Bayes rule

cause $X \sim p(x)$ prior

$p(x|y)$ $\uparrow$ $\downarrow$ $p(y|x)$ generation (likelihood)

effect Y

$p(x|y)$: posterior, inference, inversion

$$p(x|y) = \frac{p(x,y)}{p(y)} \quad \text{conditioning}$$

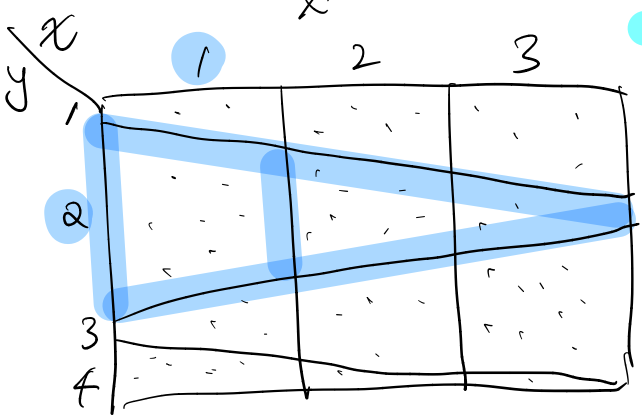
$$= \frac{p(x)\,p(y|x)}{\sum_{x'} p(x',y)} \quad \text{factorization, marginalization}$$

$$= \frac{p(x)\,p(y|x)}{\sum_{x'} p(x')\,p(y|x')}$$

disease

prob = proportion

symptom

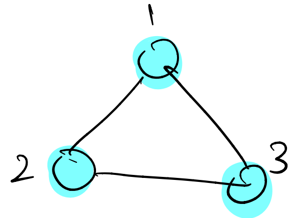


population
region
repetition

just count

Markov chain

state space



with prob $= \frac{1}{2}$, stay

$= \frac{1}{2}$, move to one of the other 2 states

$$X_0 \to X_1 \to \cdots \to X_t \to X_{t+1} \to \cdots$$

X_t : state at time t ($X_0 = 1$, e.g.)

y

$X_t \backslash X_{t+1}$	1	2	3
1			
2			
3			

x

Markov property

$p(\text{future} \mid \text{present}, \text{past})$

$= p(\text{future} \mid \text{present})$

$$p(x, y) = p(X_t = x \ \& \ X_{t+1} = y)$$

$$p_t(x) = p(X_t = x) \quad p_{t+1}(y) = p(X_{t+1} = y)$$

$$p(y \mid x) = p(X_{t+1} = y \mid X_t = x, \ X_{t-1} \cdots X_0)$$

$$= T(x, y)$$ Transition prob

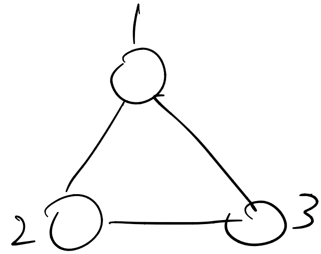
Transition matrix

$$T(x,y) = P(X_{t+1} = y \mid X_t = x)$$

$y \to j$

$x \to i$

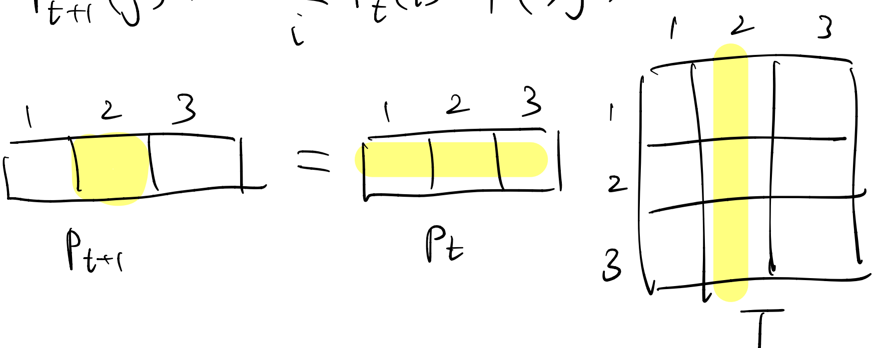
$$\begin{array}{c|ccc} X_t \backslash X_{t+1} & 1 & 2 & 3 \\ \hline 1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 2 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 3 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{array} \quad T$$



$$p(y) = \sum_x p(x,y) \quad \text{marginalization}$$

$$= \sum_x p(x)\, p(y|x) \quad \text{factorization}$$

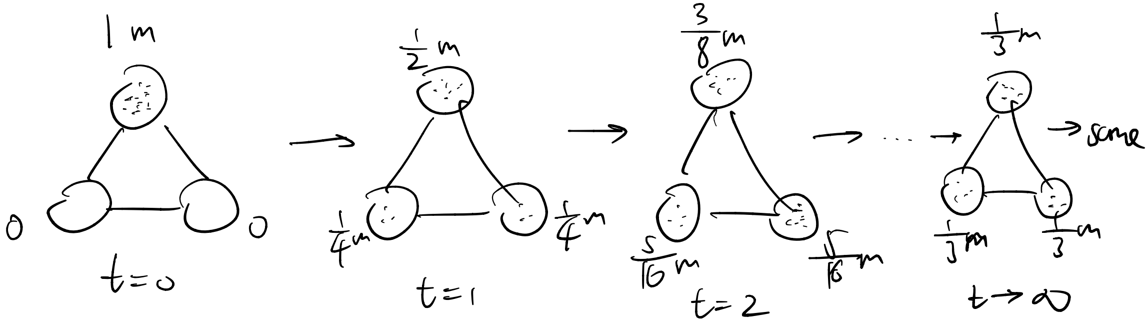
$$P_{t+1}(j) = \sum_i P_t(i)\, T(i,j)$$

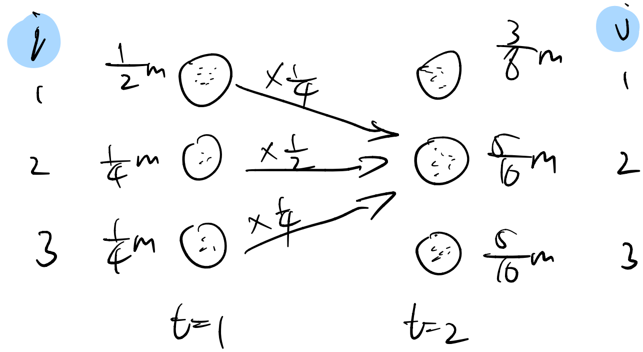


$$P_{t+1} = P_t \cdot T$$

Population migration

1 million people





arrow of time

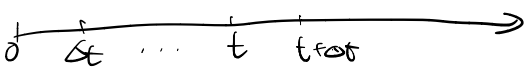
(Bayes rule ? $p(x|y)$)

$$P_{t+1}(j) = \sum_i P_t(i)\, T(i,j)$$

$$P_{t+1} = P_t \cdot T$$

$$P_t = P_0\, T^t \xrightarrow[\infty]{t} \pi \quad \boxed{\tfrac{1}{3} \,\big|\, \tfrac{1}{3} \,\big|\, \tfrac{1}{3}}$$

Continuous time : Markov jump

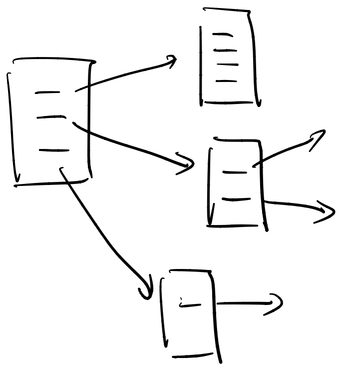


$$P(X_{t+\delta t} = j \mid X_t = i) = \lambda_{ij}\, \delta t \quad \text{(Poisson)}$$

Google Page Rank

state space : all webpages

random surfing : randomly click a link



population migration $\longrightarrow$ $\pi \approx p_0 T^{t \to \infty}$

$\pi(i)$: popularity of i, ranking

$$\pi = \pi T$$

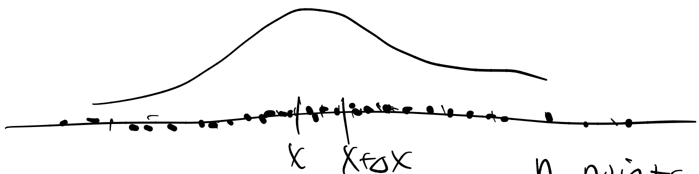
$$\pi(j) = \sum_i \pi(i) T(i, j)$$

2 random variables

Continuous

Recall one R.V.

$$X \sim f(x) \quad \text{density}$$



n points $\quad n \to \infty$

$\Delta x \to 0$

$n(x)$ = # of points in $(x, x+\Delta x)$

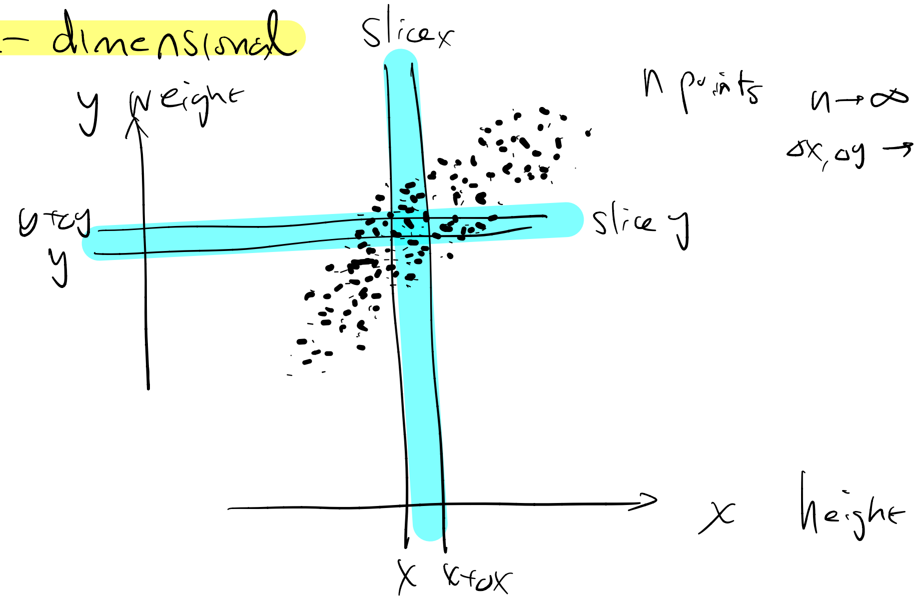
$$f(x) = \frac{n(x)/n}{\Delta x} = \frac{P(X \in (x, x+\Delta x))}{\Delta x}$$

prob = proportion

density = proportion density

$$= \frac{\text{proportion}}{\text{size}}$$

2-dimensional



n points

$n \to \infty$

$\Delta x, \Delta y \to 0$

$$n(x,y) = \#\text{ of points in } (x, x+\Delta x) \times (y, y+\Delta y)$$

joint: $$f(x,y) = \frac{n(x,y)/n}{\Delta x \Delta y} = \frac{P(X \in (x, x+\Delta x)\ \&\ Y \in (y, y+\Delta y))}{\Delta x \Delta y}$$

① **marginalization**:

$$n(x) = \#\text{ of points in } (x, x+\Delta x)$$

$$n(y) = \quad \dots \quad (y, y+\Delta y)$$

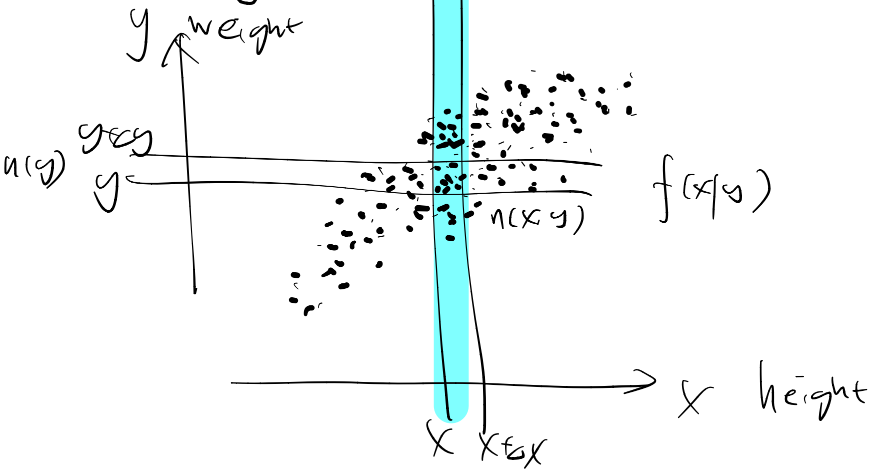
$$n(x) = \sum_y n(x,y)$$

$$f(x) = \frac{n(x)/n}{\Delta x} = \frac{\sum_y n(x,y)/n}{\Delta x}$$

$$= \frac{\sum_y f(x,y)\Delta x \Delta y}{\Delta x}$$

$$= \sum_y f(x,y)\Delta y = \int f(x,y)\,dy$$

② Conditioning



$$f(x,y) = \frac{n(x,y)/n}{\Delta x \Delta y}$$

$$f(x) = \frac{n(x)/n}{\Delta x}$$

$$f(y|x) = \frac{n(x,y)/n(x)}{\Delta y}$$

$$= \frac{f(x,y)\,\Delta x \Delta y \,/\, f(x)\,\Delta x}{\Delta y}$$

$$= \frac{f(x,y)}{f(x)}$$

③ Factorization

$$f(x,y) = f(x)\,f(y|x)$$

Bivariate Normal

$$X \sim N(0,1)$$

$$[Y \mid X = x] \sim N(\underset{\mu}{\rho x},\ \underset{\sigma^2}{1-\rho^2})$$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$f(y \mid x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

$$= \frac{1}{\sqrt{2\pi(1-\rho^2)}} e^{-\frac{(y-\rho x)^2}{2(1-\rho^2)}}$$

$$f(x,y) = f(x) f(y \mid x)$$

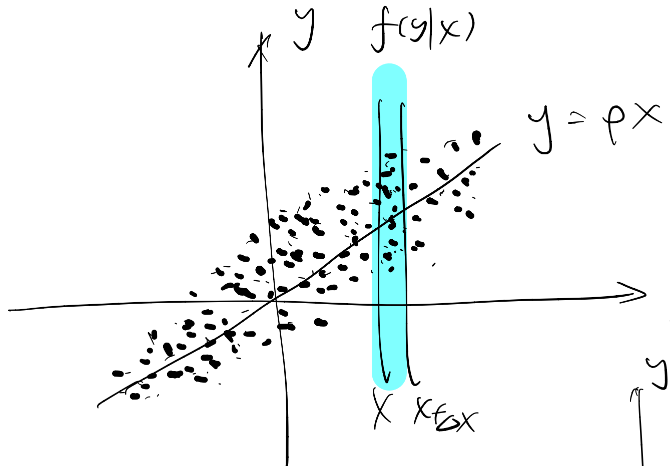
$$= \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{x^2}{2} - \frac{(y-\rho x)^2}{2(1-\rho^2)}}$$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{x^2 - \cancel{x^2\rho^2} + y^2 - 2\rho xy + \cancel{\rho^2 x^2}}{2(1-\rho^2)}}$$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}}$$

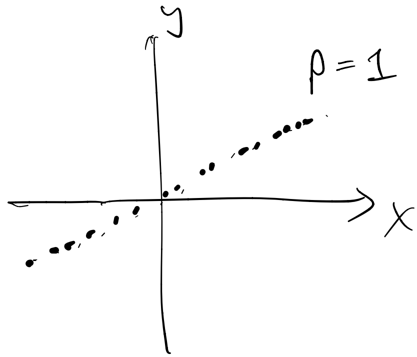
symmetric

Conditional expectation & Variance

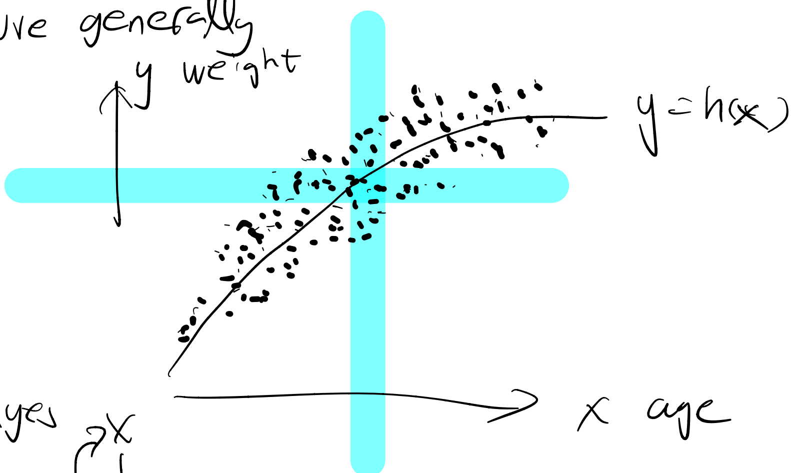


$$E(Y \mid X = x) = \rho x$$

$$Var(Y \mid X = x) = 1 - \rho^2$$



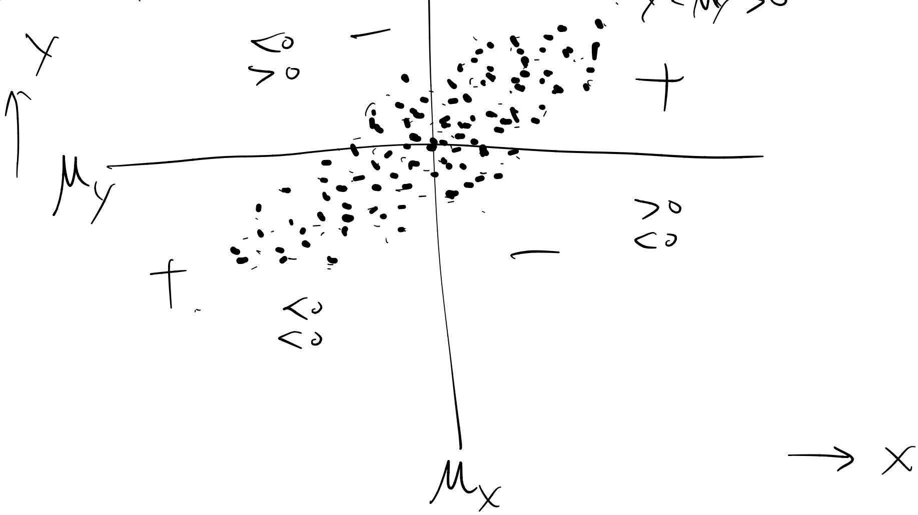
More generally



Bayes

$$\begin{tikzcd} X \arrow[d] \\ Y \arrow[u, bend left] \end{tikzcd}$$

Covariance



$$\mu_X = E(X) \qquad \mu_Y = E(Y)$$

$$Cov(X, Y) = E\big((X - E(X))(Y - E(Y))\big)$$

$$Cov(X, X) = Var(X) = \sigma_X^2$$
$$Cov(Y, Y) = Var(Y) = \sigma_Y^2$$

$$Corr(X, Y) = Cov\left(\frac{X - \mu_X}{\sigma_X}, \frac{Y - \mu_Y}{\sigma_Y}\right)$$

$$= \frac{Cov(X, Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \rho \text{ standardize}$$