n and T.

Then $\qquad E(V_T) = V_0 = (q/p)^k$

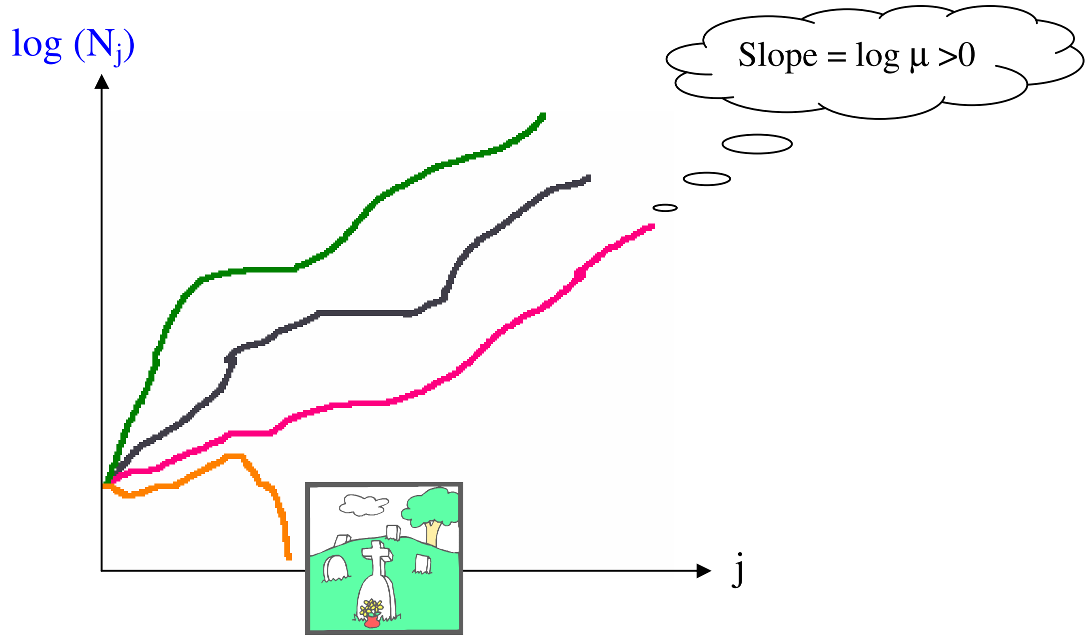
But $\qquad E(V_T) = [1 - P(Y_T = N)] \times (q/p)^0 + P(Y_T = N) \times (q/p)^N$
$$= 1 - P(Y_T = N)\,[1 - (q/p)^N]$$

from which the lemma follows.

Applications to supercritical Branching processes ($\mu > 1$)

Motivation:

N_j size of generation j.
Know $\mathrm{E}\,(N_j) = \mu^j$ so $\log \mathrm{E}\,(N_j) = j \log(\mu)$
Look at behaviour of $\log N_j$ versus j



Expect either $N_j = 0$ for j large or otherwise $\log N_j \approx j \log \mu + C$

random variable

$\log(N_j) - j \log \mu \to C$ as $j \to \infty$, i.e. $\log(N_j/\mu^j) \to C$
$\Rightarrow$

$$\frac{N_j}{\mu^j} \to e^C = K$$

random variable

and when $N_j = 0$,

$$\frac{N_j}{\mu^j} \to 0$$

How to prove this?

Let $W_j = N_j / \mu^j$

By the example on branching process, $E(W_{j+1} | N_1, N_2, \ldots, N_j) = W_j$ so that W_j is a martingale w.r.t. $[N_j, j=0,1,2,\ldots]$

We have

$$E(|W_j|) = E(W_j) = E\left(\frac{N_j}{\mu^j}\right) = 1$$

Thus $E(|W_j|)$ is uniformly bounded; consequently, we can apply MCT:

There exists random variable W such that

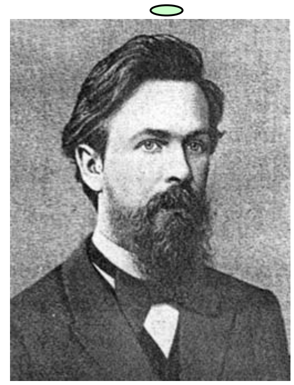
$$W_j \xrightarrow{a.s.} W \qquad as \quad j \to \infty$$

That is to say $P(N_j / \mu^j \to W) = 1$

$\Rightarrow$ Either the process becomes extinct ($W=0$)
or indeed $N_j \approx \text{Const} \times \mu^j$

Chapter 5: total 3 lectures, 5 slides per lecture

6. Markov Chains



Introduction

$[X_n: n=0,1,2,\ldots]$ stochastic process

X_n taking values in a finite or countable state space S, that is $X_n \in S$.

In *this course*, we usually take

$$S \subseteq Z = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$$

Recall

Definition. A stochastic process $\{X_t, t=0,1,\ldots,\}$ is called ***Markov* chain**, if for every n and m such that m, n≥0 and any collection of states $i_1, i_2, \ldots \in S$ and $j \in S$

$$\mathbb{P}(X_{n+m}=j \mid X_n=i_n, X_{n-1}=i_{n-1}, \ldots, X_0=i_0) = \mathbb{P}(X_{n+m}=j \mid X_n=i_n)$$

We assume that Markov Chain is time-homogeneous, that is

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

One-step transition probability

does not ever depend on n.

Matrix $P=\{p_{ij}\}$ is called **transition matrix** of the chain.

Note that

$$\sum_{j \in S} p_{ij} = 1 \qquad for \quad each \quad i \in S$$

i.e. row sum is 1 for every row. Also $0 \le p_{ij} \le 1$.

Example. Unrestricted random walk

Let $X_1, X_2, \ldots$ be i.i.d. integer-valued r.v., $X_i = \begin{cases} +1 & prob = p \\ -1 & prob = 1-p \end{cases}$

Fix k, and let $Y_0 = k$ and for $n \geq 1$,

$Y_n = k + X_1 + X_2 + \ldots + X_n$

Then $[Y_n, n = 0, 1, 2, \ldots]$ is a Markov process.

Proof.

$$\mathrm{P}(Y_{n+m} = j \mid Y_0 = i_0, Y_1 = i_1, \ldots, Y_n = i_n)$$

$$= \mathrm{P}\left(k + \sum_{i=1}^{n+m} X_i = j \mid Y_0 = i_0, \ldots, Y_n = i_n \right)$$

$$= \mathrm{P}\left(Y_n + \sum_{i=n+1}^{n+m} X_i = j \mid Y_0 = i_0, \ldots, Y_n = i_n \right)$$

Also because of independence

$$= \mathrm{P}\left(\sum_{i=n+1}^{n+m} X_i = j - i_n \mid Y_0 = i_0, \ldots, Y_n = i_n \right)$$

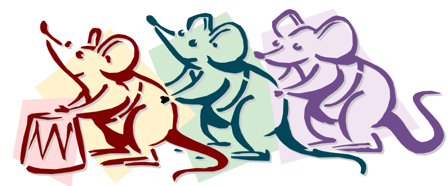
Because of independence

$$= \mathrm{P}\left(\sum_{i=n+1}^{n+m} X_i = j - i_n \right) = \mathrm{P}\left(i_n + \sum_{i=n+1}^{n+m} X_i = j \mid Y_n = i_n \right)$$

$$= \mathrm{P}\left(Y_n + \sum_{i=n+1}^{n+m} X_i = j \mid Y_n = i_n \right) = \mathrm{P}(Y_{n+m} = j \mid Y_n = i_n)$$

Example. $[Y_n : n \geq 0]$ simple random walk with absorbing barriers at 0 and N.

$$P = \begin{bmatrix} 1 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ q & 0 & p & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & q & 0 & p & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & q & 0 & p \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 & 1 \end{bmatrix}$$



Single Server Queue

$\xrightarrow{\textit{ARRIVALS}}$ $\hbar$

$\hbar$

$\hbar$

$\hbar$

$\hbar$

$\hbar$ $\xrightarrow{\textit{DEPARTURES}}$

Queue characterized by (i) distribution of interarrival time;
(ii) distribution of service time.

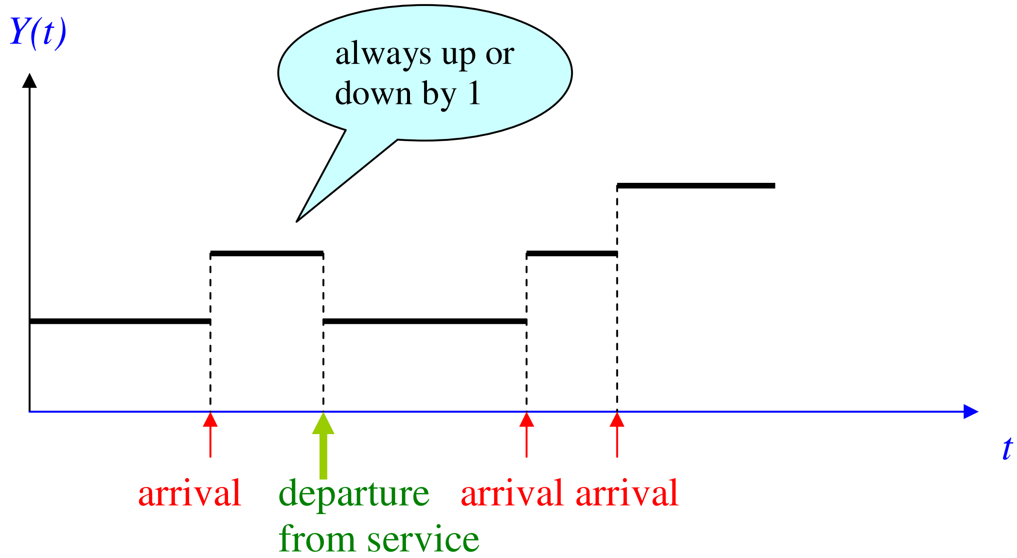
Exponential *inter*arrival times

Exponential service time

1 server

Special case: M/M/1 queue

Let $Y(t)$ = queue size at time t (as time progresses)



Let X(n) =size of queue prior to n^{th} change of size
(X(n) is called the *embedded chain*)

Suppose that

- interarrival times independent exponential (λ)
- service times independent exponential (μ)

Then P (nearest change is due to arrival)= λ / (λ+μ)

(Remember homework!)

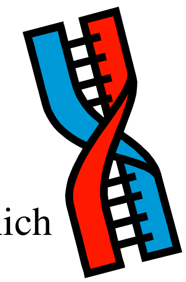
So

$$\left.\begin{aligned} \mathrm{P}\left(X(n+1)=i+1 \mid X(n)=i\right) &= \frac{\lambda}{\lambda+\mu} \\ \mathrm{P}\left(X(n+1)=i-1 \mid X(n)=i\right) &= \frac{\mu}{\lambda+\mu} \end{aligned}\right\} \quad i \geq 1$$

$$\mathrm{P}\left(X(n+1)=1 \mid X(n)=0\right)=1$$

$$P = \begin{array}{c|ccccc} & 0 & 1 & 2 & 3 & \cdots \\ \hline 0 & 0 & 1 & 0 & 0 & \cdots \\ 1 & \frac{\mu}{\mu+\lambda} & 0 & \frac{\lambda}{\mu+\lambda} & 0 & \cdots \\ 2 & 0 & \frac{\mu}{\mu+\lambda} & 0 & \frac{\lambda}{\mu+\lambda} & \cdots \\ 3 & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

Doubly infinite matrix:



Example. Change in gene frequency

Look at particular locus on a chromosome
Let X_n be the number of such loci in the population which contains the allele A after n generations.

Suppose:

- The population size is constant at N
- Given $X_n=i$, each member of the $n+1^{\text{st}}$ generation has A at the locus with probability i/N

Then

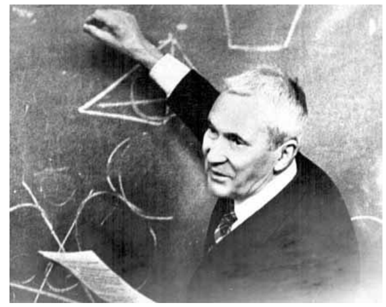
$$p_{ij} = \mathrm{P}(X_{n+1} = j \mid X_n = i) = \binom{N}{j}\left(\frac{i}{N}\right)^j\left(1-\frac{i}{N}\right)^{N-j} \qquad i, j \in S$$

Here $S=\{0,1,2,\ldots,N\}$ $\Leftarrow$ absorbing barriers at 0 and N

Andrey Nikolayevich Kolmogorov,
born April 25 1903, Tambov, Russia,
influenced many branches of modern mathematics,
especially harmonic analysis, **probability**, set
theory, information theory, and number theory

The Chapman – Kolmogorov Equations

Define $p_{ij}(n) = \mathbb{P}(X_{n+k}=j \mid X_k=i)$
i.e. n steps transition probabilities



Note: does not depend on k!

We have

$$p_{ij}(n+m) = P(X_{n+m}=j \mid X_0=i)$$
$$= \sum_{k\in S} P(X_{n+m}=j \mid X_0=i, X_n=k)\,P(X_n=k \mid X_0=i)$$

since Markov chain

$$= \sum_{k\in S} P(X_{n+m}=j \mid X_n=k)\,P(X_n=k \mid X_0=i) = \sum_{k\in S} p_{kj}(m)\,p_{ik}(n)$$

Let $P^{(n)} = \{p_{ij}(n)\}$. Then CK equations give

C.-K.

$$\boxed{P^{(n+m)} = P^{(n)}P^{(m)}}$$

Whence $P^{(n)} = P^{(n-1)}P^{(1)} = P^{(n-2)}P^{(1)}P^{(1)} = \ldots \Rightarrow P^{(n)} = P^{n}$

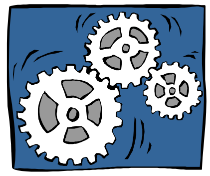
Let $\mathbb{P}(X_0=i) = p_i(0)$ and $\mathbb{P}(X_n=i) = p_i(n)$.

Denote $\vec{p}(n) = (p_0(n), p_1(n), \ldots)$ horizontal vector (maybe ∞ long)

Then $\mathbb{P}(X(n)=j) = \Sigma_i\, \mathbb{P}(X(n)=j \mid X(0)=i)\, \mathbb{P}(X(0)=i) = \Sigma_i\, p_i(0)p_{ij}(n)$

Thus $\vec{p}(n) = \vec{p}(0)\, P^{n}$

Example.



Each day a machine is either working or broken down. If it is working, it is also working next day with probability 0.9 ; and if it is broken, it remains broken next day with probability 0.75.

Let X_n = 0, if the machine is *broken* on day n
= 1, if the machine is working on day n

$[X_n: n=0,1,2,\ldots]$ is a Markov chain with transition probabilities

$$p_{00}=P(X_{n+1}=0 \mid X_n=0) = 3/4$$
$$p_{01}=P(X_{n+1}=1 \mid X_n=0) = 1/4$$
$$p_{10}=P(X_{n+1}=0 \mid X_n=1)= 0.1$$
$$p_{11}=P(X_{n+1}=1 \mid X_n=1)= 0.9$$

$$P = \begin{bmatrix} 0.75 & 0.25 \\ 0.10 & 0.90 \end{bmatrix}$$

The machine is working on day zero (given). What is the probability the machine is working on day n?
Namely, what's $P(X_n=1 \mid X_0=1)$? What happens as $n \to \infty$?

To answer this, we need to calculate P^n for $n = 2,3,\ldots$
But it's easier to look at eigenvalues of matrix P.

Solution:
Need to find eigenvectors $\underline{x}$ such that $\underline{x}\,P=\lambda\,\underline{x}$.
Find λ first:

$$\textbf{det}(P-\lambda I)=0 \Rightarrow \qquad (0.75-\lambda)(0.9-\lambda)-0.25\times 0.1=0$$

$$\lambda^2-1.65\lambda+0.65=0 \quad \lambda_{1,2}=1,\ 0.65$$

Note! $\lambda=1$ is *always* a solution of $\textbf{det}(P-\lambda I)=0$ for Markov chains.

Case λ_1=1. Let vector $\underline{\boldsymbol{x}}$ satisfy $\underline{\boldsymbol{x}}P = \lambda_1 \underline{\boldsymbol{x}}$, i.e.

$$\begin{pmatrix} x_0 & x_1 \end{pmatrix} \begin{bmatrix} 0.75 & 0.25 \\ 0.10 & 0.90 \end{bmatrix} = 1 \times \begin{pmatrix} x_0 & x_1 \end{pmatrix} = \begin{pmatrix} x_0 & x_1 \end{pmatrix}$$

$\Rightarrow$ $0.75\, x_0 + 0.1\, x_1 = x_0$ $\Rightarrow 0.1x_1 = 0.25x_0 \Rightarrow x_1 = 2.5\, x_0$

W.l.o.g. (=**w***ithout* **l***oss* **o***f* **g***enerality*) can assume $x_0 + x_1 = 1$

Thus $\underline{\boldsymbol{x}}$=(2/7, 5/7)

Case λ_2=0.65. Let vector $\underline{\boldsymbol{y}}$ satisfy $\underline{\boldsymbol{y}}P = \lambda_2 \underline{\boldsymbol{y}}$, i.e.

$$\begin{pmatrix} y_0 & y_1 \end{pmatrix} \begin{bmatrix} 0.75 & 0.25 \\ 0.1 & 0.9 \end{bmatrix} = 0.65 \times \begin{pmatrix} y_0 & y_1 \end{pmatrix} = \begin{pmatrix} 0.65\, y_0 & 0.65 y_1 \end{pmatrix}$$

$\Rightarrow$ $0.75\, y_0 + 0.1\, y_1 = 0.65\, y_0$ $\Rightarrow y_1 + y_0 = 0$ $\Rightarrow y_1 = -y_0$

W.l.o.g. $\underline{\mathbf{y}}$=(−1, 1) .

Now:

$\underline{\mathbf{p}}(0)$ = (0, 1) since initially machine is working.

We can write this in the new basis ($\underline{\mathbf{x}}$, $\underline{\mathbf{y}}$) as

$\underline{\mathbf{p}}(0) = \underline{\boldsymbol{x}} + 2/7\, \underline{\boldsymbol{y}}$

(by solving the equation $\vec{p}(0) = \alpha \quad \vec{x} \quad + \quad \beta \quad \vec{y}$ for α and β)

Since $\underline{\boldsymbol{x}}$ is an eigenvector corresponding to the eigenvalue λ_1, we have $\underline{\boldsymbol{x}}\, P = \lambda_1 \underline{\boldsymbol{x}}$ yielding

$$\underline{\boldsymbol{x}}\, P^{\,n} = (\,\underline{\boldsymbol{x}}\, P)\, P^{\,n-1} = (\,\lambda_1 \underline{\boldsymbol{x}})\, P^{\,n-1} = \lambda_1\, (\underline{\boldsymbol{x}}\, P^{\,n-1}) = \lambda_1\, [(\,\underline{\boldsymbol{x}}\, P)\, P^{\,n-2})]$$

$$= \lambda_1\, [(\,\lambda_1 \underline{\boldsymbol{x}})\, P^{\,n-2}] = \lambda_1^{\,2}\, (\underline{\boldsymbol{x}}\, P^{\,n-2}) = \ldots = \lambda_1^{\,n}\, \underline{\boldsymbol{x}}$$

Similarly, $\underline{\boldsymbol{y}}\, P^{n} = \lambda_2^{\,n}\, \underline{\boldsymbol{y}}$

Thus, since $\lambda_1=1$ and $\lambda_2= 0.65$, and $\underline{\mathbf{p}}(0) = \underline{\boldsymbol{x}} + 2/7\ \underline{\boldsymbol{y}}$

$$\underline{\mathbf{p}}(n) = \underline{\mathbf{p}}(0)P^{\,n} = \underline{\boldsymbol{x}}\, P^{\,n} + 2/7\ \underline{\boldsymbol{y}}\, P^{\,n} = \underline{\boldsymbol{x}}\ + 2/7\ (0.65)^{n}\, \underline{\boldsymbol{y}}$$

$$= (2/7,\, 5/7) + 2/7 \times (0{,}65)^{n}\, (-1,\, 1)$$

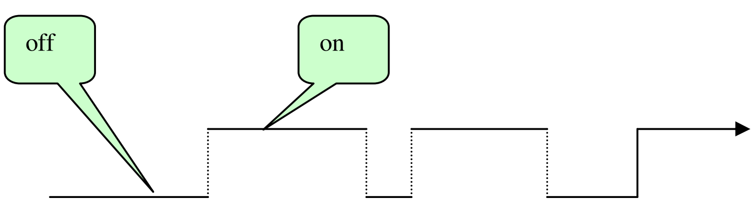
$$= \left(\frac{2}{7}\left(1 - 0.65^{n}\right) \quad , \quad \frac{5}{7} + \frac{2}{7} 0.65^{n} \right)$$

Consequently,

$$\mathbb{P}\ (X_n=1 \mid X_0=1) = 5/7 + 2/7 \times (0.65)^{n}$$

and this $\rightarrow 5/7$ as $n \rightarrow \infty$.

Heuristically:



Observe that mean working time between repairs is 10 days (=mean of geometric r.v. with rate $p=0.1$ is $1/p=1/\,0.1$) and mean repair time is 4 days ($= 1\, /\, 0.25$).

Hence proportion of working time is $10/(10+4) = 5/7$ again!

Classification of Markov Chains

Definition.

State j is *accessible* from state i if for some $n \geq 0$ $p_{ij}(n) > 0$.

Write $i \rightarrow j$.

Definition.

States i and j *communicate* if $i \rightarrow j$ and $j \rightarrow i$.

Write $i \leftrightarrow j$.

Theorem. Communications is an *equivalence class*, that is

(i) $i \leftrightarrow i$;
(ii) $i \leftrightarrow j$ if and only if $j \leftrightarrow i$;
(iii) if $i \leftrightarrow j$ and $j \leftrightarrow k$ then $i \leftrightarrow k$.

Proof. (i) and (ii) follow immediate from the definition. To prove (iii), observe that there are $m \geq 0$ and $n \geq 0$ such that $p_{ij}(m) > 0$ and $p_{jk}(n) > 0$, consequently by CK equations

$$p_{ik}(m+n) = \sum_{r \in S} p_{ir}(m) p_{rk}(n)$$
$$\geq p_{ij}(m) p_{jk}(n) > 0$$

Thus $i \rightarrow k$. Similarly prove $k \rightarrow i$.

If $i \leftrightarrow j$ then i,j are said to be in the same *communicating class*. Classes partition the state space S.

Definition. Markov chain is called *irreducible* if it has *only one* communicating class.

Definition. State i is said to have period d if

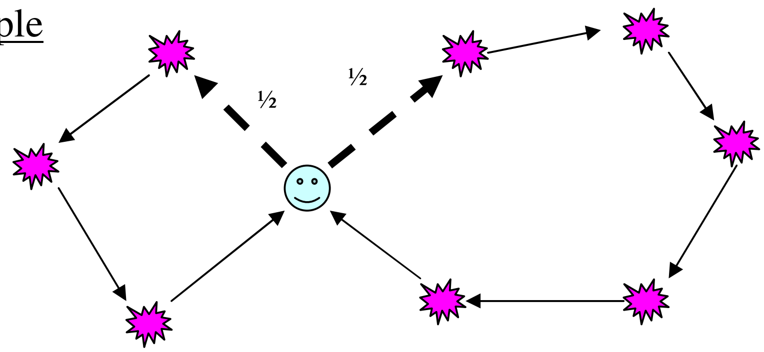
g.c.d. of $\{n: p_{ii}(n) > 0\} = \mathrm{d}$

Greatest Common Divisor

Notation. $\mathrm{d}(i) = \mathrm{d}$

Remark: if $\mathrm{d}(i) = \mathrm{d}$, then there exists $n_0 \geq 1$ such that $p_{ii}(n) > 0$ for $n \geq n_0$ if and only if n is divisible by d.

Example

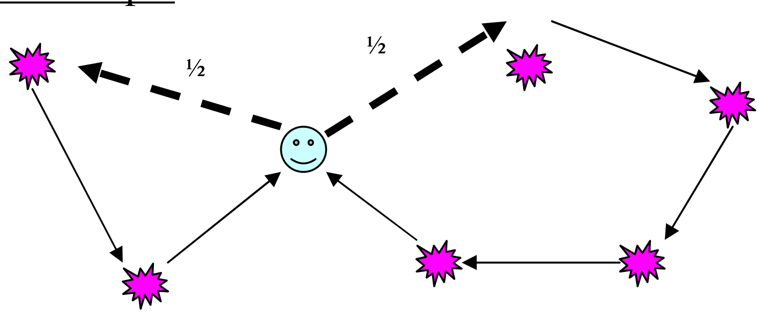


Look at return times to 

g.c.d.= 2, n_0=4

$p_{ii}(n) > 0$ for n =4, 6, 8, 10, 12, 14, …

Another example



Again look at return times to 

g.c.d.= 1, n_0=8

$p_{ii}(n) > 0$ for n =3, 5, 6, 8, 9, 10, …

Theorem. If $i \leftrightarrow j$ then $d(i) = d(j)$.

Proof. Let m and n be such that $p_{ij}(m) > 0$ and $p_{ji}(n) > 0$.
Suppose $p_{ii}(s) > 0$. Then

$$p_{jj}(m+n) \geq p_{ji}(n)\, p_{ij}(m) > 0$$
$$p_{jj}(m+n+s) \geq p_{ji}(n)\, p_{ii}(s)\, p_{ij}(m) > 0$$

Hence $d(j)$ divides $(m+n)$ and also divides $(m+n+s)$, consequently $d(j)$ divides s,

$\Rightarrow$ $d(j)$ divides g.c.d. $\{s: p_{ii}(s) > 0\} = d(i)$

$\Rightarrow$ $d(j)$ divides $d(i)$

By similar argument $d(i)$ divides $d(j)$, hence $d(j)=d(i)$. Q.E.D.

Definition. For states i,j set

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

Definition. State j is called *recurrent* (=persistent) if $f_{jj}=1$. Otherwise call it *transient*.

Remark. Suppose $X_0 = j$ and let N be the number of times the chain returns to state j. Then $f_{jj} = P\,(N \geq 1)$.

Case $f_{jj}<1$ (transient)

$\Rightarrow P\,(N \geq k) = (f_{jj})^k$. Thus N is geometric and $P\,(N=\infty)=0$.

Case $f_{jj}=1$ (recurrent)

$\Rightarrow P\,(N=\infty)=1$.

Remark. Suppose $X_0=i$, and let N be the number of times the chain visits state j.

Then $\qquad P\,(N \geq k) = f_{ij}(f_{jj})^{k-1}, \qquad$ for $k \geq 1$

In particular, if j is transient, $P\,(N=\infty)=0$.

Corollary. Suppose S has finitely many states. Then at least one state is recurrent.

Indeed, when S is finite, at least one state must be visited ∞ many times with positive probability, hence it cannot be transient.

Example. S={1,2,3,4,5,6,7} Transition matrix:

$$
P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array}
\begin{array}{c}
\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\
\left|\begin{array}{ccccccc}
0 & 1/3 & 0 & 2/3 & 0 & 0 & 0 \\
0 & 0 & 1/3 & 2/3 & 0 & 0 & 0 \\
1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 \\
0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1/2 & 0 & 1/2
\end{array}\right|
\end{array}
$$

What are the equivalence classes of communicating states?

$$
\begin{tikzcd}
1 \arrow[r] & 2 \arrow[r] & 3 \arrow[l, bend right] \arrow[ll, bend left]
\end{tikzcd}
$$

Thus {1, 2, 3} is such class

- $1 \to 4$
- $2 \to 4$
- $4 \to 5$
- $4 \to 6$
- $5 \to 4$
- $5 \to 6$
- $6 \to 6$
- $7 \to 5$
- $7 \to 7$

{4, 5} class

{6} class $\Leftarrow$ the *only* recurrent one

{7} class

Periods of states:

{1, 2, 3} return to {1} after 3, 5, 6… steps $\Rightarrow$ d(1)=d(2)=d(3)=1

{4, 5} return to {4} after 2, 4, 6, … steps $\Rightarrow$ d(4)=d(5)=2

{6} $\Rightarrow$ d=1 and {7} $\Rightarrow$ d=1

All classes except {6} are transient.

More on recurrence

Let $f_{jj}(n) = \mathbb{P}(X_1 \neq j, X_2 \neq j, \ldots, X_{n-1} \neq j, X_n = j \mid X_0 = j)$. Thus

$$f_{jj} = \sum_{n=1}^{\infty} f_{jj}(n)$$

p.m.f. of 1st return time $\tau_{jj} = \inf\{k \geq 1 : X_k = j\}$ when $X_0 = j$

Let

$$F_{jj}(s) = \sum_{n=1}^{\infty} s^n f_{jj}(n), \qquad |s| \leq 1$$

Recall that $p_{jj}(n) = \mathbb{P}(X_n = j \mid X_0 = j)$, and let

$$P_{jj}(s) = \sum_{n=0}^{\infty} s^n p_{jj}(n), \qquad |s| < 1$$

Lemma. $P_{jj}(s) = 1 + P_{jj}(s)\, F_{jj}(s)$.

Proof. As for the random walk, we have

$$p_{jj}(m) = \sum_{r=1}^{m} f_{jj}(r)\, p_{jj}(m-r) \ldots$$

… hence the result.

Corollary. State j is recurrent $\Leftrightarrow \sum_{n \geq 0} p_{jj}(n) = \infty$.

Proof. j is recurrent $\Leftrightarrow f_{jj} = 1$ (by definition) $\Leftrightarrow \sum_n f_{jj}(n) = 1$

$$\Leftrightarrow \sum_{n=1}^{\infty} f_{jj}(n) = \lim_{s \uparrow 1} \sum_{n=1}^{\infty} f_{jj}(n) s^n = \lim_{s \uparrow 1} F_{jj}(s)$$

But $F_{jj}(s) = 1 - 1 / P_{jj}(s)$ by Lemma, so

$$\lim_{s \uparrow 1} F_{jj}(s) = 1 \Leftrightarrow \lim_{s \uparrow 1} P^{-1}{}_{jj}(s) = 0 \Leftrightarrow \lim_{s \uparrow 1} P_{jj}(s) = \infty,$$

$$\Leftrightarrow \lim_{s \uparrow 1} \sum_{n=1}^{\infty} p_{jj}(n) s^n = \infty \Leftrightarrow \sum_{n=1}^{\infty} p_{jj}(n) = \infty \qquad \text{Q.E.D.}$$

Corollary. Suppose $i \leftrightarrow j$. Then i recurrent $\Leftrightarrow$ j recurrent.

Proof. Suppose i recurrent. Since $i \leftrightarrow j$ there are m and n such that $p_{ij}(n) > 0$ and $p_{ji}(m) > 0$.
From Chapman-Kolmogorov equation, applied twice, for $s \geq 0$,

$$p_{jj}(m+n+s) \geq p_{ji}(m)\, p_{ii}(s)\, p_{ij}(n)$$

Hence

$$\Sigma_s\, p_{jj}(m+n+s) \geq p_{ji}(m)\, p_{ij}(n)\, \Sigma_s\, p_{ii}(s)$$

But by above corollary, the sum on the RHS is infinite, so is

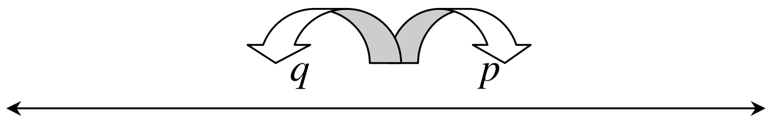
$$\Sigma_k\, p_{jj}(k) = \infty$$

and, by Corollary, j is also recurrent. Similar arguments work out for the counterpart.

Q.E.D.

Corollary. When S is finite, and the chain is irreducible, all states are recurrent.

Example. Simple random walk (unrestricted)



Then $i \leftrightarrow j$ for all i,j, so there is only 1 communicating class $\Rightarrow$ chain irreducible.
Returns to any state are possible after steps 2,4,6,… so $d(i)=2$ for all i. (So that $p_{00}(2n+1)=0$ for all n.)

Shown before,

$$p_{00}(2n) \sim \frac{(4pq)^n}{\sqrt{\pi n}} \quad as \quad n \to \infty$$

Case $p=q=½$ Then

$$p_{00}(2n) \sim \frac{1}{\sqrt{\pi n}} \quad \Rightarrow \quad \sum_{n=1}^{\infty} p_{00}(2n) = \infty$$

Thus state 0 is recurrent by Corollary, and hence all other states are recurrent too.

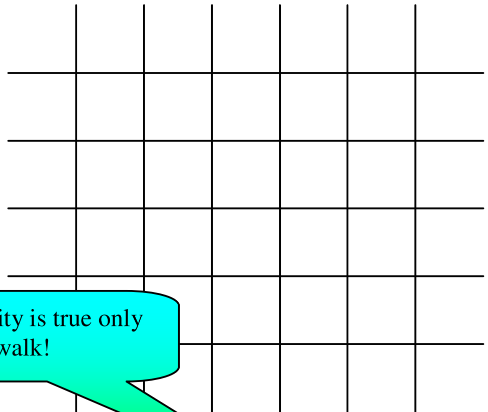
Case $p \neq q$. Then $4pq=\theta < 1$ and

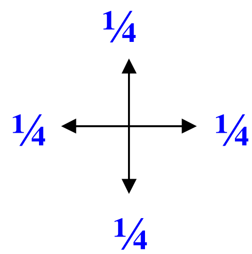
$$p_{00}(2n) \sim \frac{\theta^n}{\sqrt{\pi n}} \quad \Rightarrow \quad \sum_{n=1}^{\infty} p_{00}(2n) < \infty$$

Thus state 0 is transient by Corollary, and hence all other states are transient too.

Example.

2-dimensional simple symmetric random walk on $\mathbb{Z}^2$





This equality is true only for 2-dim walk!

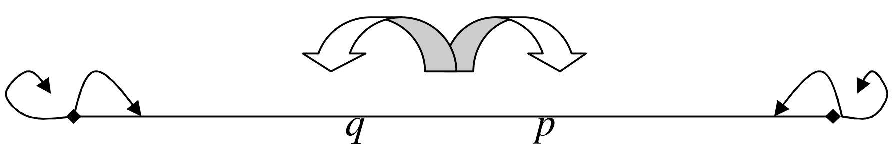
$$p^{(z2)}{}_{00}(2n) = \left[p^{(z1)}{}_{00}(2n)\right]^2 \sim \left(\frac{1}{\sqrt{\pi n}}\right)^2$$

$$\Rightarrow \qquad \sum_{n=1}^{\infty} p^{(z2)}{}_{00}(2n) = \infty$$

Thus state $\underline{\mathbf{0}} = (0,0)$ is recurrent by Corollary, and hence all other states are recurrent too.

Example.

Simple random walk with reflecting barriers at 0 and N



Then all states still communicate, i.e. $i \leftrightarrow j$ for all i,j, so the chain is irreducible (and also *aperiodic*, i.e. $d(i)=1$ since this is true for the barrier endpoints).

What happens with $\mathbb{P}(X_n=i)$ as $n \to \infty$?

- In a more general set up, I am interested in $\lim_{n\to\infty} p_{ij}(n)$

Limit theorems

First, suppose that j is transient.

Then can show $$p_{ij}(n) \to 0 \qquad \forall i \in S$$

So from now, we will deal with irreducible chains in which at least one, and hence every state is **recurrent**.

Definition. Suppose j is recurrent. Let

$$\mu_{jj} = \sum_{n=1}^{\infty} n f_{jj}(n)$$

be the mean time to return to j starting from j,
also called *the mean recurrence time of state j*.

Definition. Let j be a recurrent state. We call it

positive recurrent, if $\mu_{jj} < \infty$;
null recurrent, if $\mu_{jj} = \infty$.

Example. Simple *symmetric* random walk (unrestricted)

We know all states are recurrent. Starting from 0, mean return time is ∞; same for other states

$\Rightarrow$ *all states are null recurrent*

Theorem. Let j be a recurrent state and suppose $i \leftrightarrow j$. Then

(a) $p_{ij}(n) \to 1 / \mu_{jj}$ as $n \to \infty$ when j is aperiodic (i.e. $d(j)=1$)

(b) $p_{ij}(n) \to 1 / \mu_{jj}$ as $n \to \infty$ but only *in Cesàro sense* when j has period $d>1$, i.e. $d(j)=d$

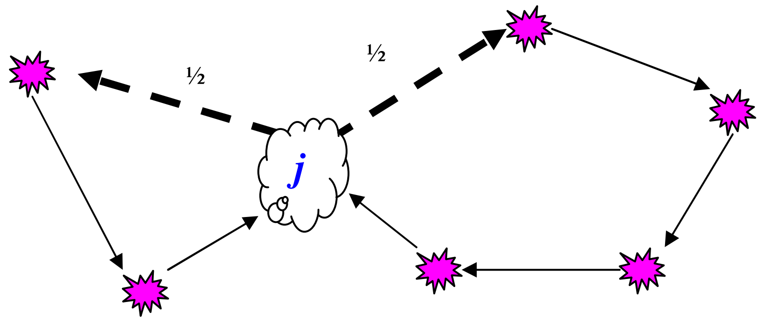
Def: $x_n \to x$ in Cesàro sense, if $(x_1 + x_2 + \ldots + x_n) / n \to x$ as $n \to \infty$.

Note: the above theorem holds even when $\mu_{jj}=\infty$.

Why (a) is this true?

Starting from i the walk eventually gets to j. Once at j it returns there ∞ many times. Mean time between returns is exactly μ_{jj}. Probability of returns even out until $P(X_n=j) \cong 1/\mu_{jj}$.

Example.



Look at state j. Suppose $X_0 = j$. Then

n	=	1	2	3	4	5	6	7
P $(X_n=j \mid X_0=j)$	=	0	0	½	0	½	¼	0

$d(j)=1$ so aperiodic. $\mu_{jj}=4$ so expect $p_{ij}(n) \to \frac{1}{4}$ for all i.

Note: in general, for $n \geq 5$,

$$
\begin{aligned}
&\mathrm{P}(X_n = j \mid X_0 = j) = \mathrm{P}(X_n = j \mid X_0 = j, X_{n-3} = j)\mathrm{P}(X_{n-3} = j \mid X_0 = j) \\
&+\mathrm{P}(X_n = j \mid X_0 = j, X_{n-5} = j)\mathrm{P}(X_{n-5} = j \mid X_0 = j) \\
&= \frac{1}{2}\mathrm{P}(X_{n-3} = j \mid X_0 = j) + \frac{1}{2}\mathrm{P}(X_{n-5} = j \mid X_0 = j)
\end{aligned}
$$

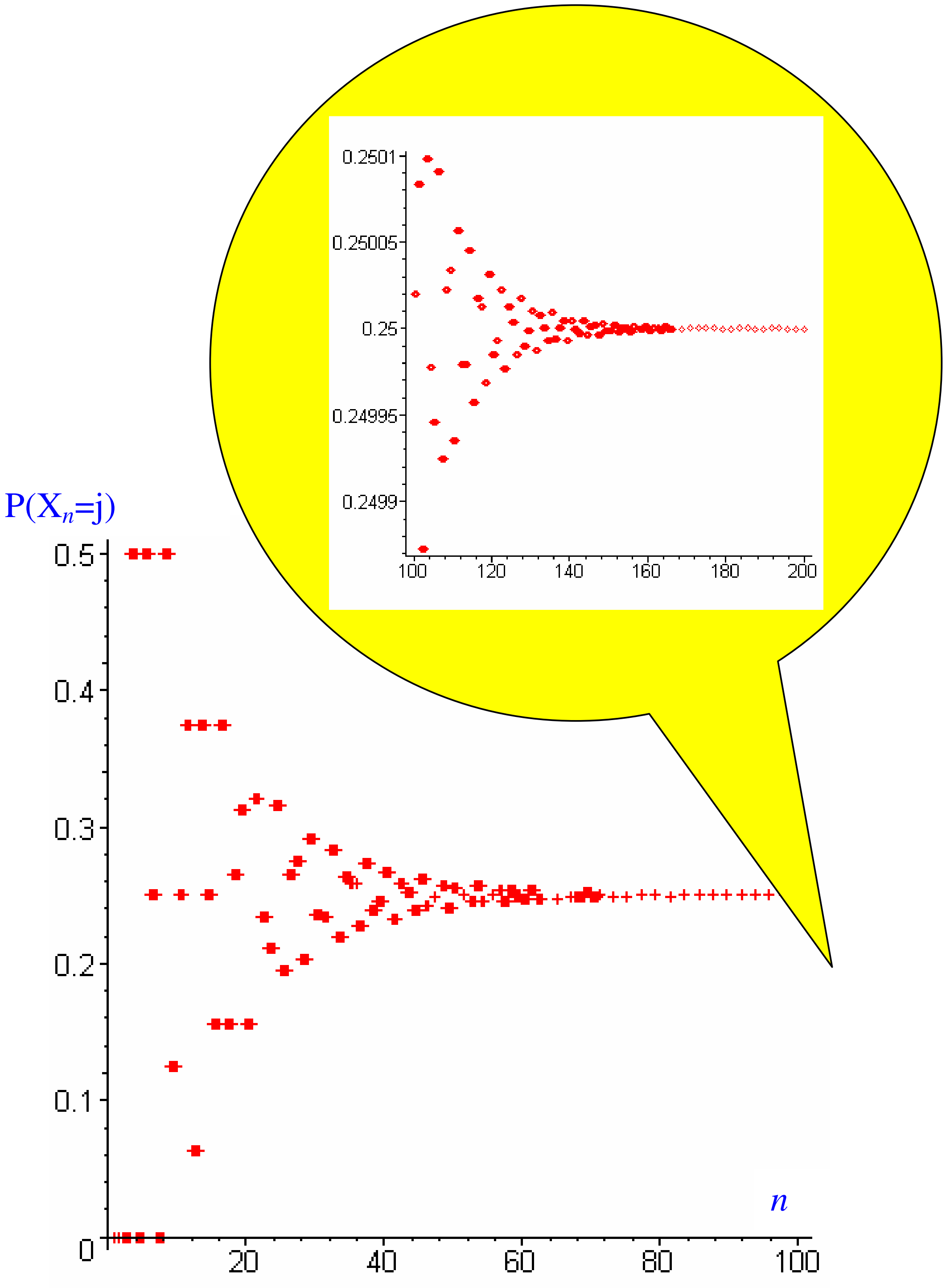
Using this recursion, we get

n	=	8	9	10	11	12	13	14
P $(X_n=j \mid X_0=j)$	=	½	1/8	¼	3/8	1/16	3/8	¼

and further

n	=	15	16	17	18	19	20	21
P $(X_n=j \mid X_0=j)$	=	5/32	3/8	5/32	17/64	5/16	5/32	41/128

etc. seems close to ¼ .



FYI almost exact formula: (beyond this course level)

$$p_n = \frac{1}{4} + 0.49(0.923)^n \cos(2.38n - 0.22) + 0.3(0.766)^n \cos(1.35n - 0.27)$$

<u>Lemma</u>. Positive (null, *respectively*) recurrence is a class property, that is whenever $i \leftrightarrow j$,

(a) i is positive recurrent $\Leftrightarrow$ j is positive recurrent;

(b) i is null recurrent $\Leftrightarrow$ j is null recurrent.

<u>Proof</u>.

Throughout the proof we suppose that j is recurrent ($\Leftrightarrow$ i is recurrent)

Since $i \leftrightarrow j$, there are fixed M and N such that $p_{ij}(M)>0$, $p_{ji}(N)>0$.

Then by CK equations, for any $n \geq 0$

$$p_{jj}(n+N+M) \geq p_{ji}(N)p_{ii}(n)p_{ij}(M)$$

$$\Rightarrow \quad 0 \leq p_{ii}(n) \leq \frac{1}{p_{ji}(N)} \frac{1}{p_{ij}(M)} p_{jj}(n+N+M)$$

Now, if j is null recurrent $\mu_{jj}=\infty$, so by Theorem $p_{jj}(k) \to 0$

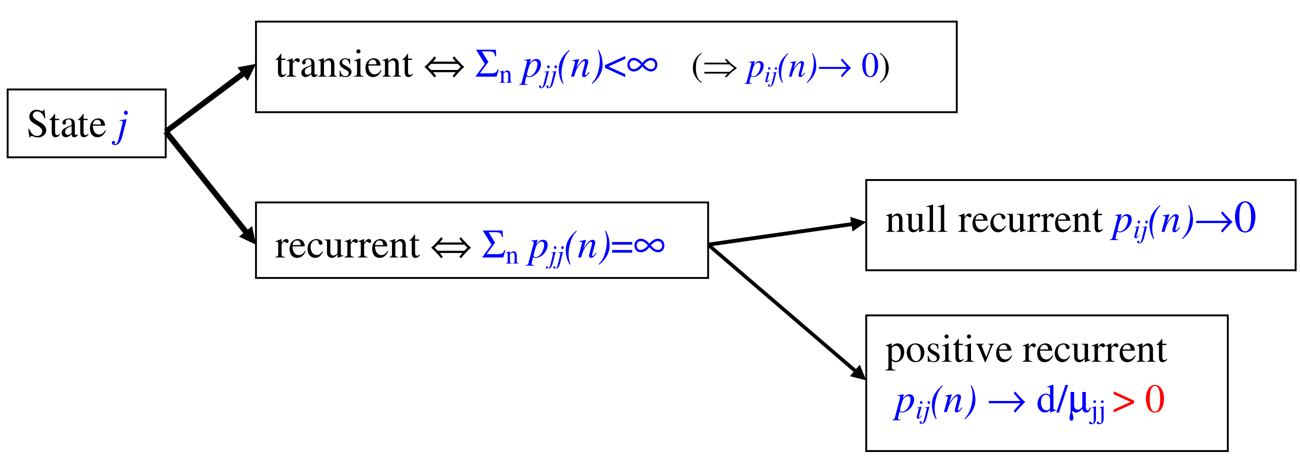
$\Rightarrow$ the RHS of the above $\to 0$ $\quad \Rightarrow \quad$ $p_{ii}(n) \to 0$

$\Rightarrow$ by the same Theorem $\mu_{ii}=\infty$ $\quad \Rightarrow \quad$ i is *also* null recurrent.

Similarly, if i is null recurrent then j is null recurrent. Thus (b) is proved and hence so is (a).

Q.E.D.

Suppose $i \leftrightarrow j$



Stationary (Equilibrium) Distribution

Markov Chain $[X_n: n = 0, 1, 2, \ldots]$, on a countable state space $S=\{0,1,2,3,\ldots\}$ is again assumed irreducible.

Definition.
Let $\{\alpha_j: j=0,1,2,\ldots\}$ be a probability distribution over S, that is

$$\begin{cases} \alpha_j \geq 0 & \forall j \in S \\ \sum_{j\in S} \alpha_j = 1 \end{cases}$$

Then $\{\alpha_j\}$ is a *stationary probability distribution* for the chain X_n if

$$\alpha_j = \sum_{i\in S} \alpha_i p_{ij} \quad \forall j \in S \quad \Rightarrow \quad \underline{\alpha} = \underline{\alpha} P \quad \Rightarrow \quad \underline{\alpha} = \underline{\alpha} P^n$$

row vector

Suppose

$\{\alpha_j\}$ is a stationary probability distribution, and
$P(X_0=j)=\alpha_j$ for all $j\in S$.

Then

$$P(X_1 = j) = \sum_{i\in S} P(X_0 = i) p_{ij} = \sum_{i\in S} \alpha_i p_{ij} = \alpha_j \qquad \forall j \in S$$
$$P(X_2 = j) = \sum_{i\in S} P(X_1 = i) p_{ij} = \sum_{i\in S} \alpha_i p_{ij} = \alpha_j \qquad \forall j \in S$$
$$\vdots$$

i.e. $P(X_n=j)=\alpha_j$ for all $j\in S$ and all n.

Remark. Let $[X_n]$ be an irreducible chain with state space S.
Suppose that either

all states are transient

or

all states are null recurrent.

Then can show that $p_{ij}(n) \to 0$ as $n \to \infty$ for all $i, j \in S$

$\Rightarrow$ ***no*** stationary probability distribution can exist.

Definition. Suppose a chain $[X_n]$ is

(i)	irreducible	(one communicating class)
(ii)	aperiodic	($d=1$)
(iii)	positive recurrent	($\mu_{jj} < \infty$)

Then it is said to be *ergodic*.

Theorem. Let $[X_n]$ be an ergodic chain with state space S.
Then

$$\lim_{n\to\infty} p_{ij}(n) = \frac{1}{\mu_{jj}} \equiv \pi_j \qquad \forall i, j \in S$$

and $\{\pi_j\}$ is unique stationary prob. distribution for the chain.

Proof. Since all states are positive recurrent we already know that

$$\lim_{n\to\infty} p_{ij}(n) = \frac{1}{\mu_{jj}} \qquad \forall i, j \in S$$

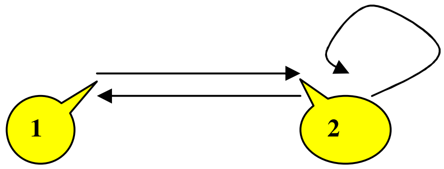
Set $\pi_j = 1/\mu_{jj}$ We must show

(a)

$$\begin{cases} \pi_j = \sum_{i\in S} \pi_i p_{ij} \quad \forall j \in S \\ \sum_{j\in S} \pi_j = 1 \end{cases}$$

(b) $\{\pi_j\}$ is the unique stationary probability distribution.

Example. Two states S = {1,2}



$$P = \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix}$$

(A) Look for the stationary distribution $\underline{\pi}\, P = \underline{\pi}$

$$(\pi_1 \quad \pi_2)\begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} = (\pi_1 \quad \pi_2) \;\Rightarrow\; 0\pi_1 + \frac{1}{2}\pi_2 = \pi_1 \;\Rightarrow\; \pi_2 = 2\pi_1$$

Since $\pi_1+\pi_2=1$, we have $\underline{\pi} = (1/3,\ 2/3)$.

(B) Can we show directly that

$$\left.\begin{aligned} p_{i1}(n) &\to {}^{1}\!/_{3} \quad as \quad n \to \infty \quad for \quad i = 1,2 \\ p_{i2}(n) &\to {}^{2}\!/_{3} \quad as \quad n \to \infty \quad for \quad i = 1,2 \end{aligned}\right\} \quad ?$$

Yes.

Let $Q = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}$, then $Q^{-1} = Q$

and easy to see that $QPQ = \begin{bmatrix} 1 & 0 \\ 0 & -1/2 \end{bmatrix}$, hence $(QPQ)^n = \begin{bmatrix} 1^n & 0 \\ 0 & (-1/2)^n \end{bmatrix}$

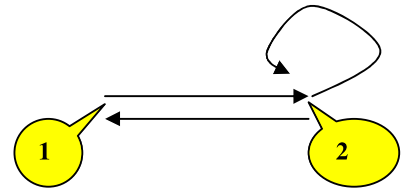
But

$$(QPQ)^n = QPQQPQ....QPQ = QP(QQ)P...(QQ)PQ = QP^nQ$$

$$\Rightarrow \quad P^n = Q^{-1}\begin{bmatrix} 1^n & 0 \\ 0 & (-1/2)^n \end{bmatrix}Q^{-1} \xrightarrow[n\to\infty]{} Q\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}Q = \begin{bmatrix} 1/3 & 2/3 \\ 1/3 & 2/3 \end{bmatrix}$$

(C) Does this agree with the mean recurrence time?

Yes.



State {1}:

Let T be time to return to {1}.

Then:

$$For \qquad k \geq 2, \qquad \mathrm{P}(T = k) = 1 \times \left(\frac{1}{2}\right)^{k-2} \times \frac{1}{2} = \left(\frac{1}{2}\right)^{k-1}$$

Thus

$$\mathrm{E}(T) = \sum_{k=0}^{\infty} k\mathrm{P}(T = k) = \sum_{k=2}^{\infty} k\left(\frac{1}{2}\right)^{k-1} = 3$$

Why is that 3?

$$\begin{aligned}\sum_{k=2}^{\infty} kx^{k-1} &= -1x^0 + \sum_{k=0}^{\infty} kx^{k-1} = -1 + \sum_{k=0}^{\infty} \frac{d}{dx}\left(x^k\right) \\ &= -1 + \sum_{k=0}^{\infty} \frac{d}{dx}\left(x^k\right) = -1 + \frac{d}{dx}\left[\sum_{k=0}^{\infty} x^k\right] \\ &= -1 + \frac{d}{dx}\left[\frac{1}{1-x}\right] = -1 + \frac{1}{(1-x)^2}\end{aligned}$$

Now plug in $x = ½$.

Similarly, can show the mean return time for state {2} is 3/2.

Example. M/M/1 queue

Inter arrival times $\exp(\lambda)$

Service times $\exp(\mu)$

$\boxed{\lambda < \mu}$

X_n = queue size after n^{th} change of queue size

$$P = \begin{array}{c|ccccc} & 0 & 1 & 2 & 3 & \cdots \\ \hline 0 & 0 & 1 & 0 & 0 & \cdots \\ 1 & \frac{\mu}{\mu+\lambda} & 0 & \frac{\lambda}{\mu+\lambda} & 0 & \cdots \\ 2 & 0 & \frac{\mu}{\mu+\lambda} & 0 & \frac{\lambda}{\mu+\lambda} & \cdots \\ 3 & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

Solve: $\underline{\pi}\, P = \underline{\pi} \;\Rightarrow$

$$\left(\frac{\mu}{\lambda+\mu}\right)\pi_1 = \pi_0 \qquad (\aleph 1)$$

$$\pi_0 + \left(\frac{\mu}{\lambda+\mu}\right)\pi_2 = \pi_1 \qquad (\aleph 2)$$

$$\left(\frac{\lambda}{\lambda+\mu}\right)\pi_{k-1} + \left(\frac{\mu}{\lambda+\mu}\right)\pi_{k+1} = \pi_k \quad , \quad k \geq 2 \quad (\aleph 3)$$

To solve, try $\pi_k = \theta^k$, $k \geq 1$

Then for $k \geq 1$, using ($\aleph$3)

$$\left(\frac{\lambda}{\lambda+\mu}\right)\theta^{k-1} + \left(\frac{\mu}{\lambda+\mu}\right)\theta^{k+1} = \theta^k \quad \Rightarrow \quad \mu\theta^2 - (\lambda+\mu)\theta + \lambda = 0$$

$$\Rightarrow \quad (\mu\theta - \lambda)(\theta - 1) = 0 \quad \Rightarrow \quad \theta_1 = 1, \quad \theta_2 = \lambda/\mu$$

Hence, general solution has the form

$$\pi_k = A + B\,(\lambda/\mu)^k \qquad (\aleph 4)$$

Now from ($\aleph$1) and ($\aleph$2),

$$\pi_0 = \left(\frac{\mu}{\lambda+\mu}\right)\pi_1 \qquad (\aleph 5)$$

$$\pi_1 = \left(\frac{\mu}{\lambda+\mu}\right)\pi_1 + \left(\frac{\mu}{\lambda+\mu}\right)\pi_2$$

$$\Rightarrow \quad \left(\frac{\lambda}{\lambda+\mu}\right)\pi_1 = \left(\frac{\mu}{\lambda+\mu}\right)\pi_2$$

$$\Rightarrow \quad \pi_2 = \frac{\lambda}{\mu}\pi_1$$

Then, from ($\aleph$4) for $k = 2$ we have

$$A + B\left(\frac{\lambda}{\mu}\right)^2 = \frac{\lambda}{\mu}\left(A + B\left(\frac{\lambda}{\mu}\right)\right) \quad \Rightarrow \quad A = \frac{\lambda}{\mu}A \quad \Rightarrow \quad A = 0$$

since $\lambda<\mu$. Therefore, $\pi_k = B\,(\lambda/\mu)^k$ for $k \ge 1$

From ($\aleph$5), we obtain

$$\pi_0 = \left(\frac{\mu}{\lambda+\mu}\right)\times B\left(\frac{\lambda}{\mu}\right)^1 = \left(\frac{\lambda}{\lambda+\mu}\right)B$$

Require

$$\sum_{k=0}^{\infty}\pi_k = 1$$

$$\Rightarrow \quad \left(\frac{\lambda}{\lambda+\mu}\right)B + \sum_{k=1}^{\infty} B\left(\frac{\lambda}{\mu}\right)^k = 1$$

$$\Rightarrow \quad B\left[\left(\frac{\lambda}{\lambda+\mu}\right) + \left(\frac{\lambda}{\mu}\right)\frac{1}{1-(\lambda/\mu)}\right] = 1$$

$$\Rightarrow \quad B\left[\frac{\lambda}{\lambda+\mu} + \frac{\lambda}{\mu-\lambda}\right] = 1 \qquad \Rightarrow \quad \frac{2B\mu\lambda}{\mu^2-\lambda^2} = 1$$

$$\Rightarrow \quad B = \frac{\mu^2-\lambda^2}{2\mu\lambda}$$

So

$$\pi_0 = \frac{\mu-\lambda}{2\mu}, \qquad \pi_k = \frac{\mu^2-\lambda^2}{2\mu\lambda}\left(\frac{\lambda}{\mu}\right)^k, \quad k \ge 1$$

Finally, compute the proportion of *real time* the queue is empty:

$$\frac{\pi_0(\lambda)^{-1}}{\pi_0(\lambda)^{-1} + \sum_{k=1}^{\infty}\pi_k(\mu+\lambda)^{-1}} = \frac{\pi_0}{\pi_0 + \sum_{k=1}^{\infty}\left(\frac{\lambda}{\mu+\lambda}\right)\frac{\mu^2-\lambda^2}{2\mu\lambda}\left(\frac{\lambda}{\mu}\right)^k}$$

$$= \frac{\pi_0}{\pi_0 + \sum_{k=1}^{\infty}\left(\frac{\mu-\lambda}{2\mu}\right)\left(\frac{\lambda}{\mu}\right)^k} = \frac{1}{1+\sum_{k=1}^{\infty}(\lambda/\mu)^k}$$

$$= \left(\sum_{k=0}^{\infty}(\lambda/\mu)^k\right)^{-1} = 1-\frac{\lambda}{\mu}$$

Chapter 6: total 8 lectures, 3 ¾ slides per lecture