Last lecture

$\arg\max_{\theta} P_\theta(X)$

Maximum Likelihood Estimation (MLE) (Ch 2.8)

- Not focus

Markov and Chebychev inequalities (Ch 2.9) $P\{X \geq c\} \leq \frac{E[X]}{c}$

- Markov inequality $\rightarrow X$
- Chebychev inequality $\rightarrow P\{|Y - \mu_Y| \geq a\sigma\} \leq \frac{1}{a^2}$
- Confidence interval

$\hookrightarrow P\{p \in [\hat{p} \pm \frac{a\sigma}{\sqrt{n}}]\} \geq 1 - \frac{1}{a^2}$

Binary Hypothesis Testing (Ch 2.11)

- Definition
- Likelihood table
- Maximum likelihood decision rule

Agenda

Binary Hypothesis Testing (Ch 2.11)

- Likelihood Ratio Test (LRT)
- Maximum likelihood decision rule
- Maximum A Posteriori (MAP) decision rule
- Examples

Union Bound/ Reliability (Ch 2.12)

Hypothesis Testing

Maximum Likelihood Table

- Table showing likelihood of two hypotheses $P(X|H_i)$

patient cares. $P(H|X)$

$P(X\|H)$	$X = 0$	$X = 1$	$X = 2$	$X = 3$
H_1	0	0.1	$\underline{0.3}$	$\underline{0.6}$
H_0	$\underline{0.4}$	$\underline{0.3}$	0.2	0.1

Claim H_0 if $X=0$.

- Decision rule can be shown on the table by underscore each column

False alarm and missing

	$X = 0$	$X = 1$	$X = 2$	$X = 3$
H_1	0	<u>0.1</u>	<u>0.3</u>	<u>0.6</u>
H_0	<u>0.4</u>	0.3	0.2	0.1

- $P_{false\ alarm} =$ $P(\text{Claim } H_1 \mid H_0) = \sum_{C \in H_0,\ \text{none underscored}} C = 0.6$
- $P_{miss} =$ $P(\text{Claim } H_0 \mid H_1) = \sum_{C \in H_1,\ \text{non underscored}} C = 0$

	$X = 0$	$X = 1$	$X = 2$	$X = 3$
H_1	0	0.1	<u>0.3</u>	<u>0.6</u>
H_0	<u>0.4</u>	<u>0.3</u>	0.2	0.1

Maximum Likelihood (ML) decision rule

Pick whichever is higher per column!

$P_i(k) = P(X = k \mid H_i)$

	$X = 0$	$X = 1$	$X = 2$	$X = 3$
H_1	$p_1(0) = 0$	0.1	0.3	0.6
H_0	0.4	0.3	0.2	0.1
Λ	$\frac{0}{0.4} = 0$	$\frac{0.1}{0.3} = \frac{1}{3}$	$\frac{0.3}{0.2} = 1.5$	$\frac{.6}{.1} = 6$

Likelihood Ratio Test (LRT) $\Lambda(k) = \dfrac{p_1(k)}{p_0(k)}$

A LRT with threshold τ:

$\Lambda(k) \geq \gamma \Rightarrow$ Claim H_1

$<$ Claim H_0

Maximum Likelihood (ML) decision rule

Pick whichever is higher per column!

$= \text{LRT w/ } \gamma = 1$

	$X = 0$	$X = 1$	$X = 2$	$X = 3$
H_1	0	0.1	0.3	0.6
H_0	0.4	0.3	0.2	0.1

What's the problem?

Good balance of $P_{\text{false alarm}}$ & P_{miss}

But

1. If we want to care $P_{\text{false alarm}}$ more?

2. Minimize total error counts?

Slidos – Midterm Review and Early Feedback

Midterm Review
- Next Thursday class time
- Vote for the contents!

5831917

Early Feedback
- Suggest improvement
- Anonymous by default

1551303

Maximum A Posteriori (MAP) decision rule

Let's pick joint probability $P(H, X)$ instead of $P(X|H)$

- Pick the higher $P(H, X)$ per column
- Lowest total error rate p_e

- But how do we get the joint probability?
- Recall conditional probability

$$P(A, B) = P(A|B)P(B)$$

$$P(\underbrace{H_1, X = k}) = P\{X = k | H_1\} \times P(H_1)$$

$P(H, X)$	$X = 0$	$X = 1$	$X = 2$	$X = 3$
H_1	0	0.02	0.06	0.12
H_0	0.32	0.24	0.16	0.08

Prior and Posterior

Σ

$$P(H_1, X = k) = P(X = k|H_1)P(H_1)$$

- $P(H_1) \triangleq \pi_1$: " Prior. ", probability assumed before observation
- $P(H_0) \triangleq \pi_0 = 1 - \pi_1$
- Bayes' rule revisited (Param/Hypothesis vs. Observation)

Likelihood. Prior

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)}$$

Posterior. ↳ Evidence.

Terminologies

Exam Safe

$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)}$$

Likelihood	$P(X\|\theta)$	Tractable, well defined	Toss $p = 0.3$ coin
Posterior	$P(\theta\|X)$	Typical goal	$x_{1:3} = \{H, T, T\}, p =?$
Prior	$P(\theta)$	Domain knowledge	# coins $p = 0.3$ in world
Evidence	$P(X)$	Approx. with large data	# H in world

Likelihood table to joint probability

Assume $\pi_1 = P(H_1) = 0.2$, $\pi_0 =$ $1 - 0.2 = 0.8$

- Decide on joint probability is same as posterior probability
- MAP rule = LRT rule with $\tau = \frac{\pi_0}{\pi_1}$ NOT $\frac{\pi_1}{\pi_0}$

$\hookrightarrow P_1(k) \times \pi_1 \geq P_0(k) \times \pi_0$ Claim H_1

$\Lambda(k) = \frac{P_1(k)}{P_0(k)} \geq \frac{\pi_0}{\pi_1}$ Claim H_1

$\times P(H)$

$P(X\|H)$	$X = 0$	$X = 1$	$X = 2$	$X = 3$
H_1	0	0.1	0.3	0.6
H_0	0.4	0.3	0.2	0.1

$\times 0.2 \rightarrow$ (row H_1), $\times 0.8 \rightarrow$ (row H_0)

$P(H, X)$	$X = 0$	$X = 1$	$X = 2$	$X = 3$
H_1	0	0.1×0.2	0.06	.12
H_0	.32	.24	.16	.08

Example

X: Draw a coin from the bag and toss it 5 times # of Heads

- Likelihood table
- Joint probability table
- Describe ML and MAP rule, compute
  - p_{false_alarm}
  - p_{miss}
  - p_e error rate $= \sum c.$

$\times \pi_0 + \times \pi_1$

$c \in$ none underscore

$\uparrow$

joint prob. or $P_i(X)\pi_i$

$H_1: p = \frac{2}{3}$ coin

$H_0: p = \frac{1}{2}$ coin

Example

$X \sim Bi(5, p)$

$P(X|H)$

	0	1	2	3	4	5
H_1	$\boxed{\left(\frac{1}{3}\right)^5}$				—	—
H_0	—	—	—	—		

$H_1 : p = \frac{2}{3}$ coin

$H_0 : p = \frac{1}{2}$ coin

$$P(X=k|H_1) = \binom{5}{k}\left(\frac{2}{3}\right)^k\left(\frac{1}{3}\right)^{5-k}$$

$$P(X=k|H_0) = \binom{5}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{5-k}$$

$$\Lambda(k) = \left(\frac{\frac{2}{3}}{\frac{1}{2}}\right)^k\left(\frac{\frac{1}{3}}{\frac{1}{2}}\right)^{1-k} = \left(\frac{4}{3}\right)^k\left(\frac{2}{3}\right)^{5-k} \geq 1$$

$k \geq 3$

≥ 4

ML

MAP

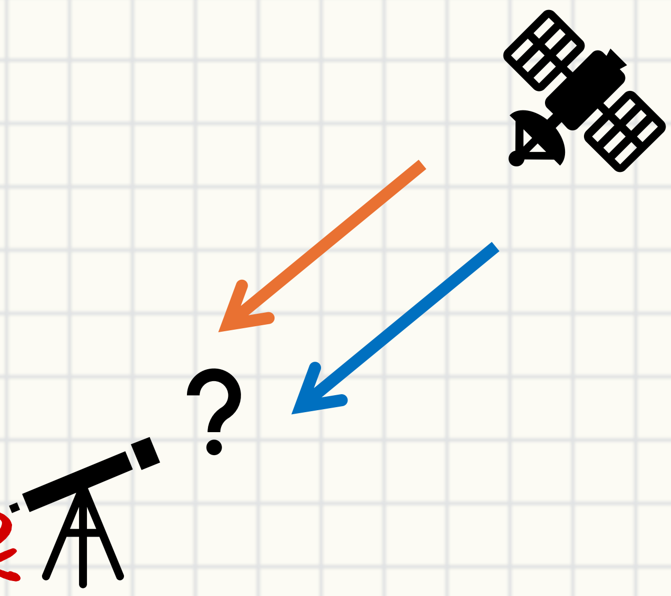
Example

4^3 2^2

3^5

Receive on-off keying (OOK) signal from a deep space Tx.

X: # of photons observe from a telescope



- $\lambda = 6$ If it's ON
- $\lambda = 2$ If it's OFF
- $\frac{\pi_0}{\pi_1} = 5$
- Describe ML and MAP rule, compute
  - p_{false_alarm}
  - p_{miss}
  - p_e

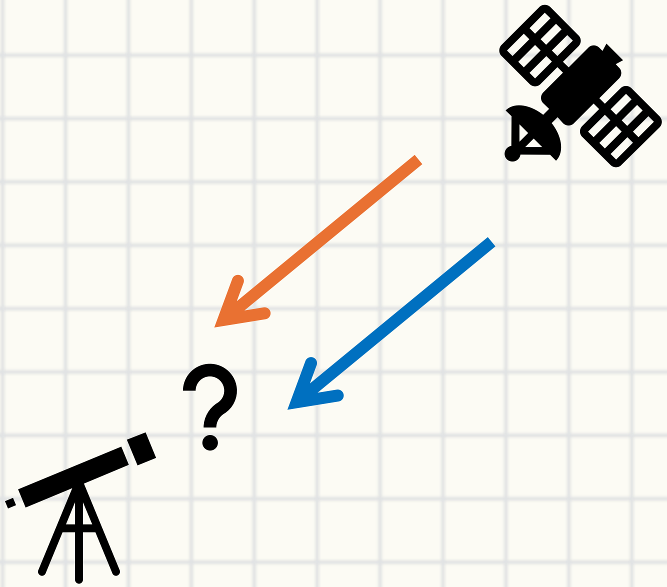
$$Poi_\lambda(k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

$H_1: Pois(\lambda = 6)$

$H_0: Pois(\lambda = 2)$

Example

$$\Lambda(k) = \frac{P_1(k)}{P_0(k)} = \frac{\dfrac{e^{-\overset{4}{\cancel{6}}}\, 6^k}{\cancel{k!}}}{\dfrac{\cancel{e^{-2}}\, 2^k}{\cancel{k!}}} = e^{-4}\, 3^k$$



$H_1: Pois(\lambda = 6)$

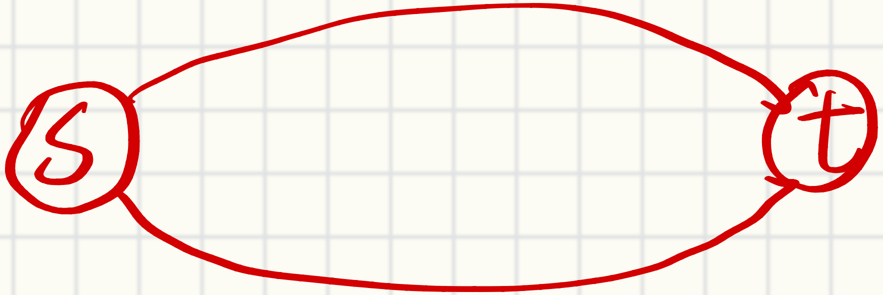
$H_0: Pois(\lambda = 2)$

ML. $k \geq k^* \Rightarrow \Lambda(k) \geq 1$ Claim H_1

MAP. $\Lambda(k) \geq \frac{\pi_0}{\pi_1} = 5,$

Union Bound/ Reliability

Skip. network flow.



Motivation and Definition

Reliability

- How likely a system will fail?
  - Each subsystem fail with small probability
  - If sub-systems fail in some pattern, the system fails

Union Bound

- Bounds for union of small probability events
- $P(A \cup B) \leq$ $P(A) + P(B)$
- $P(A_1 \cup A_2 \cup \cdots \cup A_m) \leq$ $\sum P(A_i)$
- Bound is at most 2X far from the actual value

Example – Array codes

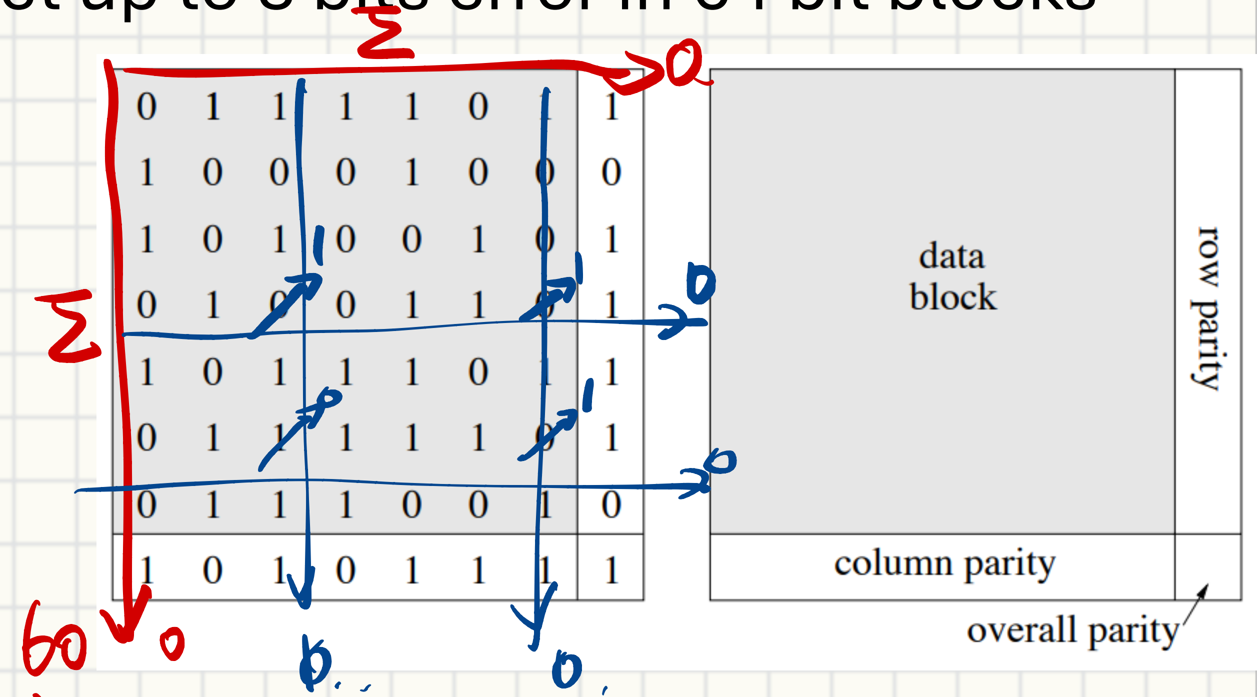
Consider a code that can detect up to 3 bits error in 64 bit blocks

- Let Y denotes # error bits
- Solve $P\{Y = 4\}$

Denote BER as p.

$$Y \sim Bi(64, p)$$

$$P\{Y=4\} = \binom{64}{4} p^4 (1-p)^{60}$$



Example – Array codes

Consider a code that can detect up to 3 bits error in 64 bit blocks

- Let Y denotes # error bits
- Bound $P\{Y \geq 4\}$

$X_i = \binom{64}{4}$ bits error.

rest bits don't care

$$P\{Y \geq 4\}$$

$$= P\{\bigcup X_i\} \leq \sum P\{X_i\}$$

$$\boxed{= \binom{64}{4} p^4}$$

Midterm #1 Until Here!