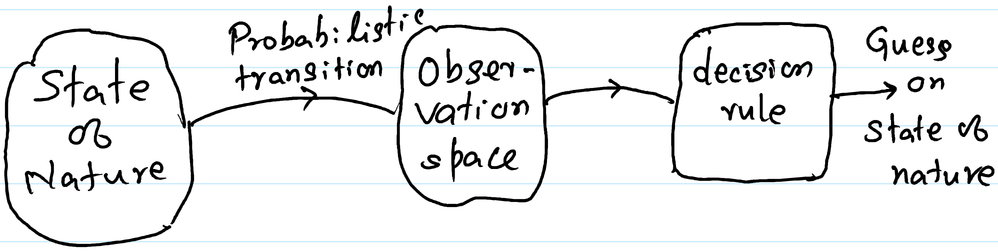
Hypothesis Testing

A general framework for solving **detection** problems



⇓
Finite number of states

(Wide Range of Applications)

- Communications
- Radar
- Medicine
- Finance
- Weather etc

To formulate & solve hypothesis testing problem we need

- Conditional cdf/pdf
- Baye's theorem
- Total Probability theorem.

Conditional pdf/cdf.

Let X be a random Variable

event (related to X)

Let X be a random ...

Let A be an event (related to X) with $P(A) > 0$

Conditional cdf (definition)

$$F_X(x|A) = \frac{P\{X \le x, A\}}{P(A)} = \Pr\{X \le x | A\}$$

Conditional pdf

$$f_X(x|A) = \frac{d}{dx} F_X(x|A)$$

Note: $\int_{-\infty}^{\infty} f_X(x|A)\, dx$

$$= F_X(\infty|A)$$

$$= \frac{P\{X \le \infty, A\}}{P(A)}$$

$$= P(A)/P(A)$$

$$= 1$$

Bayes Rule

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

$$= \Pr(X \le x | A\} \, P(A)$$

$$\Pr\{X \le x)$$

$$P(A \mid X \le x) = \frac{F_X(x \mid A)\, P(A)}{F_X(x)}$$

Now,

$$P(A \mid x < X \le x+\Delta x)$$

$$= \frac{P(x < X \le x+\Delta x \mid A)\, P(A)}{P(x < X \le x+\Delta x)}$$

$$= \frac{[F_X(x+\Delta x \mid A) - F_X(x \mid A)]\, P(A)}{F_X(x+\Delta x) - F_X(x)}$$

Taking limit as $\Delta x \to 0$

$$P(A \mid X = x) = \lim_{\Delta x \to 0} \frac{\dfrac{F_X(x+\Delta x \mid A) - F_X(x \mid A)}{\Delta x}\, P(A)}{\dfrac{F_X(x+\Delta x) - F_X(x)}{\Delta x}}$$

$$P(A \mid X = x) = \frac{f_X(x \mid A)\, P(A)}{f_X(x)}$$

$$\Delta x$$

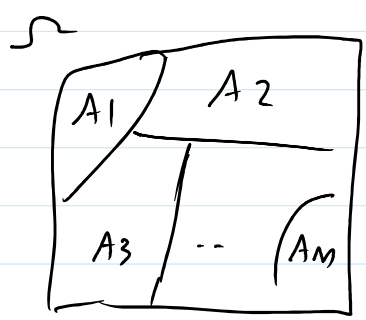
$$P(A|X=x) = \frac{f_X(x|A)\,P(A)}{f_X(x)}$$

Total Probability Theorem

Suppose $A_1, A_2, \ldots, A_M$ are M events such that

$$A_i \cap A_j = \phi \quad \text{if} \quad i \neq j$$

$$\bigcup_{i=1}^{M} A_i = \Omega \quad \text{(sample space)}$$



Now,

$$F_X(x) = P_r\{X \leq x\}$$

$$= P_r\{X \leq x, \Omega\}$$

$$= P(X \leq x, \bigcup_{i=1}^{M} A_i)$$

union & intersection distribute

$$= P(\bigcup^{M} (A_i, X \leq x))$$

intersection distribute $= P\left(\bigcup_{i=1}^{M} (A_i, X \le x)\right)$

Since A_i's are disjoint $\rightarrow$ $= \sum_{i=1}^{M} P(A_i, X \le x)$

$$= \sum_{i=1}^{M} P(X \le x | A_i) P(A_i)$$

$$F_X(x) = \sum_{i=1}^{M} F_X(x|A_i)\, P(A_i)$$

Similarly

$$f_X(x) = \sum_{i=1}^{M} f_X(x|A_i)\, P(A_i)$$

$\cancel{x}$ ———— x

Hypothesis testing

State of Nature

- $\rightarrow$ say we have M states
- $\rightarrow$ denote as $H_1, H_2, \ldots, H_M$
- $\rightarrow$ Prior information

$1 \le i \le M$

$$Pr\{H_i \text{ is true state}\} = P(H_i) = \pi_i$$

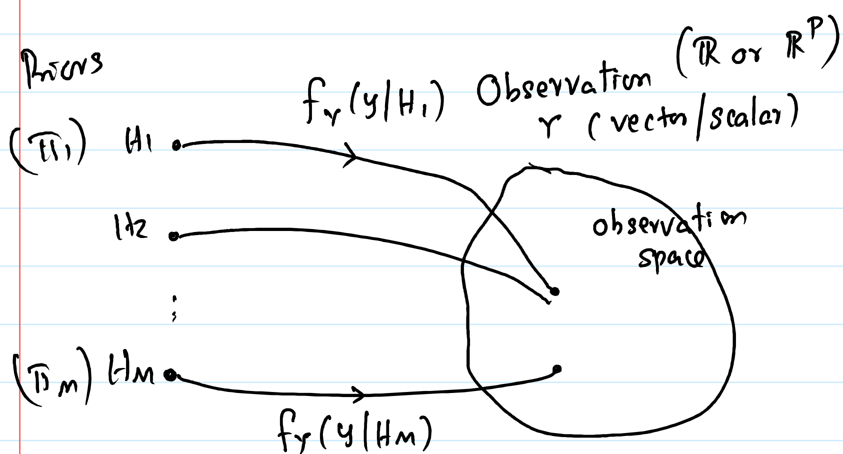
π_i's are known

<u>Observation space</u> ($\mathbb{R}$ or $\mathbb{R}^P$)

→ Measurements / observations are noisy

→ observation space contains <u>all</u> the possible (y) measurements we can get

→ observations are related to state of nature H_i thru conditional cdf $f(y|H_i)$



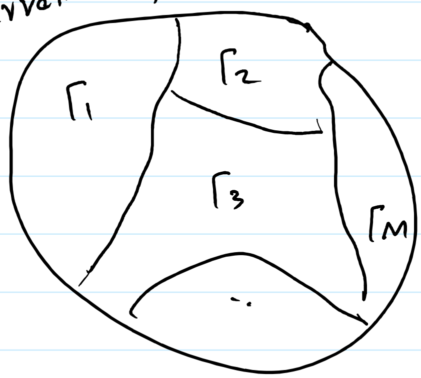
→ $f_r(y|H_i)$ known

<u>Decision Rule</u>

- For each outcome y in observation space, pick a guess on the state of nature.

observation space

observation space



$$\delta(y) = \begin{cases} H_1 & \text{if } y \in \Gamma_1 \\ H_2 & \text{if } y \in \Gamma_2 \\ \vdots & \\ H_M & \text{if } y \in \Gamma_M \end{cases}$$

- Decision rule partitions the observation space into disjoint regions $\Gamma_1, \Gamma_2, \dots, \Gamma_M$

Problem: Given π_i, $f_Y(y|H_i)$,

chose the decision rule $(\Gamma_1, \Gamma_2, \dots, \Gamma_M)$

in an optimal way

(For instance, minimize the chance of making an error)