Principles of Digital Communication

March 5, 2021

Reading for next week:

2.2.2 (m-ary Hypothesis testing)

2.3 (Q-function)

2.4 (Additive white Gaussian noise)

So far: two Hypotheses $H \in \{0, 1\}$

$P_H(0),\ P_H(1)$

observation $Y \sim P_{Y|H}(y|i)$

decision: $\hat{H}(Y)$

$\Pr(H = \hat{H}(Y))$ is largest when $\hat{H}(Y) = \hat{H}_{MAP}(Y)$

where $\hat{H}_{MAP}(y)$ is

$$\frac{P_{Y|H}(y|1)}{P_{Y|H}(y|0)} \underset{0}{\overset{1}{\gtrless}} \frac{P_H(0)}{P_H(1)}$$

In the coming reading:

- what happens when $H \in \{0, 1, \ldots, m-1\}$
  $P_H(0), \ldots, P_H(m-1)$?

$Y \sim P_{Y|H}(y|i)$

$$\hat{H}_{MAP}(y) = \arg\max_i P_{H|Y}(i|y)$$

maximizes $\Pr(H = \hat{H}(Y))$.

Gaussian Noise

Remember: A random variable

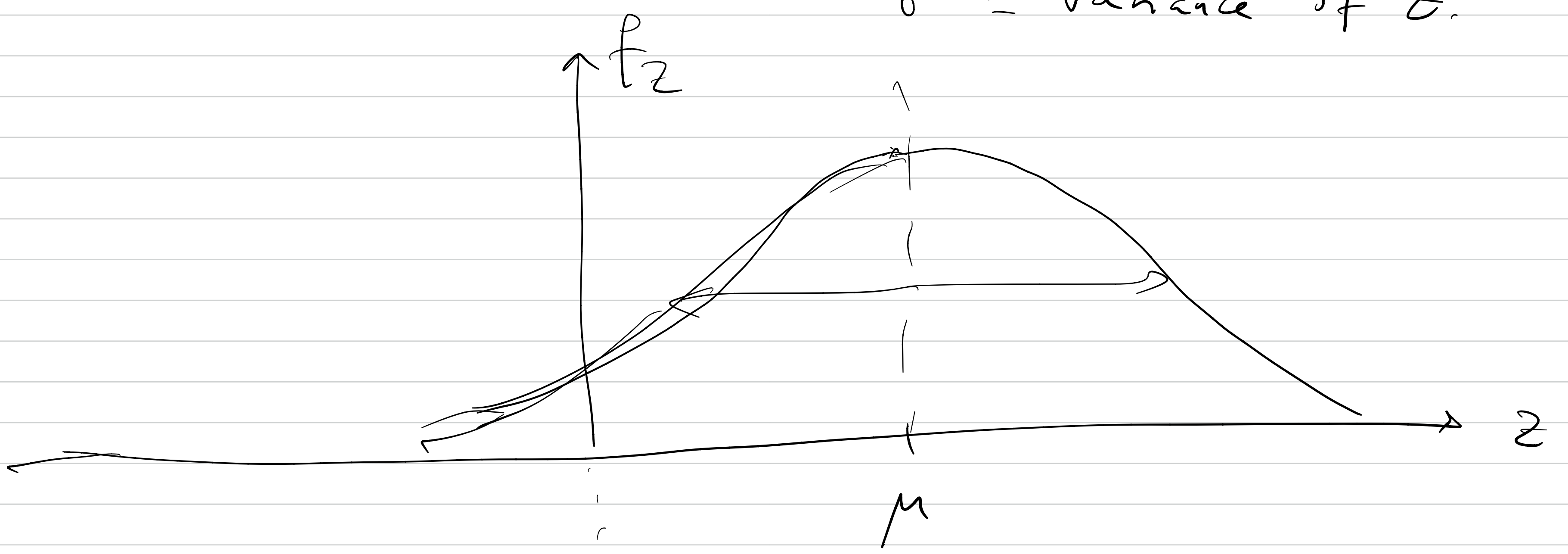
Z is said to be gaussian if its probability density function (p.d.f.) is of the following form:

$$f_Z(z) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2\sigma^2}(z-\mu)^2\right)$$

mean of Z

σ = standard deviation of Z

σ^2 = variance of Z.



Why Gaussian?

Suppose $X_1, X_2, X_3, \ldots$ are independent, identically distrib.

all with $E(X_i) = 0$, $E[X_i^2] = 1$

$S_n = \sum_{i=1}^{n} X_i$ will have $E(S) = 0$

$$E[S^2] = E\left[\left(\sum_{i=1}^{n} X_i\right)^2\right]$$

$$= E\left(\sum_i X_i \sum_j X_j\right)$$

$$= \sum_{i,j} E(X_i X_j) = n$$

$E(X_i X_j) = 0$ if $i \neq j$, 1 if $i = j$

$$S_n = \sum_{i=1}^{n} X_i \quad \text{has}$$

$$0 = E\left[\left(\frac{S_n}{\sqrt{n}}\right)\right]$$

$$1 = E\left[\left(\frac{S_n}{\sqrt{n}}\right)^2\right]$$

distribution of $\left(\frac{S_n}{\sqrt{n}}\right) \longrightarrow$ gaussian distrib.

n gets large

(convergence in distribution)

0-mean
1-variance Gaussian

$$\equiv \lim_{n\to\infty} \Pr\left(\frac{S_n}{\sqrt{n}} \le a\right) = \Pr(Z \le a)$$

$$\frac{S_n}{\sqrt{n}} = \underbrace{\sum_{i=1}^{n}}\left(\frac{X_i}{\sqrt{n}}\right)$$

n independent RVs

each is (small)

"Moral of the Central Limit Theorem"

if a Macroscopic phenomenon $(S_n/\sqrt{n})$ is the result of the sum of many microscopic phenomena $(X_i/\sqrt{n})$'s then the Macroscopic object is Gaussian.

<u>Thus</u>: It makes sense to study the case when the noise in our observation Y is Gaussian.

Let us then consider the following problem.

$H \in \{0, 1\}$

$p_H(0), \quad p_H(1).$

$Y = c_0 + Z \qquad \text{if } H=0$

$Y = c_1 + Z \qquad \text{if } H=1$

$$Z \sim \mathcal{N}(0, \sigma^2)$$

variance

0-mean

The MAP rule will do:

$$\frac{f_{Y|H}(y|1)}{f_{Y|H}(y|0)} \underset{0}{\overset{1}{\gtrless}} \eta = \frac{p_H(0)}{p_H(1)}. \qquad \text{(MAP rule)}$$

$$f_{Y|H}(y|i) = \frac{1}{\sqrt{2\pi}\cdot\sigma} \exp\left(-\frac{1}{2\sigma^2}(y-c_i)^2\right).$$

$$\text{MAP} \equiv -\frac{1}{2\sigma^2}(y-c_1)^2 + \frac{1}{2\sigma^2}(y-c_0)^2 \underset{0}{\overset{1}{\gtrless}} \log \eta$$

$$\equiv 2yc_1 - c_1^2 - 2yc_0 + c_0^2 \underset{0}{\overset{1}{\gtrless}} 2\sigma^2 \log \eta$$

$$\equiv (c_1 - c_0)y \underset{0}{\overset{1}{\gtrless}} \sigma^2 \log \eta + \frac{c_1^2 - c_0^2}{2}$$

So we get:

$$Y \underset{0}{\overset{1}{\gtrless}} \text{threshold}(\eta, c_0, c_1)$$

Suppose $c_1 > c_0$, so that we can divide by $(c_1 - c_0)$ without changing inequalities,

$$\boxed{Y \underset{0}{\overset{1}{\gtrless}} \frac{c_0 + c_1}{2} + \frac{\sigma^2}{c_1 - c_0} \log \eta}$$

$$\Pr\text{ob}(\hat{H} = 1 \mid H = 0) = \Pr\left(c_0 + Z > \frac{c_0 + c_1}{2} + \frac{\sigma^2}{c_1 - c_0} \log \eta\right)$$

$$= \Pr\left(Z > \frac{c_1 - c_0}{2} + \frac{\sigma^2}{(c_1 - c_0)} \log \eta\right)$$

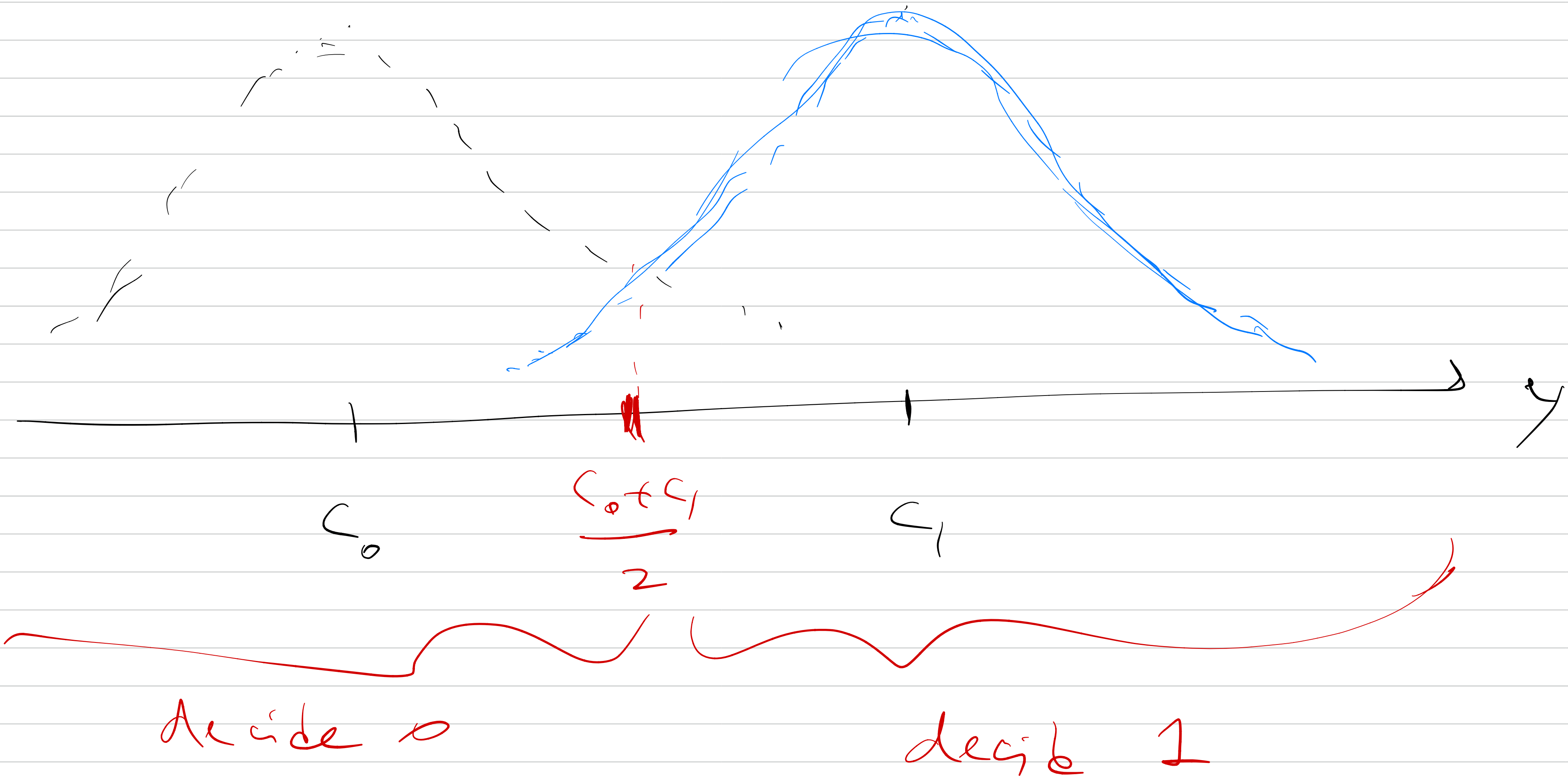
$$= \Pr\left(\underbrace{\left(\frac{Z}{\sigma}\right)}_{N(0,1)} > \frac{c_1 - c_0}{2\sigma} + \frac{\sigma}{c_1 - c_0} \log \eta\right)$$

$$\triangleq \underbrace{Q}_{\text{"Q function"}}\left(\frac{c_1 - c_0}{2\sigma} + \frac{\sigma}{c_1 - c_0} \log \eta\right)$$

Example $P_H(0) = P_H(1) = \frac{1}{2} \Rightarrow \eta = 1$

So, with $c_1 > c_0$, the decision rule is

$$y \underset{0}{\overset{1}{\gtrless}} \frac{c_1 + c_0}{2}$$



$$\Pr(\hat{H} = 1 \mid H = 0) = \Pr\left(\frac{Z}{\sigma} > \frac{c_1 - c_0}{2\sigma}\right)$$

$$= Q\left(\frac{d}{2\sigma}\right) \qquad d = |c_1 - c_0|$$

$$\Pr(\hat{H} = 0 \mid H = 1) = \text{''}$$

Vector vs Scalar observation:

$\vec{Y} \in \mathbb{R}^n$

$$H = 0 \Rightarrow \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} c_{01} \\ \vdots \\ c_{0n} \end{bmatrix} + \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{bmatrix}$$

$$H = 1 \Rightarrow \quad = \begin{bmatrix} c_{1n} \\ \vdots \\ c_{1n} \end{bmatrix} + \begin{bmatrix} Z_1 \\ \vdots \\ Z_n \end{bmatrix}$$

$(Z_1 \dots Z_n)$ are iid $N(0, \sigma^2)$.