2 & More Random Variables

$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ $\quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ 2D vector point

Discrete:

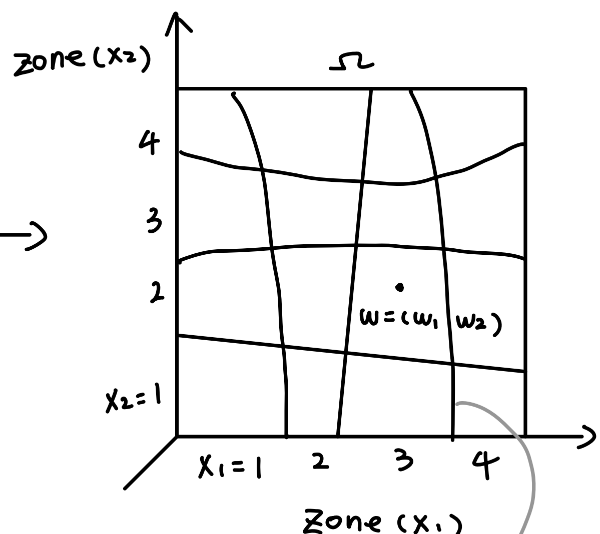
$p(x) = p(x_1, x_2)$

$p(x) = P(X = x) = P(\underset{\text{ex. eye color}}{X_1 = x_1} \,\&\, \underset{\text{hair color}}{X_2 = x_2})$

$E(h(X)) = \sum_x h(x) \cdot p(x)$

$$E(h(X_1, X_2)) = \sum_{x_1} \sum_{x_2} h(x_1, x_2)\, p(x_1, x_2)$$

X_1=eye color \ X_2=hair color	1	2	3	4
1	p(1,1)	p(1,2)	p(1,3)	p(1,4)
2	p(2,1)	⋱		⋮
3	p(3,1)		⋱	
4	p(4,1) ⋯			p(4,4)



If X_1 and X_2 are independent, the lines would become straight.

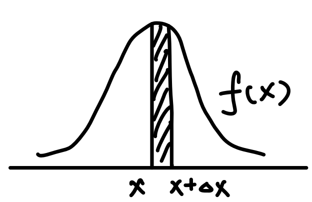
$p(x_1, x_2)$: population proportion

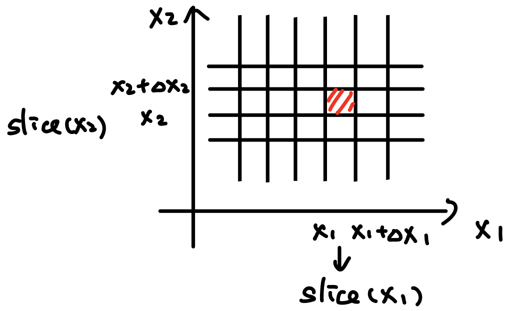
$X(w) = (\underset{\text{eye color}}{X_1(w)}, \underset{\text{hair color}}{X_2(w)})$

Continuous:

$f(x) = f(x_1, x_2)$

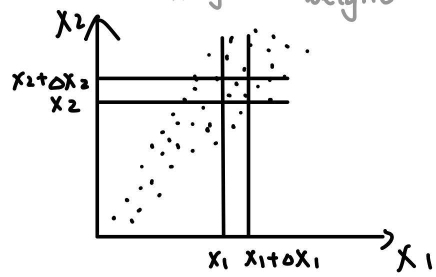
$P(X \in (x, x+\Delta x)) = f(x)\Delta x$





$$P(X_1 \in (x_1, x_1+\Delta x_1) \ \& \ X_2 \in (x_2, x_2+\Delta x_2))$$
$$= f(x_1 \ x_2)\Delta x_1 \Delta x_2$$
$$f(x_1 \ x_2) = \frac{P(\quad)}{\Delta x_1 \Delta x_2} \rightarrow \text{prob mass}$$

Ex. $X = (X_1 \quad X_2)$
height weight

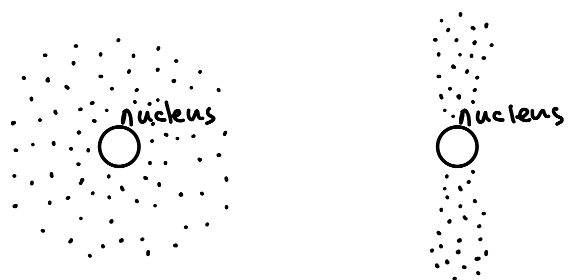


n points ($n \to \infty$)

$n(x_1 \ x_2)$: # of points in
$(x_1, x_1+\Delta x_1) \times (x_2, x_2+\Delta x_2)$

$$\frac{n(x_1 \ x_2)/n}{\Delta x_1 \Delta x_2} \xrightarrow[\infty]{n} f(x_1 \ x_2)$$

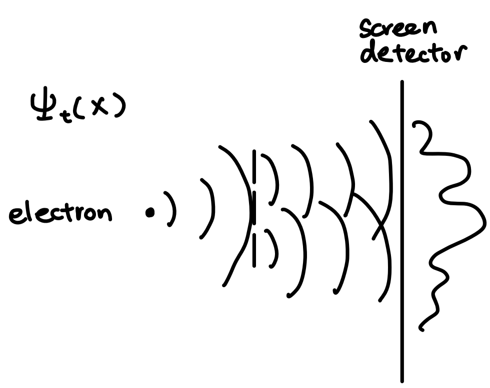
Ex. electron cloud around nucleus



$f_t(x) \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

Wave Function :

$\Psi_t(x)$



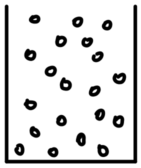
When measuring where the electron is:

$$f_t(x) = |\Psi_t(x)|^2$$
$$|a+ib|^2 = a^2+b^2$$

2 electrons :

$$X = \begin{pmatrix} x^{(1)} \\ x^{(2)} \end{pmatrix}$$

A Cup of Water :



10^{23}

$$\mathcal{X} = \begin{pmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(10^{23})} \end{pmatrix} \qquad \mathcal{X} \sim \text{Unif}(\Omega)$$

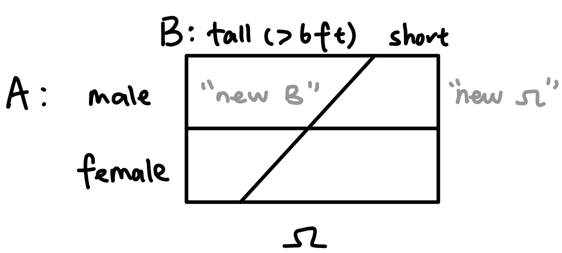
$$\Omega = \{ \underbrace{\mathcal{X}(t)}_{\text{snapshot at times}},\ t \in [0, T \to \infty] \}$$

$$E(h(\mathcal{X})) = \int h(x) f(x)\, dx$$

$$E(h(\mathcal{X}_1, \mathcal{X}_2)) = \iint h(x_1, x_2) f(x_1, x_2)\, dx_1\, dx_2$$

$$\text{Var}(h(\underset{(\mathcal{X}_1,\ \mathcal{X}_2)}{\underbrace{\mathcal{X}}})) = E((h(\underset{(\mathcal{X}_1,\ \mathcal{X}_2)}{\underbrace{\mathcal{X}}}) - E(h(\underset{(\mathcal{X}_1,\ \mathcal{X}_2)}{\underbrace{\mathcal{X}}})))^2)$$

$\boxed{\text{Conditioning}}$



Axiom : 0

$$P(B) = \frac{|B|}{|\Omega|} \begin{array}{l} - \text{ \# of tall people} \\ - \text{ total \# of whole population} \end{array}$$

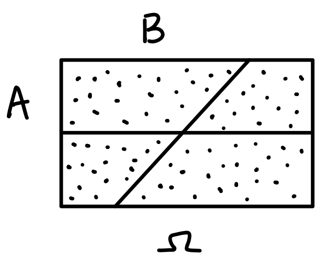
Axiom 4 (or definition) :

$$P(\underset{\text{given } A \text{, as if sample from } A \text{ ("new } \Omega\text{")}}{\underbrace{B \mid A}}) = \frac{|A \cap B|}{|A|} = \frac{|A \cap B| / |\Omega|}{|A| / |\Omega|} = \frac{P(A \cap B)}{P(A)}$$

$\Rightarrow P(A \cap B) = P(A)\, P(B \mid A)$

Independence: $P(A \cap B) = P(A) \cdot P(B)$

Another definition of independence ← $P(B \mid A) = \frac{P(A \cap B)}{P(A)} = P(B)$



repeat n times

$P(B)$: how often B happens, long run frq

$P(B|A)$: when A happens, how often B happens

ex. $P(\text{alarm} \mid \text{fire}) \approx 99.9\%$

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

2 R.V.'s

Change notation $(X_1, X_2) \longrightarrow (X, Y)$

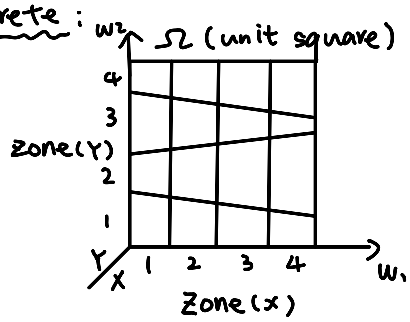
ω : random person (point)

$X(\omega)$, $Y(\omega)$

X ↓ eye, height, gender

Y ↓ hair, weight, height

Discrete :



Joint :

$p(x,y) = \text{area of cell } (x,y)$

(1) **Marginalization** :

$$p_X(x) = P(X = x) = \text{area of zone } x$$

$$= \sum_y \text{area of cell } (x,y)$$

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

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

(2) **Conditioning** :

$$p_{Y|X}(y|x) = P(\underset{B}{Y = y} \mid \underset{A}{X = x}) = \frac{P(A \cap B)}{P(A)} = \frac{p(x,y)}{p(x)} \begin{matrix} -\text{joint} \\ -\text{marginal} \end{matrix}$$

Among people of eye color x, proportion of people with hair color y.

When $X = x$, how often $Y = y$

(3)

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

joint ... marginal x conditioned

how often $X=x$ & $Y=y$	how often $X=x$	when $X=x$ how often $Y=y$

Summary:

(1) $p(x) = \sum_y p(x,y)$

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

(2) $p(y|x) = \dfrac{p(x,y)}{p(x)}$

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

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

Bayes Rule :

cause $X = \begin{cases} 1 & \text{fire} \\ 0 & \text{no fire} \end{cases}$

$\downarrow$

effect $Y = \begin{cases} 1 & \text{alarm} \\ 0 & \text{no alarm} \end{cases}$

prior: $p(x)$

x	0	1
p(x)	99%	1%

generation: $p(y|x)$

y \ x	0	1
0	98%	0.1%
1	2%	99.9%

reasoning:

$p(x|y)$ posterior prob

e.g. $P(X=1 \mid Y=1)$

$= P_{X|Y}(1|1)$

$$p(x|y) \overset{(2)}{=} \frac{p(x,y)}{p(y)} = \frac{p(x)\,p(y|x)}{\sum_{x'} p(x',y)} = \frac{p(x)\,p(y|x)}{\sum_{x'} p(x')\,p(y|x')}$$

$$\Rightarrow P_{X|Y}(1|1) = \frac{P_X(1)\,P_{Y|X}(1|1)}{P_X(0)\,P_{Y|X}(1|0) + P_X(1)\,P_{Y|X}(1|1)} = \frac{1\% \cdot 99.9\%}{99\% \cdot 2\% + 1\% \cdot 99.9\%}$$