# EECS 126: Probability & Random Processes
# Fall 2021

Digital Link

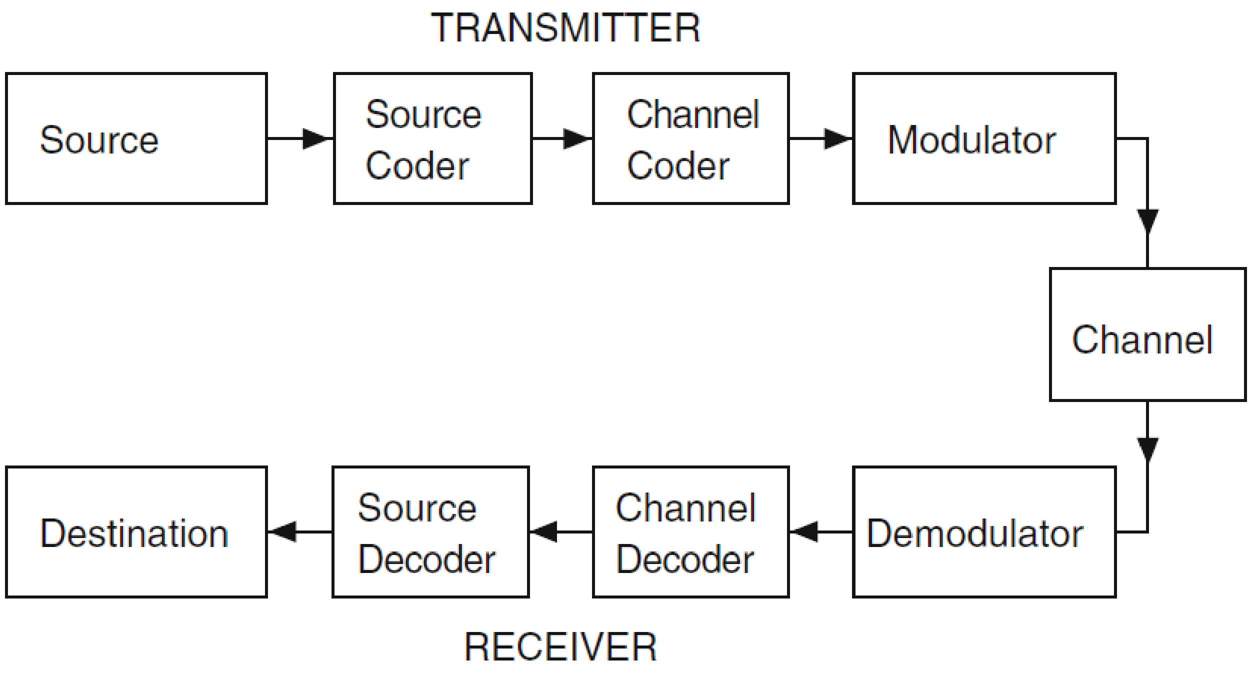
Shyam Parekh

# Topics Covered in Lectures on Digital Link

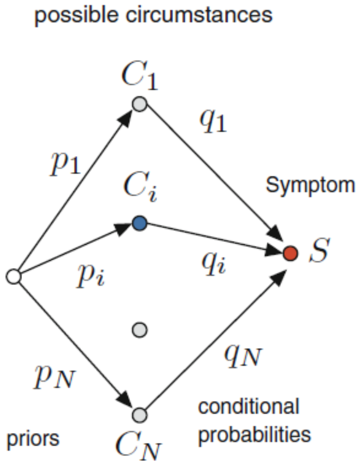
- Digital Link (Section 7.1)
- Detection and Bayes' Rule (Section 7.2)
- Huffman Codes (Section 7.3)
- Capacity of BSC (Section 15.7)
- Gaussian Channel (Section 7.4 excluding Section 7.4.1)
- Hypothesis Testing (Section 7.6)
- Proof of Neyman-Pearson Theorem (Section 8.2)
- Jointly Gaussian Random Variables (Sections 8.3)

Components

- Transfer information reliably (i.e., meeting performance requirements) using minimum resources (computation, bandwidth, storage, energy, …)
- Physical medium: phone line, cable, fiber, wireless (cellular or Wi-Fi), …



## Bayes' Rule, MAP & MLE

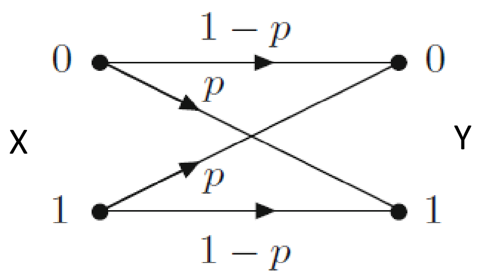



Priors: $p_i = P(C_i)$

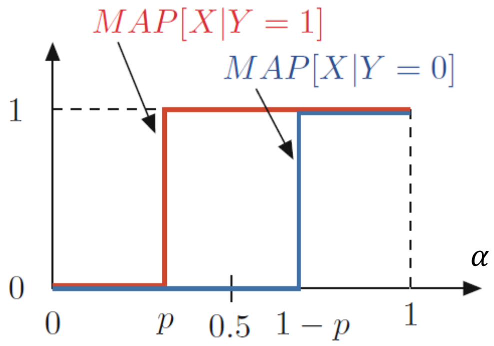
Conditionals: $q_i = P[S|C_i]$

- Theorem: Let $\pi_i = P[C_i|S]$. Then, $\pi_i = \frac{p_i q_i}{\sum_{j=1}^{N} p_j q_j}$.
- Maximum A Posteriori estimate: $MAP = \arg max_i \, P[C_i \mid S] = \arg max_i \, p_i q_i$.
- Maximum Likelihood Estimate: $MLE = \arg max_i \, P[S \mid C_i] = \arg max_i \, q_i$.
- Example: Ice Cream & Sunburn.
- General Definition: Let $X$ & $Y$ be discrete RVs.
  - $MAP[X \mid Y = y] = \arg max_x \; P(X = x, Y = y)$.
  - $MLE[X \mid Y = y] = \arg max_x \, P\,[Y = y \mid X = x]$.

Binary Symmetric Channel (BSC)

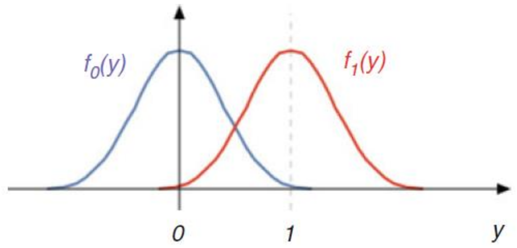



- Theorem: For BSC,
  - $MAP[X \mid Y=0] = 1\{\alpha > (1-p)\}, MAP[X \mid Y=1] = 1\{\alpha > p\}$.
  - $MLE[X \mid Y] = Y$, if $p < 0.5$.



## Gaussian Channel

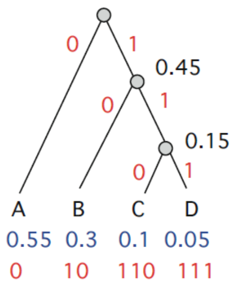
- Additive Gaussian Noise Channel: $Y = X + Z$, where $Z \equiv_D \mathcal{N}(0,\ \sigma^2)$ is independent of $X$.
  - Suppose $X \in \{0, 1\}$.
  - Let $f_0 = f_{Y|\{X=0\}}$ and $f_1 = f_{Y|\{X=1\}}$.




- Theorem: For Gaussian channel,
  - $MAP[X \mid Y = y] = 1\left\{y \geq \frac{1}{2} + \sigma^2 log_e\left(\frac{p_0}{p_1}\right)\right\}$.
  - $MLE[X \mid Y = y] = 1\{y \geq 0.5\}$.
  - Corollary: For MLE, probability of error $p = P\left(\mathcal{N}(0,1) > \frac{0.5}{\sigma}\right)$ is same as that for a BSC.

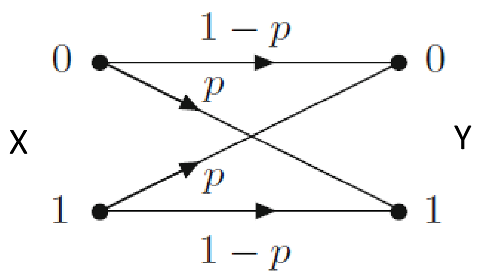
# Huffman Codes

- Consider coding of 4 symbols A, B, C, D.
  - One option: Assign 00, 01, 10, 11 to A, B, C, D, respectively.
  - We can decode a received string (without any errors) unambiguously: 0100110001 → BADAB
  - 2 bits per symbol.
- Suppose, we know probability of occurrence of each symbol: 0.55, 0.3, 0.1, 0.05, respectively.
  - Let's assign 0, 10, 110, 111 to A, B, C, D, respectively.
    - Observe we are assigning shorter codes to more frequent symbols.
  - Here also we can unambiguously decode in one pass since the codes are prefix-free: 110100111 → CBAD.
  - Average # of bits per symbol = 1.6 (20% saving in transmission).
  - These are called Huffman Codes.
- Construction:

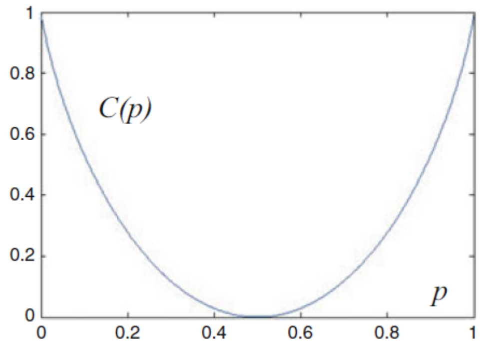



- Theorem: The Huffman Code has the smallest average number of bits per symbol among all prefix-free codes.

# Capacity of Binary Symmetric Channel (BSC)




- Theorem: Capacity of BSC $:= C(p) = 1 - H(p)$,
  where entropy $H(p) = -plogp - (1-p)log(1-p)$.



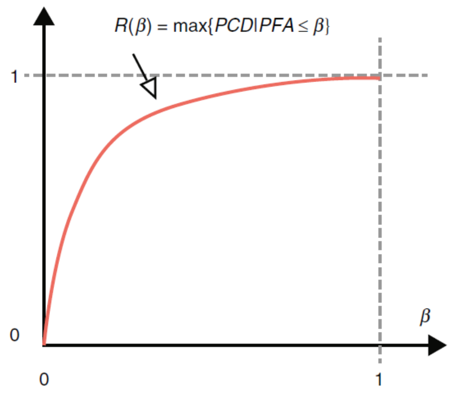
## Hypothesis Testing

- Formulation: $X \in \{0, 1\}$ and $P[Y \mid X]$ is known. Let $\hat{X}$ be the estimate of $X$.

  Maximize Probability of Correct Detection (PCD) $:= P\left[\hat{X} = 1 \mid X = 1\right]$

  Subject to Probability of False Alarm (PFA) $:= P\left[\hat{X} = 1 \mid X = 0\right] \leq \beta$.

- Receiver Operating Characteristic (ROC): PCD for the solution of the above problem as a function $R(\beta)$.




- Solution is provided by the following theorem.

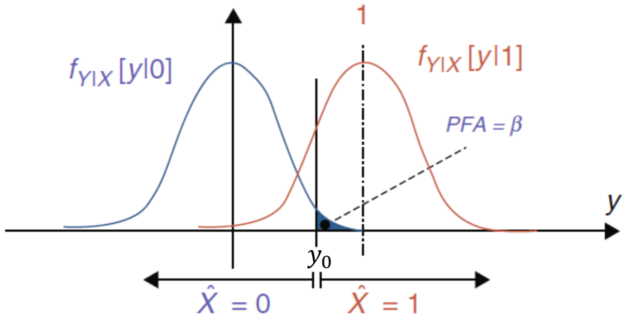
- Neyman-Pearson Theorem:

$$\hat{X} = \begin{cases} 1, if\ L(Y) > \lambda \\ 1\ w.p.\gamma, if\ L(Y) = \lambda \\ 0, if\ L(Y) < \lambda \end{cases}$$

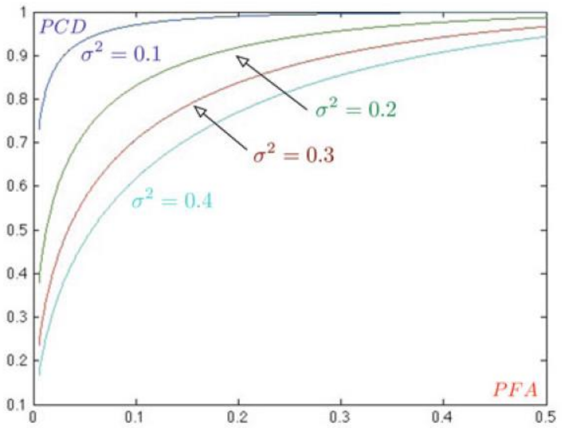
  where the likelihood ratio $L(y) = \frac{f_{Y|X}[y \mid 1]}{f_{Y|X}[y \mid 0]}$ and $\lambda, \gamma$ are chosen s. t.
  $P\left[\hat{X} = 1 \mid X = 0\right] = \beta$.

## Hypothesis Testing – Example (1)

- Gaussian Channel: $Y = X + Z$, where $Z \equiv_D \mathcal{N}(0,\ \sigma^2)$ is independent of $X$.
  - Receiver wants to guess $X$ from the received signal $Y$ with $PFA \leq \beta$.
- Solution of the hypothesis problem:




- ROC:



# Hypothesis Testing – Example (2)

- Mean of Exponential RVs: A machine produces lightbulbs with IID $Exp(\lambda_x)$ lifespans, where $x \in \{0, 1\}$, $\lambda_0 < \lambda_1$, and $x = 1$ indicates a defective machine.
  - By observing the lifespans of $n$ lightbulb, we want to detect if the machine is defective with $PFA \leq \beta$.

# Hypothesis Testing – Example (3)

- Bias of a coin: A coin is $B(p_x)$, where $x \in \{0, 1\}$ with $p_1 > p_0 = 0.5$
  - By observing $n$ coin flips, we want to detect if it's a biased coin with $PFA \leq \beta$.

# Hypothesis Testing – Example (4)

- Discrete observations: $X \in \{0, 1\}$ and $Y \in \{A, B, C\}$.
  - Given $P[Y \mid X]$, guess $X$.
  - $P[Y \mid X]$:

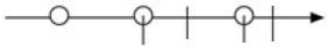
| | Y=A | Y=B | Y=C |
|---|---|---|---|
| X=0 | 0.2 | 0.5 | 0.3 |
| X=1 | 0.2 | 0.2 | 0.6 |

  - $L(Y)$:

| | | |
|---|---|---|
| 1.0 | 0.4 | 2.0 |

- Solution:

| $Y$ | $B$ | $A$ | $C$ |
|---|---|---|---|
| $P[Y\|X=1]$ | 0.2 | 0.2 | 0.6 |
| $P[Y\|X=0]$ | 0.5 | 0.2 | 0.3 |
| $L(Y)$ | 0.4 | 1 | 2 |



$\lambda = 2.1 \Rightarrow PCD = 0, PFA = 0$

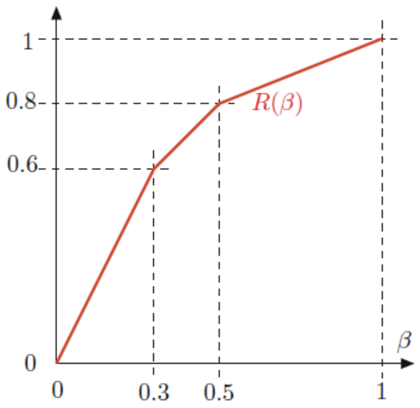
$\lambda = 2 \Rightarrow PCD = 0.6\gamma, PFA = 0.3\gamma$

$\lambda = 1.4 \Rightarrow PCD = 0.6, PFA = 0.3$

$\lambda = 1 \Rightarrow PCD = 0.6 + 0.2\gamma, PFA = 0.3 + 0.2\gamma$

$\lambda = 0.4 \Rightarrow PCD = 0.8 + 0.2\gamma, PFA = 0.5 + 0.5\gamma$

- ROC:



## Recall from Basic Probability, Section B.7

### Density of a Function of RVs

- Suppose a RV $X$ has PDF $f_X(\mathrm{x})$ and $Y = aX + b$. Then,

  $f_Y(y) = \frac{1}{|a|} f_X(\mathrm{x})$ where $ax + b = y$.

- Suppose a random vector $\boldsymbol{X}$ has JPDF $f_{\boldsymbol{X}}(\boldsymbol{x})$, and $\boldsymbol{Y} = A\boldsymbol{X} + \boldsymbol{b}$ where $A$ is a nonsingular matrix. Then,

  $f_{\boldsymbol{Y}}(\boldsymbol{y}) = \frac{1}{|A|} f_{\boldsymbol{X}}(\mathbf{x})$ where $A\boldsymbol{x} + \boldsymbol{b} = \boldsymbol{y}$, and $|A|$ is the absolute value of the determinant of $A$.

- Suppose $\boldsymbol{Y} = g(\boldsymbol{X})$ where $\boldsymbol{X}$ has density $f_{\boldsymbol{X}}(\boldsymbol{x})$. Then,

  $f_{\boldsymbol{Y}}(\boldsymbol{y}) = \sum_i \frac{1}{|J(\boldsymbol{x}_i)|} f_{\boldsymbol{X}}(\boldsymbol{x}_i)$ where the sum is over all $\boldsymbol{x}_i$ such that $g(\boldsymbol{x}_i) = \boldsymbol{y}$ and $|J(\boldsymbol{x}_i)|$ is the absolute value of Jacobian evaluated at $\boldsymbol{x}_i$.

  - Recall $J_{i,j}(\mathbf{x}) = \frac{\partial}{\partial x_j} g_i(\boldsymbol{x})$.

## Jointly Gaussian (JG) RVs

- **Definition:** $Y_1, \dots, Y_n$ are JG RVs if $\boldsymbol{Y} := \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} = A\boldsymbol{Z} + \mu_{\boldsymbol{Y}}$, where $Z_i$'s in $\boldsymbol{Z} := \begin{bmatrix} Z_1 \\ \vdots \\ Z_k \end{bmatrix}$ are IID $\mathcal{N}(0,1)$, and $A$ and $\mu_{\boldsymbol{Y}}$ are $n \times k$ and $n \times 1$ constant matrices, respectively.
  - Mean and covariance of $\boldsymbol{Y}$ are $\mu_{\boldsymbol{Y}}$ and $\Sigma_{\boldsymbol{Y}} := AA'$.
  - We write $Y \equiv_D \mathcal{N}(\mu_{\boldsymbol{Y}}, \Sigma_{\boldsymbol{Y}})$.
- **Theorem:** Assuming $\Sigma_{\boldsymbol{Y}}$ is invertible, $f_{\boldsymbol{Y}}(\boldsymbol{y}) = \frac{1}{\sqrt{|\Sigma_{\boldsymbol{Y}}|(2\pi)^{n/2}}} exp\{-\frac{1}{2}(\boldsymbol{y} - \mu_{\boldsymbol{Y}})'\Sigma_{\boldsymbol{Y}}^{-1}(\boldsymbol{y} - \mu_{\boldsymbol{Y}})\}$.
- **Example:** IID $\mathcal{N}(0, \sigma_i^2)$ RVs
- **Fact:** Level curves for the JG PDF are elliptical.
- **Example:** Marginals are Gaussian, but not JG.
- **Theorem:** JG RVs are independent if and only if they are uncorrelated.
- **Theorem:** If $\boldsymbol{V}$ and $\boldsymbol{W}$ are JG, their linear combinations $A\boldsymbol{V} + \boldsymbol{a}$ and $B\boldsymbol{W} + \boldsymbol{b}$ are also JG.
- **Example:** Given IID $\mathcal{N}(0,1)$ $X, Y$, $X + Y$ and $X - Y$ are JG.