11/1/22

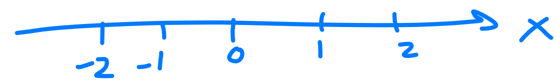
# Part 3 : Conditioning

H = move forward
T = move backward

Two scenarios

(1) Forward $P(A|B)$

random walk




$$P(X_{t+1}=y \mid X_t = x) = \begin{cases} \frac{1}{2} & y = x+1 \\ \frac{1}{2} & y = x-1 \\ 0 & \text{others} \end{cases}$$

Transition probability $T(x,y)$

$$X_{t+1} = X_t + \varepsilon_t$$

$$\varepsilon_t \perp X_t$$

| $\varepsilon$ | $-1$ | $+1$ |
|---|---|---|
| prob | $\frac{1}{2}$ | $\frac{1}{2}$ |

ex.) Dynamics

$P(X_{t+1} \mid X_t, a_t)$
action

Policy

$P(a_t \mid X_t$

ex.) A : alarm
B : fire

$P(A|B) = 99.9\%$

$P(A|B^c) = .001\%$

can be estimated by frequencies under $B$ or $B^c$

$$P(A \cap B) = P(B)P(A|B)$$

how often both happen — how often B happens — when B happens, how often does A happen

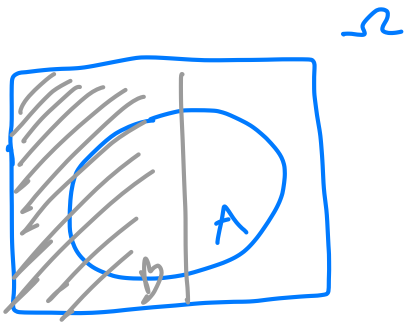
## (2) Inverse inference

$$\text{cause} \xrightarrow{①} \text{effect}$$
$$\text{cause} \xleftarrow{②\text{: inference}} \text{effect}$$

subjective belief

A : alarm    B : fire

$P(B|A)$




randomly sample point from $\Omega$

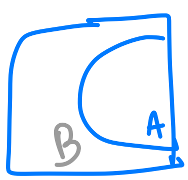
$$P(A) = \frac{|A|}{|\Omega|}$$ prior belief

Now you are told the point falls in B.

axiom 4

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

Re-imagine as if we randomly throw point in B




↳ like new $\Omega$

$$P(A|B) = \frac{|\text{new } A|}{|\text{new } \Omega|} = \frac{|A \cap B|}{|B|}$$

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

$P(A) = P(A|\Omega)$

rewrite as:

$$P(A \cap B) = P(B)P(A|B)$$

wave function $\Psi(x)$

$$f(x) = |\Psi(x)|^2$$

If we observe $X = x$    $P(X = x) = 1$

$$f(x | X = x) = \delta_x$$

Meta rule: Insert the same condition into any equations

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A \cup B|C) = P(A|C) + P(B|C) - P(A \cap B|C)$$

subregion/
sub popo'

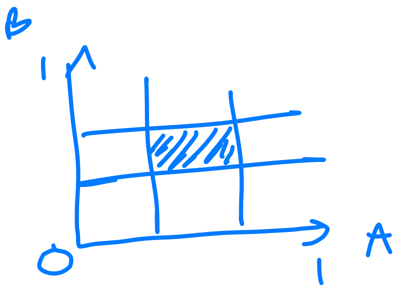
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

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

$$P(A \cap B \cap C) = P(B \cap C)\,P(A|B \cap C)$$

$$= P(C \qquad P($$

$$A \perp B \iff P(A \cap B) = P(A)P(B)$$




$$\iff P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)\,\cancel{P(B)}}{\cancel{P(B)}} = P(A)$$

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

# Conditional Independence

$A \perp B \mid C$

(1) $P(A \cap B \mid C) = P(A \mid C)\, P(B \mid C)$

(2) $P(A \mid B \cap C) = P(A \mid C)$

(1)
$$\begin{tikzcd} & C \text{ parent} \arrow[dl] \arrow[dr] & \\ \text{child } A & & B \text{ child} \end{tikzcd}$$

shared cause

(2) $B \longrightarrow C \longrightarrow A$

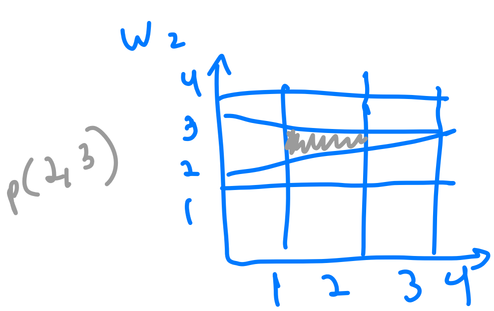
past present future

markovian process

Switch notation to random variables

<u>Discrete</u> $(X, Y) \sim p(x,y) = p(X = x \,\&\, Y = y)$

| X : eye color | Y hair color 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | p(1,1) | p(1,2) | | |
| 2 | | | | |
| 3 | | | | |
| 4 | | | | |




"Marginalization"

$p(x,y)$ = area of cell

$p(x)$ = area of zone ( $\sum_y$ area of cell $(x,y)$ )

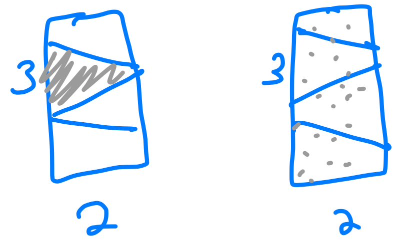
$= \sum_y p(x,y)$

$p(y) = \sum_x p(x,y)$

## "Conditioning"

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

among sub population of eye color 2
what is the proportion of these w/ hair color 3




$P(\text{alarm} | \text{fire}) = 1$

$P(\text{fire} | \text{alarm}) \approx 0$

## "Factorization"

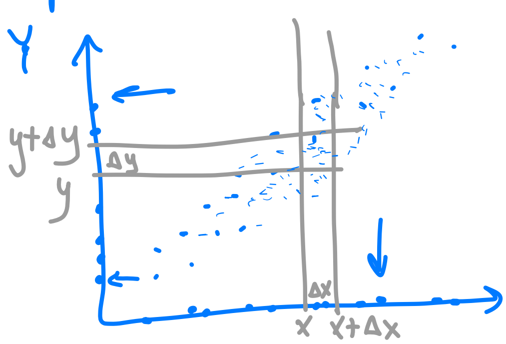
$$p(x,y) = p(x)\,p(y|x) = p(y)\,p(x|y)$$

## Continuous

$(X, Y) \sim f(x,y)$

$$P(X \in (x, x+\Delta x) \;\&\; Y \in (y, y+\Delta y)) = f(x,y)\,\Delta x\,\Delta y$$

Scatterplot




$n \to \infty$

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

to get density of $x$, project points onto $x$ axis. same for $y$

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

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

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

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

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

"marginalization"

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

similarly $f(y) = \int f(x,y)\,dx$

normalizing over size

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

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

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

"Conditioning" $= \frac{\text{Joint}}{\text{Marginal}}$

**Marginalization**

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

$$p(y) = \sum_x p(x,y) \qquad f(y) = \int f(x,y)\,dx$$

**Conditioning / Normalization**

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

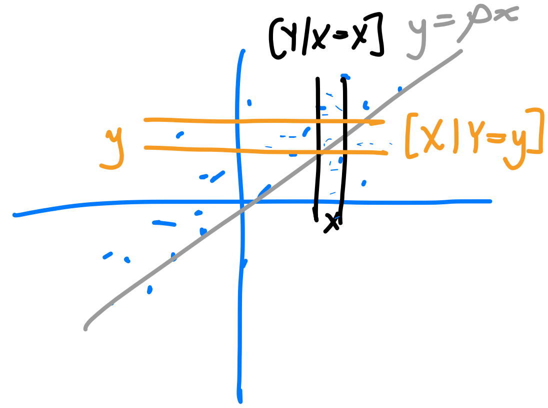
$$p(x|y) = \frac{p(x,y)}{p(y)} \qquad f(x|y) = \frac{f(x,y)}{f(y)}$$

**Factorization**

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

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

Bivariate Normal




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

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

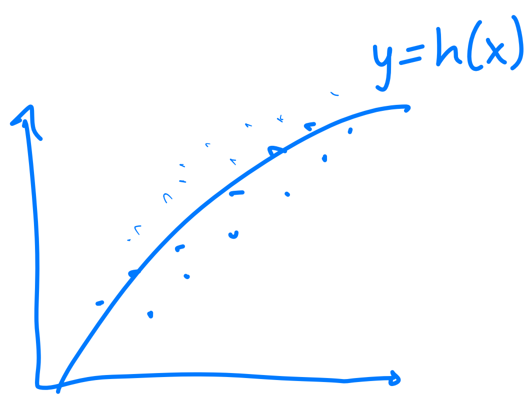
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

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

$$f(x,y) = f(x)f(y|x) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2}\left(x^2 + \frac{(y-\rho x)^2}{(1-\rho^2)}\right)}$$

$$\frac{x^2 - x^2\rho^2 + y^2 - 2\rho xy + \rho^2 x^2}{1-\rho^2}$$

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




$$[X \mid Y=y] \sim N(\rho y, 1-\rho^2)$$

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

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

conditional expectation

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

$$Var(Y \mid X=x)$$

$$Y = h(x) + \varepsilon$$