 Jodian Mishtaku Aug 2022

I b) Probability Law assigns a non-negative value to an event(set) A, $P(A)$ → called probability of A, that encodes our knowledge or belief about the collective likelihood of the (elements) outcomes of the event A.

$A \subset \Omega$ $(A \in 2^{\Omega})$

Def $\boxed{\text{Probability Law}}$

(E, Ω, P) Ω is the sample space i.e. the set of all the possible outcomes of experiment E

Then a Probability Law ~~assigns~~ :

- Specifies the likelihood of any outcomes or/and of any set of possible outcomes (i.e. of an event)

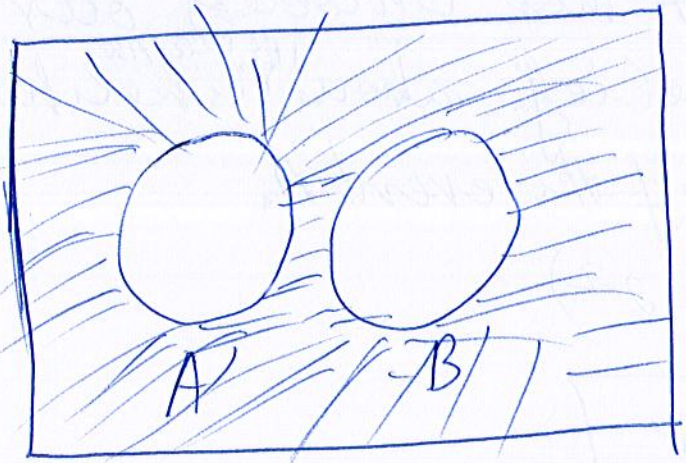
- Or alternatively assigns to any event A a ~~non-negative~~ number $P(A)$, called the probability of A, satisfying the folling axioms.

1. Nonnegativity $P(A) \geq 0, \forall A \subset \Omega$
2. Additivity : A, B disjoint events (i.e. $A \cap B = \emptyset$) then $P(A \cup B) = P(A) + P(B)$

   That is true for an infinite sequence of disjoint events $A_1, A_2, A_3 \dots$

   $P(A_1 \cup A_2 \cup A_3 \cup \dots) = P(A_1) + P(A_2) + \dots$
3. Normalization $P(\Omega) = 1$

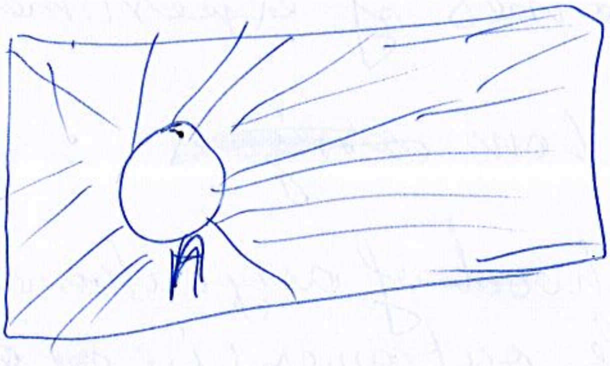
2



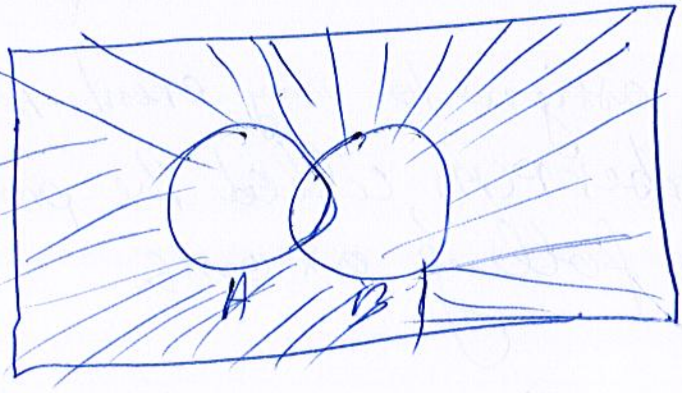
$A^c \cap B^c$

$\begin{cases} A \not\subset A^c \Rightarrow A \not\subset A^c \cap B^c \\ B \not\subset B^c \Rightarrow B \not\subset A^c \cap B^c \end{cases}$

Ω



Add to the diagram a B set



$\begin{cases} A \not\subset A^c \Rightarrow A \not\subset A^c \cap B^c \\ B \subset B^c \Rightarrow B \subset A^c \cap B^c \end{cases}$

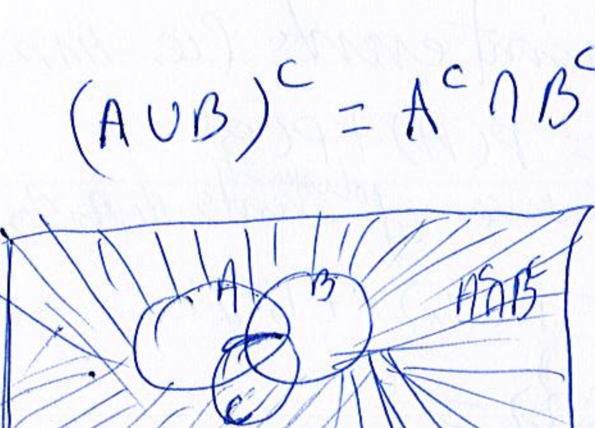
$\Downarrow$

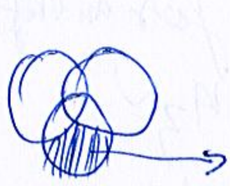
$A \cap B \not\subset A^c \cap B^c$

$A \cup B \not\subset A^c \cap B^c$

$(A \cup B)^c = A^c \cap B^c$

$A^c \cap B^c = (A \cup B)^c$





$\rightarrow (A^c \cap B^c) \cap C$

or $(A \cup B)^c \cap C$

③ $P(\Omega) = 1$ $\quad P(\emptyset) = 0$

Theorem:

a) $A \subset B \Rightarrow P(A) \leq P(B)$

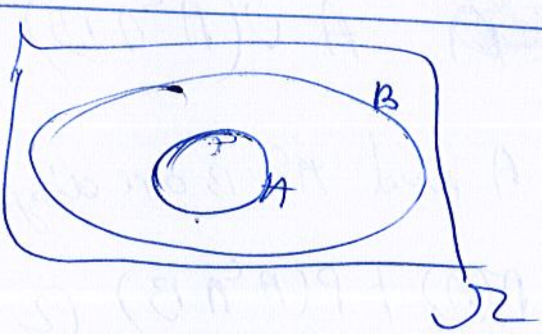
b) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

c) $P(A \cup B) \leq P(A) + P(B)$

d) $P(A \cup B \cup C) = P(A) + P(A^c \cap B) + P(A^c \cap B^c \cap C)$

Proof:

a)



$A \subset B \subset \Omega$

(1) $P(B) = P(A) + P(A^c \cap B)$ because $A, A^c \cap B$ are disjoint

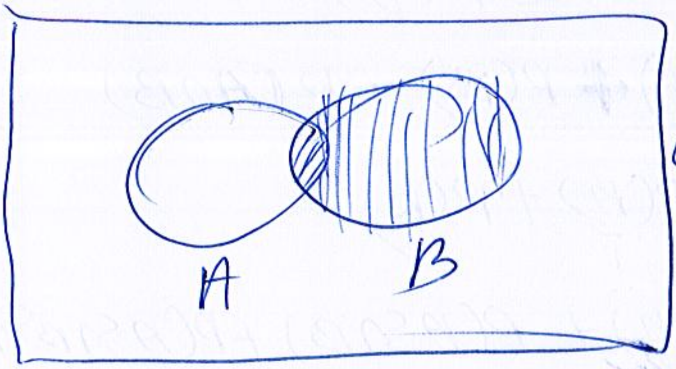
(2) $P(B) - P(A) = P(A^c \cap B) \geq 0 \Rightarrow$

(3) $P(B) \geq P(A) \Leftrightarrow P(A) \leq P(B)$

or from (1) $P(B) = P(A) + P(A^c \cap B) \geq P(A)$

(4)

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

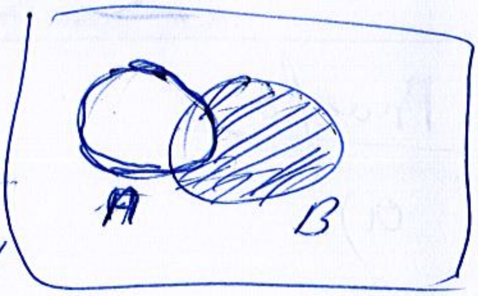


$B = \underbrace{(A \cap B) \cup (A^c \cap B)}_{\text{disjoint}}$

axiom 2 $\Downarrow$

$P(B) = P(A \cap B) + P(A^c \cap B)$ (1)

$A \cup B = A \cup (A^c \cap B)$



axiom 2 $\Downarrow$ A and $A^c \cap B$ are disjoint

$P(A \cup B) = P(A) + P(A^c \cap B)$ (2)

$\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$ (3)

(2)

~~(1) $P(B) = P(A \cup B) + P(A^c \cap B) = P(A) + P(A^c \cap B) + P(A \cap B)$~~

~~$P(A) = P(A^c \cap B) - P(A \cup B)$~~

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

c) probve that $P(A \cup B) \leq P(A) + P(B)$

(1) $\Rightarrow$ $P(A \cap B) = P(B) - P(A^c \cap B)$

(2) $\Rightarrow$ $P(A \cup B) = P(A) + P(A^c \cap B)$

$2P(A \cup B) = P(B) + P(A)$

$P(A \cup B) = \dfrac{P(B) + P(A)}{2} \leq P(B) + P(A)$

one as $P(A \cap B) \geq 0$ $\Rightarrow$ $P(A \cup B) \leq P(A) + P(B)$

5.

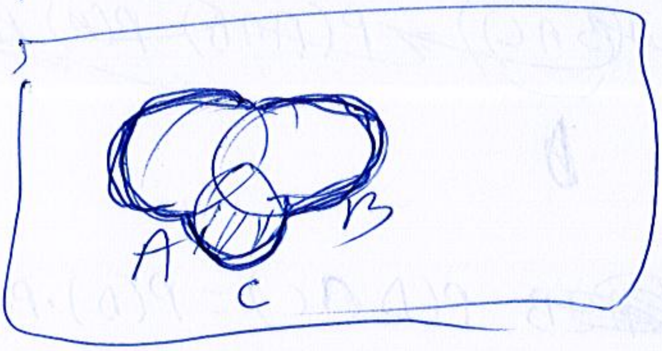
Proove that

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

fig 1

$$(A \cup B)^c = A^c \cap B^c$$

$\cancel{P(A \cup B)}$



$$A \cup B \cup C = \underbrace{A \cup A^c \cap B}_{\text{fig 1}} \cup (A^c \cap B^c) \cap C$$

$$A \cup B = A \cup (A^c \cap B)$$

$$(A \cup B) \cup C = \quad \cup A^c \cap B^c \cap C$$

$A, A^c \cap B, A^c \cap B^c \cap C$ are disjoint, Hence

$\cancel{P(A \cup B \cup C) = P(A) + P(B) + P}$

$$P(A \cup B \cup C) = P(A) \cup P(A^c \cap B) \cup P(A^c \cap B^c \cap C)$$

$\therefore$

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

Again on conditional independency

Def

A, B are ind given $C \iff P(A \cap B | C) = P(A|C) P(B|C)$

$\iff \begin{cases} P(B \cap C) \geq 0 \\ P(A | B \cap C) = P(A|C) \end{cases}$

Question ~~What~~ would like $P(A \cap B | C)$ if A, B are not ind given C?

$P(A \cap B \cap C) = \cancel{P(A \cap \underbrace{(B \cap C)}_{D})} = \cancel{P(A \cap B) = P(A) \cdot P(A|D)}$

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

$= P((A \cap B) \cap C) = \cancel{P(A \cap B} \quad P(D \cap C) = P(D) \cdot P(D|C)$ (1)

$= P(C|D) \cdot P(D)$ (2)

(1) $\cancel{= P(A \cap B) \cdot P(A \cap B | C)} \Rightarrow P(A \cap B | C) = \dfrac{P(A \cap B \cap C)}{\cancel{P(A \cap B)}}$

(2) $\cancel{= P(C | A \cap B)} \cdot P(A \cap B)$

$= \dfrac{P(C \cap A \cap B)}{\cancel{P(A \cap B}} = \dfrac{P(C \cap A|B) P(B)}{P(A|B) \cdot P(B)}$

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

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

7)

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

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

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

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

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

$$\boxed{P(A\cap B|C) = \frac{P(A|B\cap C)\cdot P(B|C)\cdot \cancel{P(C)}}{\cancel{P(C)}}} \quad (1)$$

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

$$= \boxed{P(B|A\cap C)\cdot P(A|C)} \quad (2)$$

$$(1')\ \boxed{P(A\cap B|C) = P(A|B\cap C)\cdot P(B|C)} \Big\} \overset{\text{if ind}}{=} P(A|C)\cdot P(B|C)$$

$$(2')\ \boxed{P(A\cap B|C) = P(B|A\cap C)\cdot P(A|C)} = P(B|C)\cdot P(A|C)$$

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

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

(8)

$X \to$ Geometric R.V. representing the number of attempts until the first success (pass the test) come up.

PMF: $p_X(x) = (1-p)^{x-1} \cdot p$

Let A = the event student passes the test within n attempts $A = \{X \le n\}$

$$\boxed{P(A) = \sum_{k=1}^{n} (1-p)^{k-1} p}$$

$$P_{X|A}(x) = \frac{P(\{X = x\} \cap A)}{P(A)} = \frac{P(X = x, A)}{P(A)} =$$

$p_X(x) = (1-p)^{x-1} p$ $\quad p_X(k) = (1-p)^{k-1} \cdot$ $\quad P(\{X=k\} \cap A) = (1-p)^{k-1} \cdot p$ where $1 \le k \le n$ $\to A$

$$P_{X|A}(k) = \begin{cases} \dfrac{(1-p)^{k-1} \cdot p}{\sum_{m=1}^{n} (1-p)^{m-1} p} & \text{for } k = 1, 2, 3 \ldots n \\ 0 & \end{cases}$$

$P(\{X=k\} \cap A)$ / $P(A)$

$$P_{X|A}(x) = \frac{P(\{X=x\} \cap A)}{P(A)}$$

$$P_{X|A}(k) = \frac{P(\{X=k\} \cap A)}{P(A)}$$

$$P_{X|A}(k) = (1-p)^{k-1} \cdot p \quad \underbrace{1 \le k \le n}_{A}$$

(9)

A → event that student succeeds within n attemps

$\boxed{p_X(k) = (1-p)^{k-1} p}$

$Val(X) = \{1, 2, 3, \dots\}$

~~$P_{X|A}(k) = (1-p)^{k-1} p \quad 1 \le k \le n$~~

$\boxed{P(\{X = k\} \cap A) = (1-p)^{k-1} \cdot p \text{ and } 1 \le k \le n}$

$$\boxed{p_{X|A}(k) = \begin{cases} \dfrac{(1-p)^{k-1} \cdot p}{\sum_{m=1}^{n} (1-p)^{m-1} p} & \text{for } 1 \le k \le n \\ 0, & \text{otherwise} \end{cases}}$$

$$\boxed{P(A) = \sum_{m=1}^{n} (1-p)^{m-1} \cdot p}$$

- A_1 event that $\{X = 1\}$
- A_2 event that $\{X > 1\}$

$E[X] = ?$

$$E[X] = \sum_{i=1}^{2} E[X|A_i] \cdot P(A_i)$$

$$E[X] = E[X|A_1] P(A_2) + E[X|A_2] P(A_2)$$

$\underline{P(A_1)} = P(\{X=1\}) = \sum_{m=1}^{m=1} (1-p)^{m-1} p = \underline{p}$

$\underline{P(A_2)} = P(\{X > 1\}) = 1 - P(\{X \le 1\}) = 1 - P(A_1) = 1 - P(\{X=1\}) = \underline{1-p}$

$E[X|A_1] = \sum x \cdot p_{X|A_1}(x)$

(10)

$$\boxed{E[X|A_1] = \sum_{x=1}^{\infty} x \cdot P_{X|A_1}(x)}$$

$$\boxed{E[X|A_2] = \sum_{x=1}^{\infty} x \, P_{X|A_2}(x)}$$

$$\boxed{P_{X|A_1}(x)} = \frac{(1-p)^{x-1} \cdot p}{P(A_1)} \quad \text{bilance} = \frac{P(\{X=x\} \cap A_1)}{P(A_1)} \quad A_1 = \{X=1\}$$

$$= \frac{(1-p)^{1-1} \cdot p}{P(A_1)} = \frac{p}{p} = 1 \Big] = \begin{cases} 1, & x=1 \\ 0, & \text{for the rest of } x \in R(X) \end{cases}$$

0, otherwise

we could write $P_{X|A}(k)$
(so a konkret value k from $\{2,3,\ldots\}$)

konkret value

$$\boxed{P_{X|A_2}(x)} = \frac{P(\{X=x\} \cap \{A_2)}{P(A_2)} = \frac{P(\{X=x\} \cap \{X>1\})}{P(A_2)}$$

$$= \frac{(1-p)^{x-1} \cdot p}{P(A_2)} = \frac{(1-p)^{x-1} p}{1-p} = \begin{cases} (1-p)^{x-2} \cdot p, & \text{for } x>1 \\ 0, & \text{otherwise i.e. for } x=1 \end{cases}$$

$$= \cancel{(1-p)p + (1-p)^2 p + (1-p)^3 p + \ldots}$$

$$(1-p)$$

$$= p + (1-p)p + (1-p)^2 p + (1-p)^3 p + \ldots + (1-p)^4 p + \ldots =$$

$\cdot(1-p)$ $\quad \cdot(1-p)$ $\quad q = 1-p$

$$= p + q \cdot p + q^2 p + q^3 p + q^4 p = p(1 + q + q^2 + q^3 + \ldots +) =$$

$$= p \cdot \frac{1}{1-q} = \frac{p}{1-q} = \frac{p}{1-(1-p)} = \frac{p}{p} = 1$$

(11)

$$E[X] = E[X|A_1]\, P(A_1) + E[X|A_2] \cdot P(A_2)$$

$$E[X|A_1] = \sum_{x=1}^{\infty} x\, p_{X|A_1}(x) = \overset{x}{1} \cdot p_{X|A_1}(1) + \sum_{x=2}^{\infty} x \cdot 0 = 1 \cdot 1 + 0 = 1$$

$$E[X|A_2] = \sum_{x=1}^{\infty} x \cdot p_{X|A_2}(x) = 1 \cdot 0 + \sum_{x=2}^{\infty} x\, p_{X|A_2}(x) =$$

$$= \underbrace{\{x=1\} \cdot 0}_{1 \cdot 0} + \sum_{x=2}^{\infty} x \cdot (1-p)^{x-2} \cdot p \overset{x'=x-1}{=} \sum_{x'=1} (x'+1) \cdot (1-p)^{x'-1} p =$$

$$x = x'+1$$

$$= \sum_{x'=1}^{\infty} x' (1-p)^{x'-1} p + \underbrace{\sum_{x'=1}^{\infty} (1-p)^{x'-1} p}_{\sum_x p_X(x) = 1} = 1 =$$

$$= E[X] \sum_{x=1}^{\infty} x (1-p)^{x} p + 1 = E[X] + 1$$

Hence $E[X] = 1 \cdot p(A_1) + (E[X]+1) \cdot p(A_2)$

$$= 1 \cdot p + (E[X]+1) \cdot (1-p) =$$

$$= \cancel{p} - p\,E[X] - \cancel{p} + E[X] + 1 =$$

$$E[X] = 1 + (1-p) \cdot E[X] \Rightarrow E[X](\cancel{1} + p - \cancel{1}) = 1 \Rightarrow$$

$$\therefore \boxed{E[X] = \frac{1}{p}}$$

(12)

$$\sum_{x=2}^{\infty} f(x) = \overset{2}{\downarrow}f(2) + \overset{3}{\downarrow}f(3) + \overset{4}{\downarrow}f(4) + \overset{5}{\downarrow}f(5) + \ldots$$

$x' = 1 \quad 2 \quad 3 \quad 4$

$x = x'+1 = 2 \quad 3 \quad 4 \quad 5$

$$\sum_{x'=1}^{\infty} f(x') = f(2) + f(3) + f(4) + f(5) + \ldots$$

we want

$$= \sum_{x'=1}^{\infty} f(x'+1) \qquad \begin{matrix} 1 & 2 & 3 & 4 & \ldots \\ \downarrow & \downarrow & \downarrow & \downarrow & \\ 2 & 3 & 4 & 5 & \ldots \end{matrix}$$

$$Var(Z) = E[(Z - E[Z])^2]$$

$$\sigma_Z = \sqrt{Var(Z)}$$

$$Var(X+Y) = E[(X+Y-E[X+Y])^2] =$$

$$= E[(X+Y-E[X]-E[Y])^2] = E[(X-E[X]+Y-E[Y])^2]$$

$\tilde{X} = X - E[X]$

$\tilde{Y} = Y - E[Y]$

$$E[(\tilde{X}+\tilde{Y})^2] = E[\tilde{X}^2 + 2\tilde{X}\tilde{Y} + \tilde{Y}^2] =$$

$$= E[\tilde{X}^2] + 2E[\tilde{X}\cdot\tilde{Y}] + E[\tilde{Y}^2] =$$

$$= E[(X-E[X])^2] + 2E[(X-E[X])(Y-E[Y])] + E[\tilde{Y}^2] =$$

$$= Var(X) + 2E[X\cdot E[Y] - X\cdot E[Y] - E[X]\cdot Y + E[X]E[Y]]$$

$$= Var(X) + 2E[X]E[Y] - 2E[X]E[Y] - 2E[X]E[Y] + 2E[X]E[Y] + Var(Y)$$

$$+ Var(Y) \qquad (E[E[Y]] = E[Y])$$

$$Var(X+Y) \overset{=}{=} Var(X) + Var(Y)$$

$$E[X] = \int_{-\infty}^{+\infty} x\, p_X(x)\, dx$$

$$\boxed{E[X|Y=y] = \int_{-\infty}^{+\infty} x\, p_{X|Y}(x|y)\, dx}$$

$$\boxed{E[g(X)|Y=y] = \int_{-\infty}^{+\infty} g(x)\, p_{X|Y}(x|y)\, dx}$$

small x

$$E[X|A] = \int_{-\infty}^{+\infty} x\, f_{X|A}(x)\, dx \qquad f() \text{ is pdf}$$

$$\boxed{E[X] = \sum_{i=1}^{M} E[X|A_i]\, P(A_i)}$$

$$\boxed{E[g(X)] = \sum_{i=1}^{M} E[g(X)|A_i] \cdot P(A_i)}$$

conditions

joint

$$f_{X,Y}(x,y) = f_{X|Y}(x|y) \cdot f_Y(y)$$

$$\boxed{f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\, dy} = \boxed{\int_{-\infty}^{+\infty} f_{X|Y}(x|y) \cdot f_Y(y)\, dy}$$

$$= \int_{-\infty}^{+\infty} f_{Y|X}(y|x) f_X(x)\, dx$$

Let $g = f_{X|Y}(x|y)$, then

$$\int_{-\infty}^{+\infty} \underbrace{f_{X|Y}(x|y)}_{g} f_Y(y)\, dy = \underset{Y\sim p}{E}[f_{X|Y}(x|y)] = E[f_{Y|X}(y|x)]$$

$$\boxed{f_X(x) = E_Y[f_{X|Y}(x|y)]}$$

$$p_x(x) = \int_y p(x,y)\,dy = \int_y \underbrace{p_{X|Y}(x|y)}_{g()} \cdot p_Y(y)\,dy = E_Y\left[p_{X|Y}(x|y)\right]$$

$$\boxed{p_X(x) = E_Y\left[p_{X|Y}(x|y)\right]}$$

$$\boxed{p_X(x) = E_{Y\sim p(y)}\left[p_{X|Y}(x|y)\right]}$$

$$\Downarrow \; y = \theta$$

$$\boxed{p_X(x) = E_{\theta\sim p(\theta)}\left[p_{X|\theta}(x|\theta)\right]}$$

θ

$p(x|\theta)$ $\quad$ $P(\theta|x)$

$y = \theta$

$$p(x) = E_\theta\left[p_{X|\theta}(x|\theta)\right]$$

$$\boxed{p(\theta) = E_X\left[P_{\theta|X}(\theta|x)\right]}$$

$$\boxed{p(\theta) = E_{x\sim p(\theta|y)}\left[p_{\theta|x}(\theta|x)\right]}$$

$$p(\theta) = E_{X\sim p(\theta|x)}\left[p_\theta\right]$$

marginal of theta is expeted value of x

$$\begin{cases} x = \theta \\ y = x \end{cases}$$

$$\boxed{P_\theta(\theta) = E_{y\sim p(y)}\left[P_{\theta|y}(\theta|y)\right]}$$

If discret

$$P_\theta(\theta) = \sum_y P_Y(y) \cdot P_{\theta|Y}(\theta|y)$$