Lecture 3

Note Title 1/18/2012

* $\hat{X}_{MAP}(y)$, $\hat{X}_{ML}(y)$, $f(y)$

* Performance of a detector.
  - FA / MD prob.
  - Average error prob

① $P(f(Y) \neq X)$

② $\hat{X}_{MAP}(y)$ minimizes $P(f(Y) \neq X)$

③ Properties of $\hat{X}_{MAP}(y)$

* Chernoff Bound: (Mostly used for i.i.d. R.V.s)

① $P(X \geq d) \leq \min_{s \geq 0} \frac{E(e^{sX})}{e^{sd}}$

② $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$, for i.i.d X_i'

$$P(\bar{X} \geq d) \leq \left(\min_{s \geq 0} \frac{E(e^{sX_i})}{e^{sd}} \right)^n$$

* Application of the Chernoff Bound to hypothesis testing with i.i.d. observations

H_0: $Y_1 \cdots Y_n$ are i.i.d with $Y_i \sim P_0$

H_1: $Y_1 \cdots Y_n$ are i.i.d with $Y_i \sim P_1$

Q: Find the expression of the misdetection prob of a Log likelihood ratio test with threshold γ.

Ans: $P_1\left(\log \frac{P_0(Y_1, Y_2, \cdots, Y_n)}{P_1(Y_1, Y_2, \cdots, Y_n)} \geq \gamma \right)$

Or equivalently

$$\because \log \frac{P_0(Y_1, \cdots, Y_n)}{P_1(Y_1, \cdots, Y_n)} = \log \frac{P_0(Y_1) P_0(Y_2) \cdots P_0(Y_n)}{P_1(Y_1) P_1(Y_2) \cdots P_1(Y_n)}$$

$$= \sum_{i=1}^{n} \log \frac{P_0(Y_i)}{P_1(Y_i)} \triangleq \sum_{i=1}^{n} T(Y_i) \text{ where}$$

$$T(y) = \log \frac{P_0(y)}{P_1(y)}$$

$$\Rightarrow P_1\left(\frac{1}{n} \sum_{i=1}^{n} T(Y_i) \geq \tau \right)$$

$$\tau \triangleq \frac{\gamma}{n}$$

If $\tau = 0 \Rightarrow$ MD prob of the ML detector

If $\tau = \frac{\log(P(X=1)) - \log(P(X=0))}{n}$

$\Rightarrow$ MD prob of the MAP detector

What if $\tau = E_0(T(Y))$?

or more rigorously $\tau = E_0(T(Y)) - \varepsilon$ for some small $\varepsilon > 0$

What is the physical meaning of

$$① = P_1\left(\frac{1}{n}\sum_{i=1}^{n} T(Y_i) \geq E_0(T(Y)) - \varepsilon\right)?$$

Ans: $T(Y_i)$ are i.i.d. when H_0 is true, the sample mean $\frac{1}{n}\sum_{i=1}^{n} T(Y_i)$ concentrate around $E_0(T(Y))$

$\Rightarrow$ ① is the prob when H_1 is true but the observations $Y_1, \ldots, Y_n$ look as if H_0 is true.

What is the value of

$$P_1\left(\frac{1}{n}\sum_{i=1}^{n} T(Y_i) \geq \underbrace{E_0(T(Y))}_{\tau^*}\right)?$$

By Chernoff bound.

$$P_1\left(\frac{1}{n}\sum_{i=1}^{n} T(Y_i) \geq \tau^*\right)$$

$$\approx \left(\min_{s\geq 0} \underbrace{\frac{E_1(e^{sT(Y)})}{e^{s\tau^*}}}_{\triangleq f(s)}\right)^n$$

We need to solve

$$f'(s) = E_1\left((T(Y)-\tau^*)\, e^{s(T(Y)-\tau^*)}\right) = 0$$

Try $s=1$

$$e^{\tau^*} f'(1) = E_1\left((T(Y)-\tau^*)\, e^{T(Y)}\right)$$

$$= \int_y \left(\log\frac{P_0(y)}{P_1(y)} - \tau^*\right) e^{\log\frac{P_0(y)}{P_1(y)}} \times P_1(y)\, dy$$

$$= \int \left(\log \frac{p_0(y)}{p_1(y)} - \tau^* \right) \frac{p_0(y)}{p_1(y)} \, p_1(y) \, dy$$

$$= \int \left(\log \frac{p_0(y)}{p_1(y)} - \tau^* \right) p_0(y) \, dy$$

$$= E_0 \left(T(Y) - \tau^* \right)$$

$$= E_0(T(Y)) - \tau^* = 0$$

recall that $\tau^* \triangleq E_0(T(Y))$

One can actually show that $f(s)$ is a convex function of s. So $s^* = 1$ is the minimum. H.V. Poor pp. 86-91

$$\Rightarrow \min_{s \geq 0} f(s) = f(1)$$

$$= \frac{E_1\left(e^{T(Y)}\right)}{e^{\tau^*}}$$

$$= \frac{\int_y e^{\log \frac{p_0(y)}{p_1(y)}} \cdot p_1(y) \, dy}{e^{\tau^*}}$$

$$= \frac{\int \frac{p_0(y)}{p_1(y)} p_1(y) \, dy}{e^{\tau^*}} = e^{-\tau^*}$$

This special $\tau^* \triangleq E_0\left(\log \frac{P_0(Y)}{P_1(Y)}\right)$

is denoted as $D(P_0 \| P_1)$ the divergence between P_0 distribution & the P_1 distribution

$$D(P_0\|P_1) = E_0\left(\log \frac{P_0(Y)}{P_1(Y)}\right) = \int_y P_0(y) \log \frac{P_0(y)}{P_1(y)}\, dy$$

(Also known as the Kullback-Leibler information number.)

The intuition behind $D(P_0 \| P_1)$ is:

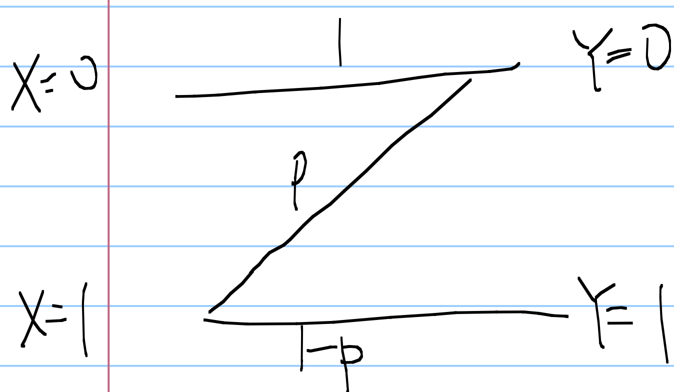
The prob that H_1 is true but the i.i.d. $Y_1, \cdots, Y_n$ look as if H_0 is true. is upper bounded by $e^{-nD(P_0\|P_1)}$

(which is asymptotically tight due to the Chernoff bound.)

$$\text{Volume} = e^{-nD(P_0 \| P_1)}$$

if H_0 is true, the most likely values of $(y_1, \cdots, y_n)$ is called the <u>typical set</u> of P_0

Example: The Z channel



If $X=0$. $P_0(y) = \begin{cases} 1 & \text{if } y=0 \\ 0 & y=1 \end{cases}$

If $X=1$ $P_1(y) = \begin{cases} p & \text{if } y=0 \\ 1-p & y=1 \end{cases}$

Q: $D(P_0 \| P_1)$ & $D(P_1 \| P_0)$

Ans: $D(P_0 \| P_1) = E_0\left(\log\left(\frac{P_0(Y)}{P_1(Y)}\right)\right)$

$$= 1 \times \log \frac{P_0(0)}{P_1(0)} + 0 = \log\left(\frac{1}{p}\right)$$

$D(P_1 \| P_0)$

$$= p \log \frac{P_1(0)}{P_0(0)} + (1-p) \log \frac{\overbrace{P_1(1)}^{=(1-p)}}{\underbrace{P_0(1)}_{=0}}$$

$$= \infty \ \#$$

Intuition :

$P(Y_1, \cdots, Y_n$ look like H_0 is true $| H_1$ is true$)$

$$\underset{(\approx)}{\leq} e^{-nD(P_0 \| P_1)} = \left(\frac{1}{p}\right)^{-n} = p^n$$

I.e. receiving n consecutive zeros even though $X=1$.

$P(Y_1, \cdots, Y_n$ look like H_1 is true $| H_0$ is true$)$

$$\underset{(\approx)}{\leq} e^{-nD(P_1 \| P_0)} = 0$$

Why it is zero? $\because$ When $X=0$, the output (with $P_0(Y=1)=0$) can never mimic (look like) $X=1$, which has $P_1(Y=1) \neq 0$

Properties of $D(P_0 \| P_1)$

① $D(P_0 \| P_1) \geq 0$

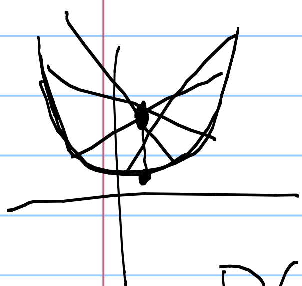
Intuition:

$e^{-nD(P_0 \| P_1)}$ is a tight prob bound $\Rightarrow$ $e^{-nD(P_0 \| P_1)} \in [0,1]$

$\Leftrightarrow$ $0 \leq D(P_0 \| P_1) \leq \infty$

pf: By Jensen's inequality:

For any convex function $f(\cdot)$



$$\underbrace{E(f(X))}_{y \text{ coordinate}} \geq f(\underbrace{E(X)}_{x \text{ coordinate}})$$

$$D(P_0 \| P_1) = E_0\left(\log \frac{P_0(Y)}{P_1(Y)}\right)$$

$$= E_0\left[-\log\left(\frac{P_1(Y)}{P_0(Y)}\right)\right]$$

$\because -\log(\cdot)$ is convex

$$\geq -\log\left(E_0\left(\frac{P_1(Y)}{P_0(Y)}\right)\right)$$

$$= -\log 1 = 0$$

② In general

$$D(P_0 \| P_1) \neq D(P_1 \| P_0)$$