BIOMATH 203 - Stochastic Models in Biology

No textbook - many available

Combination of formal definitions and models of physical/biological systems.

~2 homework + midterm

~ final project or final exam

Review probability and definitions

Sets a collection of distinct objects

an "experiment" (an event such as rolling a die, possibly more than once) is described by

(1) Outcome space — set of all possible outcomes

"Ω" $= \{1, 2, 3, 4, 5, 6\}$ for the dice problem

(2) Events $\mathcal{F}$ are subsets of Ω ($\mathcal{F}$: event space)

for rolling a die $\Omega = \{1, 2, 3, 4, 5, 6\}$,

event E_1: roll an even number $E_1 = \{2, 4, 6\}$

event E_2: roll a number > 4, $E_2 = \{5, 6\}$

event E_3: roll a one, $E_3 = \{1\}$ (typically, $\mathcal{F}$ is a σ-algeb

(3) Probability of an event is defined

Probability of event E is the number of outcomes in E divided by total number of outcomes

$$P(E) = \frac{\#(E)}{\#(\Omega)}$$

Microcanonical Ensemble

all replica's have same energy

for the die, $\Omega = \{1, 2, 3, 4, 5, 6\}$, $\#(\Omega) = 6$

* $\#(E_1) = 3$, $P(E_1) = \frac{1}{2}$
* $\#(E_2) = 2$, $P(E_2) = 1/3$
* $\#(E_3) = 1$, $P(E_3) = 1/6$

$0 \leq P(E) \leq 1$, $\underbrace{P(\Omega) = 1}_{\text{prob. of anything happening}}$, $\overbrace{P(\phi) = 0}^{\text{Something has to have happened by defn.}}$

Rolling die twice, or two dice (independent)

$$\Omega = \{ (1,1), (1,2), (1,3) \ldots (1,6), \\ (2,1), (2,2), (2,3) \ldots (2,6), \\ \vdots \\ (6,1), (6,2), \ldots (6,6) \}$$

$\Omega = \{(i,j) \mid i,j \text{ between 1 and 6}\}$

$\#(\Omega) = 36$

In this scenario, consider <u>event A_1</u>:
The sum of the two numbers $= 5$;

$$A_1 = \{(1,4), (2,3), (3,2), (4,1)\} \subset \Omega, \quad \#(A_1) = 4$$

$$P(A_1) = \frac{4}{36} = \frac{1}{8}$$

<u>Event A_2</u>: second number greater than first $j > i$

$$\begin{aligned} A_2 = \{ & (1,2), (1,3), (1,4) \cdots (1,6), \\ & (2,3), (2,4) \cdots (2,6), \\ & (3,4), (3,5), (3,6), \\ & (4,5), (4,6), \\ & (5,6) \} \end{aligned}$$

$$\#(A_2) = 15$$

$$P(A_2) = \frac{15}{36}$$

<u>Event A_3</u>: at least one number is a 1,

$$\begin{aligned} A_3 = \{ & (1,1), (1,2), \ldots (1,6), \\ & (2,1), \\ & (3,1), \\ & \vdots \\ & (6,1) \} \end{aligned}$$

$$\#(A_3) = 11$$

$$P(A_3) = \frac{11}{36}$$

* Our construction of probabilities assumes "equally likely outcomes" i.e. a fair dice. $\boxed{\text{"microcanonical ensemble"}}$

if probabilities (elements of Ω) are not equal, they have to be <u>weighted</u>.

e.g. Boltzmann weighting according to the "energy" of an event

---

<u>Frequency $\Leftrightarrow$ probability</u>:

Repeat experiment, each repetition is a trial.

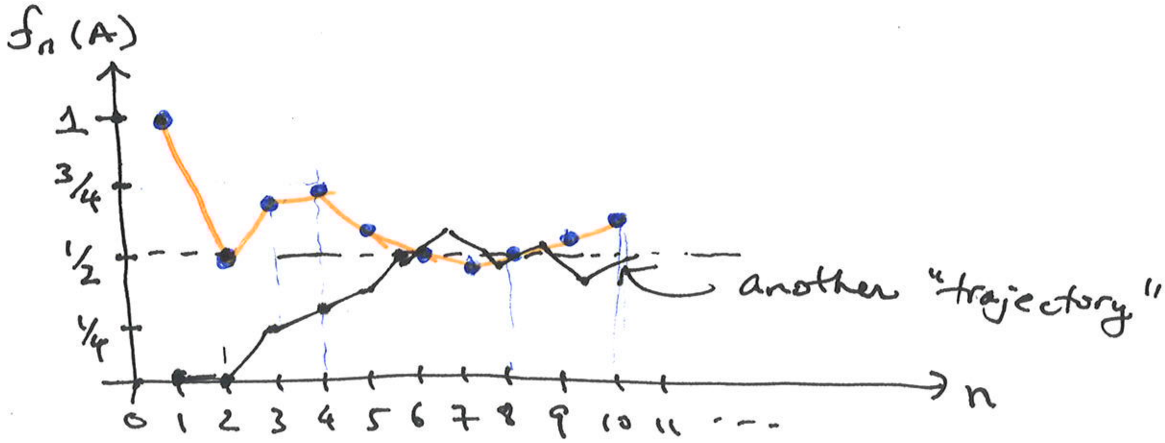
If trials are independent, $P(A)$ can be estimated by the proportion of times A occurs.

Frequency $f_n(A) \equiv \dfrac{n(A)}{n}$ n is the number of trials

<u>example</u>

coin toss. suppose $A = \{H\}$, $\Omega = \{H, T\}$

$n =$	1	2	3	4	5	6	7	8	9	10
outcome $=$	H	T	H	H	T	T	T	H	H	H
$n(H) =$	1	1	2	3	3	3	3	4	5	6
$f_n(A) =$	$\frac{1}{1}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{3}{4}$	$\frac{3}{5}$	$\frac{3}{6}$	$\frac{3}{7}$	$\frac{4}{8}$	$\frac{5}{9}$	$\frac{6}{10}$



$f_n(A)$ is not unique, depends on realization

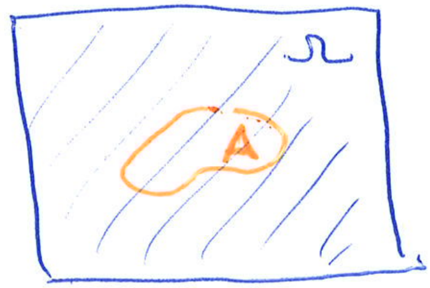
$P(A) = \lim_{n \to \infty} f_n(A)$, but $f_n(A)$ depends on specific sequence, nonetheless limit is well-defined as $n \to \infty$

frequency interpretation formally requires repeated measures, however a single sample can still have a probability.

Probability

...

Events are subsets of Ω,
labelled as $A, B, C, \ldots$



...

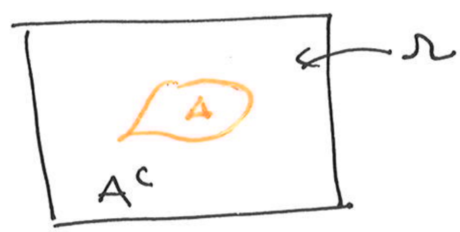
probability is a
function on the set of events

$P \longrightarrow [0, 1]$

if $\Omega = \{1, 2, 3, 4, 5, 6\}$, and $A = \{1, 6\}$, $P(A) = \frac{1}{3}$

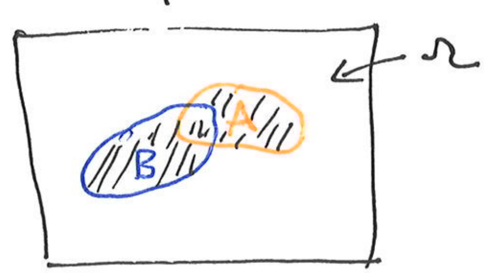
probability measures the "mass" of an event

- For an event A, complement of A, A^c is the set of all outcomes in Ω that does not belong to A



- Union of two events $A \cup B$ is the set of outcomes

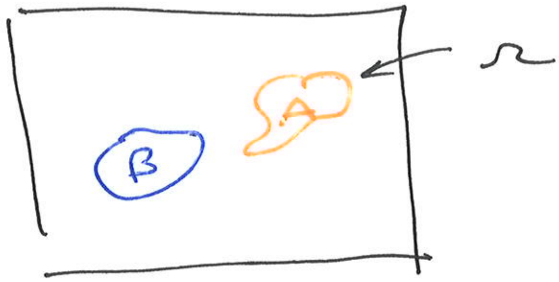
Either A or B happens



- Intersection of two events A and B

$A \cap B$ means ~~either~~ both A and B happens

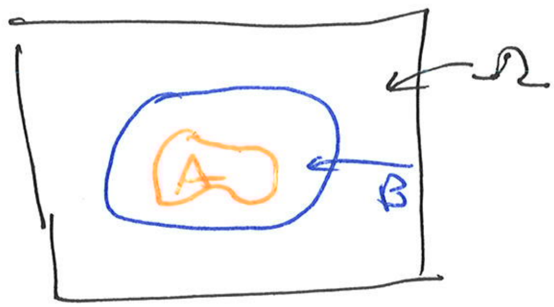
- Disjoint sets



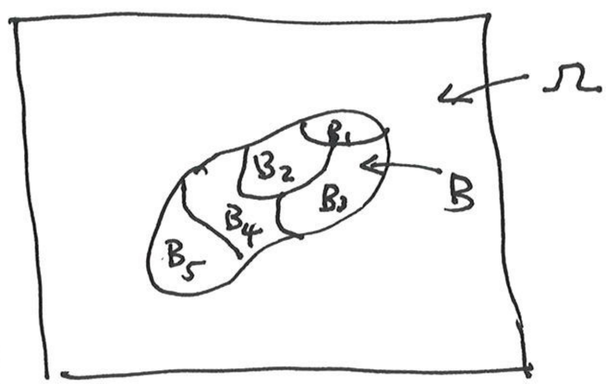
A and B are mutually exclusive, they cannot happen at the same time

- $A \subseteq B$, A is a subset of B

If A happens, B must also happen



Partitions



$B = B_1 \cup B_2 \cup B_3 \dots$

B_i are mutually exclusive

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

$B_1 = \{3, 6\}$, $B_2 = \{4\}$, $B_3 = \{5\}$

* for any $B \subseteq \Omega$, $P(B) \geq 0$

* for any $B \subseteq \Omega$, and partitions B_i, $P(B) = P(B_1) + P(B_2) + \dots$

* $P(\Omega) = 1$ * If $A \subseteq B$, $P(A) \leq P(B)$

Conditional probabilities

probability that A happens, given that you know another (B) happened.

Ex toss coin 3 times

$\{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$ $\#(\Omega) = 8$ (2^3)

Suppose event $A \equiv$ 'at least 2 Hs.

$A = \{HHH, HHT, HTH, THH\}$ $\#(A) = 4$

$$P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{1}{2}$$

event $B \equiv$ first toss lands H:

$B = \{HHH, HHT, HTH, HTT\}$ $\#(B) = 4$

Conditional probability of A given B

$$P(A|B) = \frac{3}{4} = \frac{\#(A \cap B)}{\#(B)}$$

$$= \frac{\frac{\#(A \cap B)}{\#(\Omega)}}{\frac{\#(B)}{\#(\Omega)}} = \frac{P(A \cap B)}{P(B)}$$

Conditional probabilities

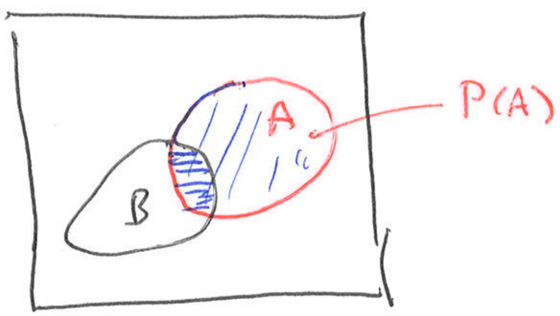
Suppose A & B are two events in sample space

The conditional probability of event B given event A has occurred is $P(B|A)$

$$= \frac{P(A \cap B)}{P(A)}$$

or

$$\boxed{\begin{aligned} P(A\cap B) &= P(B|A)\,P(A) \\ &= P(A|B)\,P(B) \end{aligned}}$$



Events are independent if $P(B|A) = P(B)$

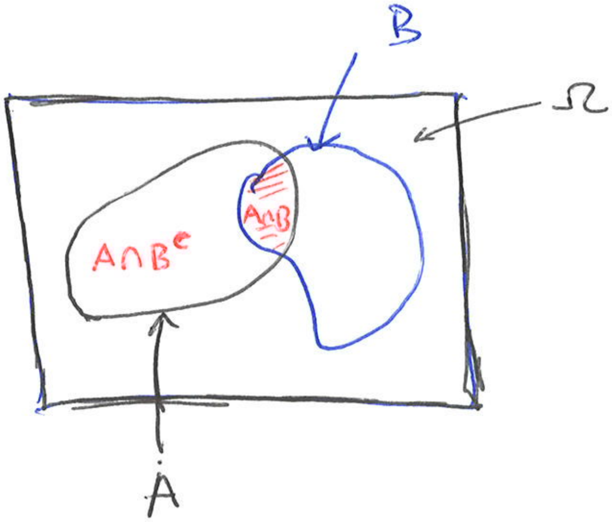
(probability doesn't depend on A)

or $P(A|B) = P(A)$

$$\boxed{P(A\cap B) = P(A)\,P(B)}$$

disjoint A and B

Average conditional prob



$$P(A) = P(A \cap B^c) + P(A \cap B)$$
$$= P(AB^c) + P(AB)$$

if $P(\overbrace{A \cap B}^{AB}) = P(A|B)P(B)$ and $P(\overbrace{A \cap B^c}^{AB^c}) = P(A|B^c)P(B^c)$,

then $\boxed{P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)}$

Dependence can depend on setting / protocol :

ex drawing two cards from a deck

A = first card red

B = second card red

ecology example

With replacement $P(B|A) = 1/2, \; P(B) = 1/2$ independent

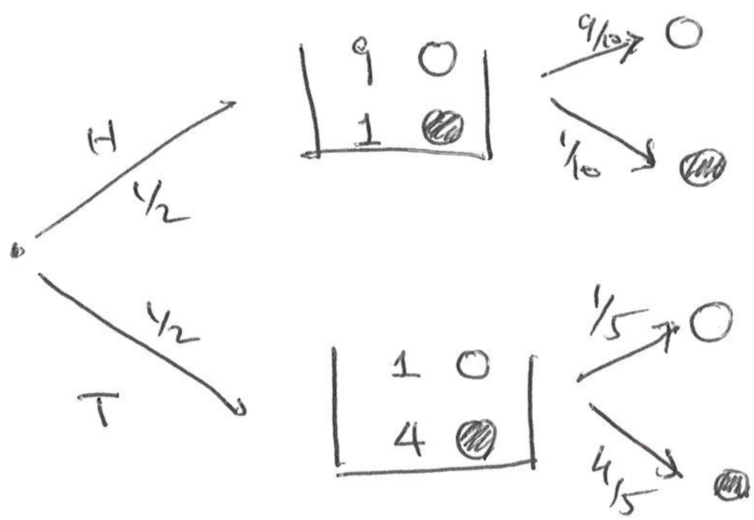
without replacement $P(B|A) = \frac{25}{51}$ not independent

$$P(B) = P(B|A)P(A) + P(B|A^c)P(A^c) = \frac{25}{51}\cdot\frac{1}{2} + \frac{26}{51}\cdot\frac{1}{2} = \frac{1}{2}$$

$\boxed{\text{Compound process}}$ Sequence of events

- flip coin H or T
- if H, then draw ball from one bin
  if T, then draw ball from a different bin

the bins have different numbers of black & white balls



$P(W|H) = 9/10$

$$P(W) = P(W|H)\,P(H) + P(W|\underset{H^c}{T})\,P(T)$$

$$= \frac{1}{2}\frac{9}{10} + \frac{1}{2}\frac{1}{5} = \frac{11}{20}$$

Now, suppose we picked a white ball (given) (W is given), what is the probability we chose H?

i.e. what is $P(H|W)$?

$$P(W|H) = \frac{P(WH)}{P(H)}$$ ← white and heads

$$P(H|W) = \frac{P(HW)}{P(W)}$$

equivalent so $P(H|W)\,P(W) = P(W|H)\,P(H)$

$$\Rightarrow \boxed{P(H|W) = \frac{P(W|H)\,P(H)}{P(W)}}$$

Bayes Rule

$\boxed{\text{Baye's Rule}}$

$$P(H|W) = \frac{9/10 \cdot 1/2}{11/20} = \frac{9}{11}$$

In general

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$$

Classic FPR, TPR diagnosis:

estimate of ground truth disease D prevalence $\overbrace{P(D)=0.01}^{P(D^c)=0.99}$

Test parameters $P(+|D) = 0.95$ probability of true positive

$P(+|D^c) = 0.02$

not diseases

all these numbers must be estimated from exper.

Pick a random individual to test, result is +

given result is +, what is probability of disease?

$P(D|+)$?

$$P(D|+) = \frac{P(+|D)P(D)}{P(+)} = \frac{(0.95)(0.01)}{(0.95)(0.01)+(0.02)(0.99)}$$

$$= 0.32$$

random testing of a rare disease amplifies the FPR. For $P(D)=0.3$, $P(D|+) = 0.953$

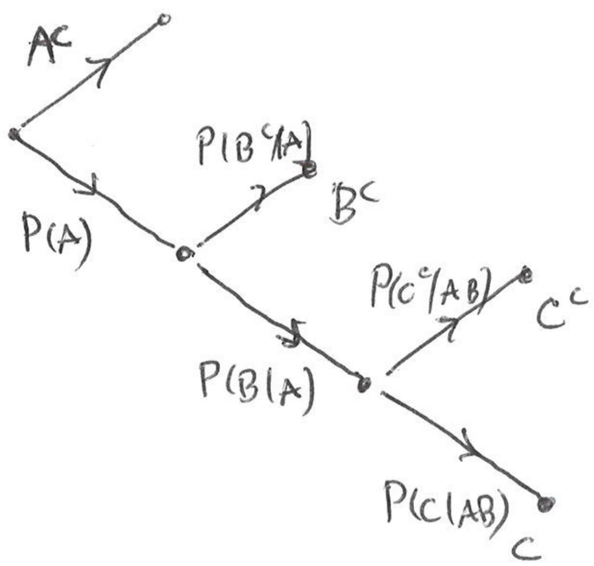
Sequences of events

(closer to "stochastic process")

A, B, C are sequential events

ABC, event A and B and C occurred

$$P(ABC) = P(C|AB)\underbrace{P(AB)}$$
$$= P(C|AB)\;P(B|A)P(A)$$



Probability distributions

* for discrete state space, consider histograms of

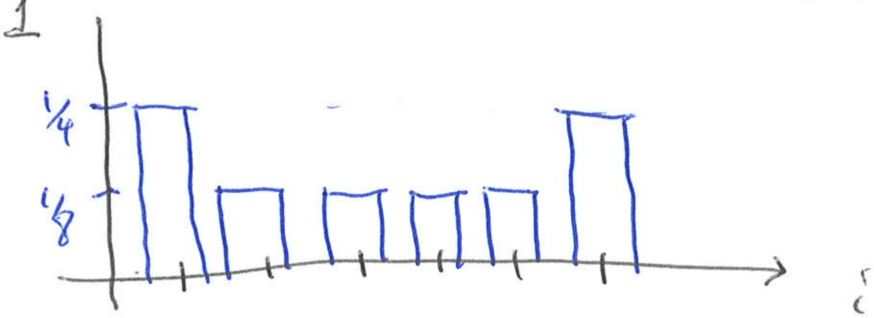
$P(\{i\}),\ i \in \Omega$

- bar diagram with outcomes on x-axis
- height above each even q is the probability of q $P(q)$
- Normalized

Ex biased die, 1 and 6 mor likely

Suppose $P(1) = P(6) = \frac{1}{4}$ and $P(\{i\}) = \frac{1}{8}$ for $i = 2, 3, 4, 5$

$$\sum_{j=1}^{\#(\Omega)=6} P(j) = 1$$



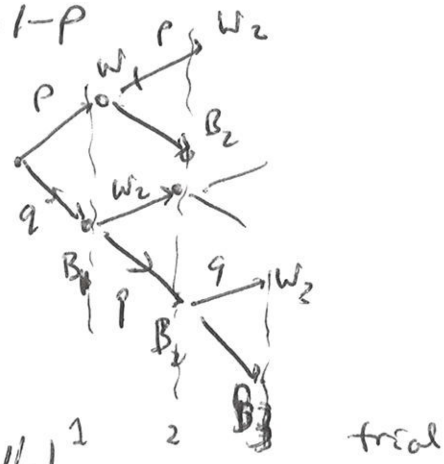
$\boxed{\text{probability distribution}}$

Geometric distribution

probabilities:

Binary tree: binary outcome, $p =$ white

$q =$ black $= 1-p$



independent trials

P(it takes 3 trials or fewer to get white)

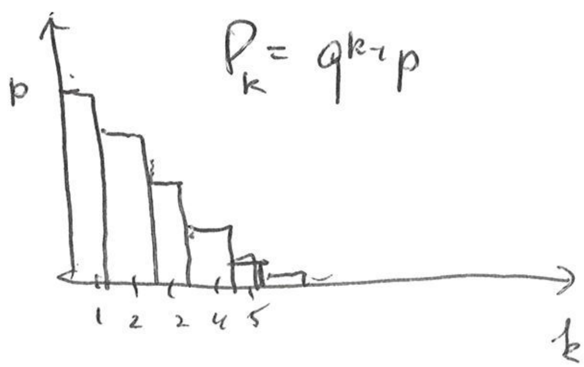
$$= \underbrace{P(W_1)}_{1^{st}\text{ attempt}} + \underbrace{P(B_1)P(W_2)}_{2^{nd}\text{ attempt}} + \underbrace{P(B_1)P(B_2)P(W_3)}_{3^{rd}\text{ attempt}}$$

$$p + qp + q^2 p = (1+q+q^2)p$$

P(getting white for the first time on the k^{th} trial)

$= q^{k-1} p$ $\quad$ $(k-1)$ trials failed, then success.

$$P_k = q^{k-1} p$$



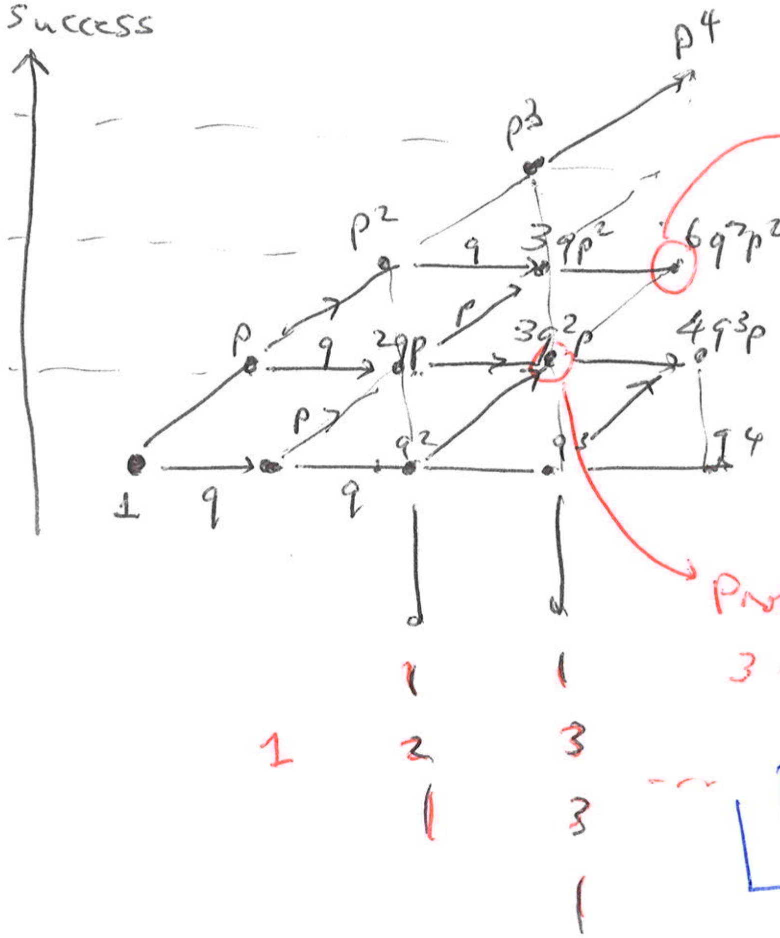
(discrete version of exponentially distributed waiting time in a Markov process)

$$\approx e^{(k-1)\ln((1-p)p)}$$

Binomial Distribution

n trials, how many success?

(each trial has success probability p)



probability of two successes out of 4 trials

Prob of 1 success out of 3 trials

Pascal's triangle

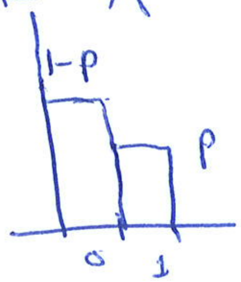
$$\equiv \binom{n}{k} = \frac{n!}{(n-k)!\,k!}$$

Probability of k successes out on n trials:

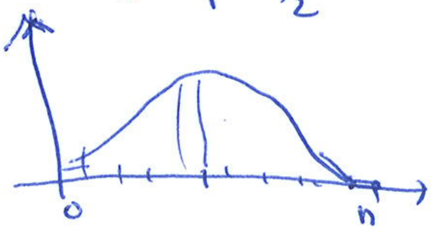
$$P(k|n) = \binom{n}{k} p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

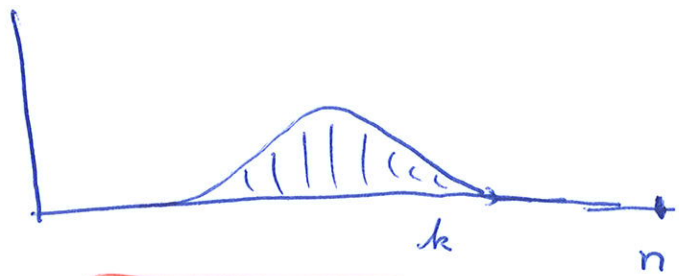
Binomial distribution $B(n,p)$

$B(1,p)$ = Bernoulli distribution Bernoulli(p)

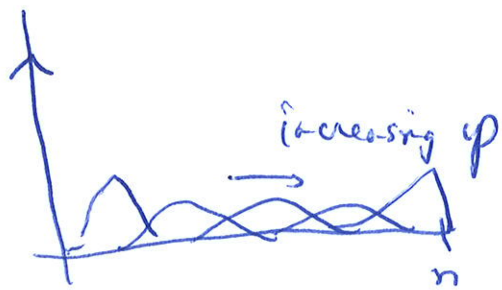


for $p = 1/2$





larger



$$\boxed{\begin{aligned} \mathbb{E}[k] &= np \\ Var[k] &= np(1-p) \end{aligned}}$$

* approaches Gaussian for large n and $p \sim 1/2$

<u>Poisson distribution</u> approximation for Binomial for small p (rare events)

if $p = \lambda/n$ we expect $\sim \lambda$ events ($n \to \infty$, $p \to 0$ s.t. $np = \lambda$)

success :

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

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

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \approx \frac{n^k}{k!}$$

$$p(k) \approx \frac{n^k}{k!} \left(\frac{\lambda}{n}\right)^k e^{-\lambda} \approx \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{(Erlang)}$$

$$\boxed{\mathbb{E}[k] = Var[k] = np}$$

$\boxed{\text{Negative Binomial}}$

of trials needed to record k successes

$P(n \geq k)$

n^{th} trial must be a success with probability p

the first $n-1$ trials must contain $k-1$ successes and $n-k$ failures

$\binom{n-1}{k-1}$ ways of arranging $k-1$ successes in $n-1$ trials

probability is $p^{k-1} q^{n-k}$

$$P(n \text{ trials to get } k \text{ successes}) = \binom{n-1}{k-1} p^{k-1} q^{n-k} \cdot p$$

$$= \binom{n-1}{k-1} p^{k} q^{n-k}$$

$$\boxed{E[n] = \frac{k}{p}, \ Var[n] = \frac{kq}{p^2}}$$

Continuous probability densities

$f(x) \equiv P(X = x)$ is a density if x, X is continuous

x can be time, position, etc

$f(x)\,dx$ is the probability that X has value in $[x, x+dx)$

Uniform

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

$$\mathbb{E}[X] = \frac{1}{2}\frac{(b^2-a^2)}{b-a} = \frac{a+b}{2}$$

$$\mathrm{Var}[X] = \frac{1}{12}(b-a)^2$$

Exponential

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & x < 0 \end{cases} \quad (\lambda)$$

$$\mathbb{E}[X] = \lambda^{-1}$$

$$\mathrm{Var}[X] = \lambda^{-2}$$

Gamma

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\,\beta^\alpha} x^{\alpha-1} e^{-x/\beta} & x \ge 0 \\ 0 & x < 0 \end{cases} \quad (\alpha, \beta)$$

$$\mathbb{E}[X] = \alpha\beta$$

$$\mathrm{Var}[X] = \alpha\beta^2$$

Gaussian

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (\mu, \sigma) \qquad -\infty < x < \infty$$

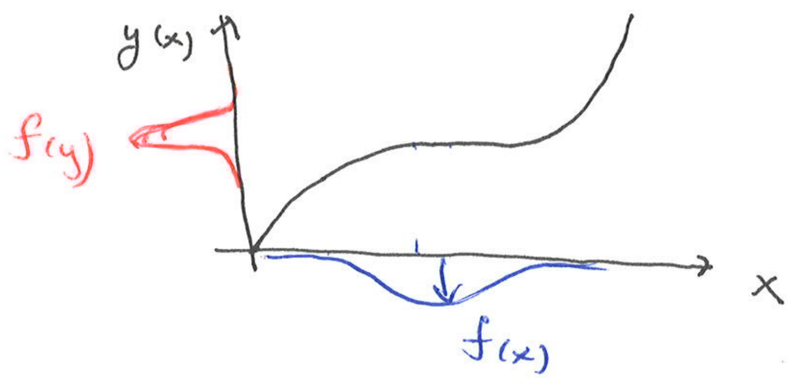
$$\mathbb{E}[X] = \mu$$

$$\mathrm{Var}[X] = \sigma^2$$

Tranformations

$$\int_{x_{min}}^{x_{max}} f(x)\, dx = 1 \quad \text{if } f(x) \text{ is a probability density}$$

Suppose theres a function on x $y = g(x)$, what is $f(y)$?



In 1D, $f_Y(y) = f_X(g^{-1}(y)) \left| \dfrac{d g^{-1}(y)}{dy} \right| = \dfrac{f_X(g^{-1}(y))}{\left|\frac{dy}{dx}\right|}$

monotonic $g(x)$,

For nonmonotonic $g(x)$, must consider each monotonic branch

Suppose $g(x) = x^2$, $y > 0$ only, but there are two branches

$X = \sqrt{y}$, $X = -\sqrt{y}$

$$f_Y(y) = f_X(\sqrt{y}) \left| \frac{1}{2\sqrt{y}} \right| + f_X(-\sqrt{y}) \left| \frac{1}{2\sqrt{y}} \right|$$

for Gaussian $f_X(x)$, ($\mu = 0$)

$$\boxed{f_Y(y) = \frac{e^{-y/2\sigma^2}}{\sqrt{2\pi y}\ \sigma}, \quad \int_0^{\infty} f_Y(y)\, dy = 1}$$

<u>Transformations</u>

$$f_{X,Y}(x,y) = f_X(x)\, f_Y(y) \quad \text{and} \quad \underline{Z = X+Y}$$

$$f_Z(z) = \iint f_X(x) f_Y(y)\, \delta(z-(x+y))\, dx\, dy$$

δ-function enforces $y = z-x$

$$\boxed{f_Z(z) = \int dx\, f_X(x) f_Y(z-x)\, dx \quad \text{convolution!}}$$

$\underline{Z = XY}$ $\quad f_Z(z) = \iint f_{X,Y}(x,y)\, \delta(z-xy)\, dx\, dy$

coordinate transformation of δ-function $\delta(z-xy) \to \dfrac{\delta(y-\frac{z}{x})}{|x|}$

$$\boxed{f_Z(z) = \iint f_{X,Y}\left(x, \frac{z}{x}\right) \frac{dx}{|x|}}$$

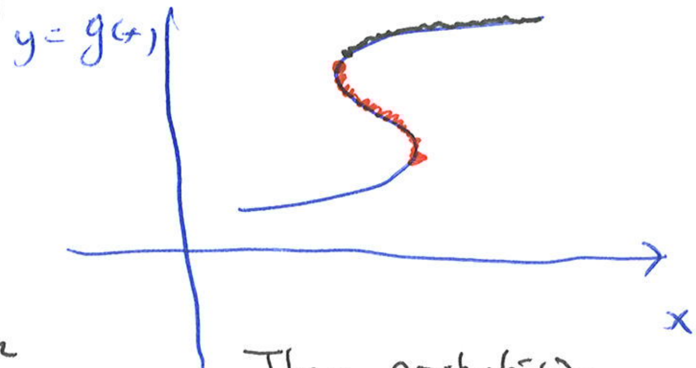
Transformations

$g(x) = x^3 \qquad f_Y(y) = f_X(y^{1/3})\left|\frac{dy^{1/3}}{dy}\right| \qquad \frac{f_X(y^{1/3})}{3y^{2/3}}$

for normal $f_X(x)$,

$$\boxed{f_Y(y) = \frac{2\, e^{-|y|^{2/3}/2\sigma^2}}{3\sqrt{2\pi}\,\sigma\,|y|^{2/3}}}$$

what if $g(x)$ is not single-valued?



Then probability on the branches need to be considered, esp. if x-dynamics does not reach unstable branch...

homework? $y = \sin x$

$y = \frac{1}{1+e^{-x}}$?

Multivariate transformations

let $X = (x_1, x_2, x_3 \ldots x_n)$ be a random vector and

$f_X(x)$ be the joint prob. density.

Consider the transformation

$Y = g(X)$ invertible

$\frac{\partial x_i}{\partial y_j}$

$f_Y(y) = f_X(g^{-1}(y))\,|\det J_{ij}|$

Suppose $f_{R,\Theta}(r,\theta) \longrightarrow$ transform to x, y

Transformations

$x = r\cos\theta, \quad y = r\sin\theta$

$r = \sqrt{x^2+y^2}\,, \quad \theta = \tan^{-1}(y/x)$

$$J = \begin{vmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{vmatrix} = \begin{pmatrix} x/r & y/r \\ -y/r^2 & \frac{x}{r^2} \end{pmatrix} = \frac{1}{r(x,y)}$$

$$f_{X,Y}(x,y) = f_{R,\Theta}\left(\sqrt{x^2+y^2}, \tan^{-1}(y/x)\right) \frac{1}{\sqrt{x^2+y^2}}$$

---

* transform to fewer variables (cannot transform to more → meaningless)

$f_{X,Y}(x,y) \rightarrow f_Z(z)$ if $Z = g(x,y)$

<u>Cumulative method</u>

$$F_Z(z) = P(Z \le z) = \iint\limits_{(\{x,y\},\ g(x,y) \le z)} f_{X,Y}(x,y)\,dx\,dy$$

then, that $\frac{dF_Z(z)}{dz} \equiv f_Z(z)$

<u>δ-function method</u>

$$f_Z(z) = \iint f_{X,Y}(x,y)\,\delta(z - g(x,y))\,dx\,dy$$

Generating Functions

Discrete variables assume integer state space $\{0, 1, 2, \dots\}$

$f(j) \equiv$ probability distribution $= P(X=j) \equiv p_j$
(aka "mass function")

mean $\mu_X = \mathbb{E}[X] \equiv \sum_{j=0}^{\infty} j\, p_j$

$$\sigma_X^2 = \mathbb{E}[(X-\mu)^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$
$$= \sum_{j=0}^{\infty} j^2 p_j - \mu_X^2$$

"generating" functions generate (moments / probabilities / useful quantities) of X
and is a Laplace transform:
("z-transform")

$$g(z) \equiv \mathbb{E}[z^X]_{\{p_i\}} = \sum_{j=0}^{\infty} p_j z^j$$

(probabily generating function)

$$g(1) = 1, \quad \frac{1}{n!}\frac{d^n g(z)}{dz^n}\bigg|_{z=0} = p_n$$

Generating Functions

Moment generating functions

$$M(z) \equiv \mathbb{E}[e^{zX}] = \sum_{j=0}^{\infty} p_j e^{zj} \equiv \sum_{j=0}^{\infty} p_j s^j, \quad s \equiv e^z$$

$$\left.\frac{d^n M(z)}{dz^n}\right|_{z=0} = \mathbb{E}[X^n]$$

Characteristic Functions

$$\phi_X(z) = \mathbb{E}[e^{izX}]$$

Fourier counterpart
most useful for proofs

$$\boxed{\text{Cumulative distributions}}$$

$$F_X(x) = P(X \leq x)$$

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

let $b = a + \varepsilon$.

$$\lim_{\varepsilon \to 0} P(a < X < a+\varepsilon) = F_x(a+\varepsilon) - F_x(a)$$

$$= \cancel{F_X(a)} + \varepsilon \left.\frac{dF}{dx}\right|_a + \dots - \cancel{F_X(a)}$$

$$= \underbrace{\varepsilon \left.\frac{dF}{dX}\right|_a}_{f(x) \text{ probability density}} \equiv \text{probability } X \in (a, a+\varepsilon)$$

Now Suppose $Y = g(X)$ is monotonicay increasing

CDF for Y: $F_Y(y) = P(Y \leq y) = P(g(X) \leq y)$

$$= P(X \leq g^{-1}(y))$$

$$\equiv F_X(g^{-1}(y))$$

then, $f_Y(y) = \dfrac{dF_Y(y)}{dy}$

Ex. Linear $Y = aX + b$,

$$F_Y(y) = F_X\left(\frac{y-b}{a}\right), \quad f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{a}\frac{dF_X}{dy}$$

$$= \frac{1}{a} f_X\left(\frac{y-b}{a}\right)$$

Cumulant generating function

$$K_X(z) = \ln M_X(z)$$

Continuous versions

probability gen func. $\int f(x)\, z^x \, dx \equiv \mathbb{E}[z^X]$

Laplace is $x \geq 0$ and $z \equiv e^{-s}$

moment gen. func. $\int f(x) e^{zx} \, dx \equiv \mathbb{E}[e^{zX}]$

Characteristic func. $\int f(x)\, e^{ixz} \, dx = \mathbb{E}[e^{izX}]$

("Fourier")

Example: prob. gen. func. for Binomial

$$\mathbb{E}[z^X] = \sum_{j\geq 0}^{n} \binom{n}{j} p^j (1-p)^{n-j} z^j$$

$$= \sum_{j\geq 0}^{n} \binom{n}{j} (pz)^j (1-p)^{n-j} \equiv [pz + (1-p)]^n$$

PGF "generates" certain averages (in addition to P_i)

$$\frac{d}{dx} \mathbb{E}[z^X]\Big|_{z=1} = \mathbb{E}[X], \quad \frac{d^2}{dz^2} \mathbb{E}[z^X]\Big|_{z=1} = \mathbb{E}[X(X-1)]$$

Now consider probability distributions like P_n evolve in "time"

In other words, the recursion may be related to a rule governing the evolution of P_n for all n.

Markov process

$$\star \quad \mathbb{P}\text{rob}\,[X_k = i_k \mid X_{k-1} = i_{k-1}, X_{k-2} = i_{k-2}, \ldots X_0 = i_0)$$

$i_k \in \{1, 2, 3 \ldots\}$ for $k = 0, 1, 2 \ldots$ (think of k as discrete time steps)

$$\equiv \mathbb{P}\text{rob}\,[X_k = i_k \mid X_{k-1} = i_{k-1}]$$

probability now depends on only previous state

$$p_i(k) = \mathbb{P}\text{rob}\,[X_k = i]$$ probability distribution at time k

Define transition probabilities as a conditional probability over one step:

$$p_{ij}(k) \equiv \mathbb{P}\text{rob}\,[X_{k+1} = i \mid X_k = j]$$

this is the probability that state had value j at time step k given the state acquires value i at step $k+1$

p_{ij} represents $j \to i$ (different conventions)

Transition matrix (Stochastic matrix)

$$P = \begin{pmatrix} P_{11} & P_{12} & P_{13} & P_{14} & \cdots \\ P_{21} & P_{22} & P_{23} & P_{24} & \cdots \\ P_{31} & P_{32} & P_{33} & & \\ \vdots & & & & \end{pmatrix}$$

In this convention columns sum to 1

In principle P can also depend on k (time-inhomogeneous)

P_{11} prob 1 stays 1

P_{12} prob $1 \rightarrow 2$

P_{13} prob $1 \rightarrow 3$

+

total prob of all steps $= 1$

In this notation,

$$\boxed{p_i(k+1) = \sum_j P_{ij}\, p_j(k)}$$

Discrete time Markov process

$\downarrow$

$$\vec{p}\,(k+1) = P\,\vec{p}(k)$$

$\vec{p}(k)$ is a right eigenvector

Sometimes use $\pi_i(k)$

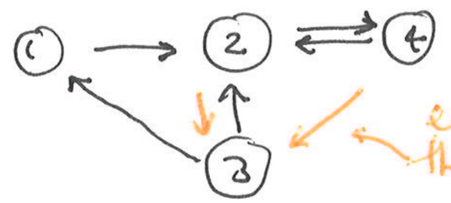
$$\pi_i(k+1) = \sum_j P_{ij}\, \pi_j(k)$$

Example

$1 \to 2$ $\qquad$ $3 \to 1,\ 3 \to 2$

$$P = \begin{pmatrix} 0 & 0 & p_{13} & 0 \\ \boxed{p_{21}} & 0 & p_{23} & p_{24} \\ 0 & 0 & 0 & 0 \\ 0 & p_{42} & 0 & 0 \end{pmatrix} \qquad 2 \leftrightarrow 4$$

graph



either of these will make matrix irreducible

① and ③ cannot be returned to once left

$\{1\}$, $\{3\}$, and $\{2, 4\}$ are the three communication classes

---

Consider Gambler's Ruin again

State space $\{0, 1, 2, \dots N\}$ stochastic matrix is $(N+1) \times (N+1)$

$$p_{ji} \qquad \begin{pmatrix} 1 & \overset{1\to 0}{q_1} & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & \overset{(2\to 1)}{q_2} & 0 & 0 & 0 & \cdots \\ 0 & \overset{2\to 1}{p_2} & 0 & \overset{3\to 2}{q_3} & 0 & 0 & \cdots \\ 0 & 0 & \overset{3\to 2}{p_3} & 0 & \overset{4\to 3}{q_4} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$

directed graph for gambler's ruin

$$
\begin{tikzcd}
0 \arrow[loop left, "1"] & 1 \arrow[l, "q_1"] \arrow[r, shift left, "p_1"] & 2 \arrow[l, shift left, "q_2"] \arrow[r, shift left, "p_2"] & 3 \arrow[l, shift left, "q_3"] \arrow[r, shift left] & 4 \arrow[l, shift left] \arrow[r, shift left] & 5 \arrow[l, shift left] \arrow[r, shift left] & 6 \arrow[l, shift left] \arrow[r, shift left] & \cdots \arrow[l, shift left] \arrow[r, "p_{N-1}"] & N \arrow[loop right, "1"]
\end{tikzcd}
$$

end states (boundaries) are "absorbing"

3 communication classes

"transient" states

$$\underbrace{\{0\}, \{1, 2, 3, \ldots N-1\}, \{N\}}_{\text{absorbing states}}$$

unbounded random walk

$$
\begin{tikzcd}
\cdots \arrow[r, shift left] & -2 \arrow[l, shift left] \arrow[r, shift left] & -1 \arrow[l, shift left] \arrow[r, shift left] & 0 \arrow[l, shift left] \arrow[r, shift left] & 1 \arrow[l, shift left] \arrow[r, shift left] & 2 \arrow[l, shift left] \arrow[r, shift left] & \cdots \arrow[l, shift left]
\end{tikzcd}
$$

$$
P = \begin{pmatrix}
 & 0 & p & 0 & q & 0 & 0 & \cdots \\
\cdots & 0 & 0 & p & 0 & q & 0 & 0 & \cdots \\
 & & \cdots & & p & 0 & q & \cdots
\end{pmatrix}
$$

Many other interesting/important motifs

$$\begin{tikzcd} \textcircled{1} \arrow[r] & \textcircled{2} \arrow[r] & \textcircled{3} \arrow[r, dashed] & \cdots \arrow[r] & \textcircled{N} \arrow[llll, bend left] \end{tikzcd}$$

loops (energy dissipation)

$$P = \begin{pmatrix} 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \\ 0 & 0 & 1 & \cdots & \\ \vdots & & & & \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

* **discrete step dynamics is more likely to have periodicity**

$d(i) \equiv$ periodicity of state i.

defined as the greatest common divisor of all integers n such that

$$[P^n]_{ii} > 0$$

if $d(i) > 1$, then state i is "periodic"

if $d(i) = 1$, then state i is "aperiodic"

Suppose $\textcircled{0} \rightleftharpoons \textcircled{1}$ and $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

State flips between $\textcircled{0}$ and $\textcircled{1}$ with each time step

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$ so $[P^2]_{ii} = 1$, same with $n = 4, 6 \ldots$

g.c.d $= 2$ $\quad d(i) = 2$

$d(i)$ holds for all i in each communication class.

Infinite random walk has $d(i) = d = 2$

(can return in two steps)

First return and first passage times

$f_{ii}^{(n)} \equiv$ prob. that starting from i, $X_0 = i$, the first return to state i occurs at the n^{th} step

$$f_{ii}^{(n)} \equiv \text{Prob}\left[X_n = i,\ X_m \neq i,\ m = 1, 2, 3 \ldots, n-1 \mid X_0 = i\right]$$

$f_{ii}^{(0)} = 0$; $\quad f_{ii}^{(1)} \equiv p_{ii}$ but $f_{ii}^{(n)} \neq [P^n]_{ii}$

this represents <u>first</u> time returning to i.

we must have

$$\boxed{0 \leq \sum_{n=1}^{\infty} f_{ii}^{(n)} \leq 1}$$

State i is transient if $\sum_{n=1}^{\infty} f_{ii}^{(n)} < 1$

and "recurrent" if $\sum_{n=1}^{\infty} f_{ii}^{(n)} = 1$

First return time is random variable (with a distribution)

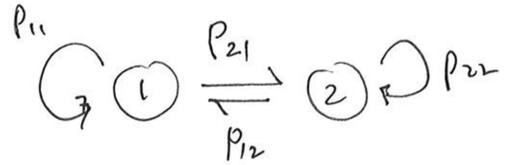
$$T_{ii} = \inf_{m \geq 1} \{ m \mid X_m = i \text{ and } X_0 = i \}$$

$T_{ii} = n$ with probability $f_{ii}^{(n)}$, $n = 0, 1, 2, \ldots$

mean of $T_{ii} \equiv \mathbb{E}[T_{ii}] = \sum_{n=1}^{\infty} n f_{ii}^{(n)}$

if i is a transient state, recurrence time $\to \infty$, and mean $\mathbb{E}[T_{ii}] \to \infty$ is not defined.

Consider 2-state model $\quad P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$



$f_{11}^{(1)} = p_{11}, \quad f_{11}^{(2)} = p_{21} p_{12}$

$f_{11}^{(3)} = p_{21} p_{22} p_{12}$

$f_{11}^{(n)} = p_{21} p_{22}^{n-2} p_{12}, \quad n \geq 3$

$$\sum_{n=1}^{\infty} f_{11}^{(n)} = p_{11} + p_{21} p_{12} \sum_{n=0}^{\infty} p_{22}^{n}$$

$$= p_{11} + \frac{p_{12} p_{21}}{1 - p_{22}} = 1$$, state 1 is recurrent

First passage times

Define $f_{ji}^{(n)}$ $(j \neq i)$ as $\text{Prob}[X_n = j, X_m \neq j, m = 1, 2, \ldots n-1 \mid X_0 = i]$

$\equiv$ First passage time probabilities $\quad f_{ji}^{(0)} = 0$

$$T_{ji} = \inf_{m \geq 1} \{ m \mid X_m = j \text{ and } X_0 = i \}$$

Mean first passage time

$$\mathbb{E}[T_{ji}] = \sum_{n=1}^{\infty} n f_{ji}^{(n)}$$

Connecting n-step probabilities and first return or first passage probabilities:

$[P^n]_{ii}$ could have first return at step $j = 1, 2, 3, \ldots n$

then, $$[P^n]_{ii} = \underset{\downarrow \atop 0}{f_{ii}^{(0)}} [P^n]_{ii} + f_{ii}^{(1)} [P^{n-1}]_{ii} + \cdots + f_{ii}^{(n)} \underbrace{P_{ii}^{(0)}}_{\equiv 1}$$

$$= \sum_{k=1}^{n} f_{ii}^{(k)} [P^{n-k}]_{ii}$$

$[P^n]_{ii}$ can have first return at $n = 1, 2, 3, \ldots$

Similarly, $$[P^n]_{ji} = \sum_{k=1}^{n} f_{ji}^{(k)} [P^{n-k}]_{jj}$$

Define generating function for sequence $f_{ji}^{(n)}$

as $$F_{ji}(z) = \sum_{n=0}^{\infty} f_{ji}^{(n)} z^n$$

and for $[P^n]_{ji}$ $$P_{ji}(z) = \sum_{n=0}^{\infty} [P^n]_{ji} z^n$$

$$F_{ii}(z) P_{ii}(z) \equiv \sum_{r=1}^{\infty} c_{ii}^{(r)} z^r = P_{ii}(z) - 1$$

$r \geq 1$ (not zero) because:

where $$c_{ii}^{(r)} = \sum_{k=1}^{r} f_{ii}^{(k)} [P^{r-k}]_{ii} \quad \left(= [P^r]_{ii} \right)$$

$$\therefore \quad P_{ii}(z) = \frac{1}{1 - F_{ii}(z)}$$

$$P_{0i}(z) = \overset{1}{P_{00}} \, F_{0i}(z)$$

P_0^0

Similarly, show (HW?)

$$F_{ji}(z) P_{jj}(z) = P_{ji}(z) \qquad i \neq j$$

Prove that if state i is $\overbrace{\text{recurrent}}^{\sum_{n=1}^{\infty} f_{ii}^{(n)} = 1}$, if & only if $\sum_{n=0}^{\infty} [P^n]_{ii}$ diverges,

that is $$\sum_{n=0}^{\infty} [P^n]_{ii} = \infty$$