Probability vs. Statistics

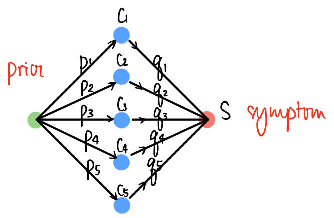
- Prob: model, axiom, clean analysis
- Stat: data-driven

2 Classes of Statistics:

1/ Bayesian - treat unknowns as RVs w/ known distributions & priors

2/ Frequentist - treat unknowns as parameters to be estimated

MLE/MAP



Detection & Bayes' Theorem

- N possible causes
- each cause has a prior p_i & a prob q_i of generating the observed symptom

π_i = posterior of cause i given symptom S

$\pi_i = P(C_i | S)$

$$= \frac{P(S|C_i)P(C_i)}{\sum_j P(S|C_j)P(C_j)} = \frac{p_i q_i}{\sum_j p_j q_j}$$

MAP: Maximum A Posteriori

$\arg\max_i \pi_i = \arg\max_i p_i q_i$

$\text{MAP}[X | Y=y] = \arg\max_x P(X=x | Y=y)$.

↳ Which cause best explains the observed symptom?

MLE: Maximum Likelihood Estimation

$\arg\max_i q_i$ = MAP under uniform prior

$\text{MLE}[X | Y=y] = \arg\max_x P(Y=y | X=x)$

↳ Which cause best generates the observed symptom?

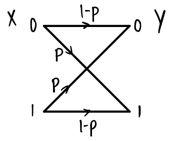
B-T 8.1: 4 Versions of Bayes

Observed Symptom (X) \ Cause (θ)	Dis	Cont
Dis	BSC $\frac{P(\theta)P(X|\theta)}{\sum P(\theta')P(X|\theta')}$	Finding bias of a coin
Cont	AWGN $\frac{P(\theta)f(X|\theta)}{\underbrace{\sum P(\theta')f(X|\theta')}_{f_X(x)}}$	Romeo & Juliet $\frac{f(\theta)f(X|\theta)}{\underbrace{\sum f(\theta')f(X|\theta')}_{f_X(x)}}$

detection/classification

estimation/regression

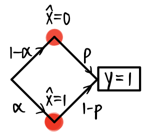
Ex 1: BSC



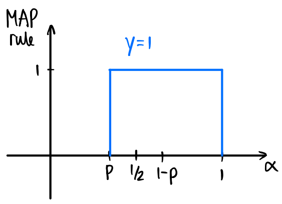
MAP: $\underset{i\in\{0,1\}}{\operatorname{argmax}}\ p_i q_i$

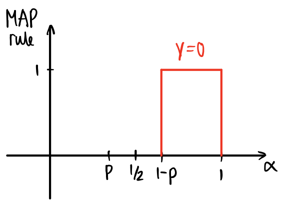
Notation:
$\boxed{p_0 q_0 \underset{\hat{x}(1)=1}{\overset{\hat{x}(1)=0}{\gtrless}} p_1 q_1}$

$y=1 \rightarrow (1-\alpha)p \underset{1}{\overset{0}{\gtrless}} \alpha(1-p) \Rightarrow \boxed{p \underset{1}{\overset{0}{\gtrless}} \alpha}$



$y=0 \rightarrow (1-\alpha)(1-p) \underset{1}{\overset{0}{\gtrless}} \alpha p \Rightarrow \boxed{1-p \underset{1}{\overset{0}{\gtrless}} \alpha}$

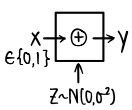


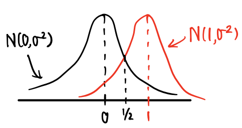


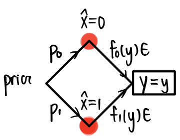
MLE for BSC(p) w/ $p<\frac{1}{2}$:

$\left.\begin{array}{l} \text{MLE}[X|Y=0]=0 \\ \text{MLE}[X|Y=0]=1 \end{array}\right\} \text{MLE}[X|Y]=X$

Ex 2: Additive White Gaussian Noise (AWGN) channel







$Y\in\{y, y+\varepsilon\}$

MAP: $p_0 f_0(y)\cancel{\varepsilon} \underset{1}{\overset{0}{\gtrless}} p_1 f_1(y)\cancel{\varepsilon}$

$\underbrace{\frac{f_1(y)}{f_0(y)}}_{L(y)\ \leftarrow \text{ likelihood}} \underset{1}{\overset{0}{\lessgtr}} \frac{p_0}{p_1}$

$$L(y) = \frac{\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(y-1)^2}{2\sigma^2}\right\}}{\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{y^2}{2\sigma^2}\right\}} = \exp\left\{\frac{2y-1}{2\sigma^2}\right\}$$

$LL(y) = \log L(y)$ ← log likelihood

$LL(y) \underset{1}{\overset{0}{\lessgtr}} \log\left(\frac{p_0}{p_1}\right)$

$\frac{2y-1}{2\sigma^2} \underset{1}{\overset{0}{\lessgtr}} \log\left(\frac{p_0}{p_1}\right)$

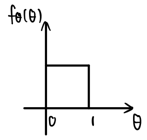
$\Rightarrow \boxed{y \underset{1}{\overset{0}{\lessgtr}} \frac{1}{2} + \sigma^2 \log\left(\frac{p_0}{p_1}\right)}$

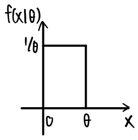
e.g. $\frac{p_0}{p_1} = e$

$\sigma^2 = 0.1$

↳ $y \underset{1}{\overset{0}{\lessgtr}} 0.5 + 0.1\log e = 0.6$

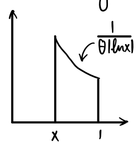
Ex 3: Romeo & Juliet (B-T)



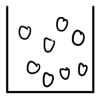


$$f_{X|\Theta}(x|\theta) = \begin{cases} 1/\theta, & 0 \le x \le \theta \le 1 \\ 0, & \text{else} \end{cases}$$

$f(\theta|x) \leftarrow$ Bayes



German Tank (GT) Problem



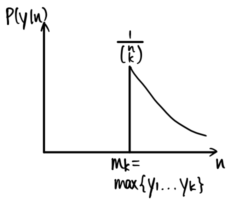
Draw a ball randomly from bin (without replacement) from 1 to N, where N is a rand int.

$MLE[N|Y=m] = m$

e.g. $m=7 \Rightarrow MLE[N|Y=7] = P(\text{observe a } 7 \mid N=n)$

$$= \begin{cases} 0 & \text{if } n<7 \\ 1/n & \text{if } n \ge 7 \text{ (chooses max)} \end{cases}$$

MLE of $N=n$ given k observations



MLE is $m_k = \max\{y_1, \ldots, y_k\}$