جامعة تكريت
كلية الهندسة – الشرقاط
قسم الهندســـة الكهربائيـــة
Department of Electrical Engineering

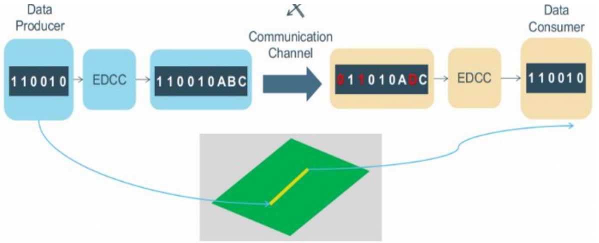
EE426
Information Theory

Asst. Prof. Dr. Ayad A. ABDULKAFI

2024-2025

- **1. Information Theory:**
- Review Of Probability
- Self Information, Source Entropy, Source Entropy Rate
- Mutual Information , Transinformation, Marginal Entropies , Joint and conditional Entropies
- Channels Types and Venn Diagram , Channel Capacity, Efficiency and Redundancy and Cascading of Channels
- **2. Source Coding**
- Source Coding of Discrete Sources, Coding Efficiency and Redundancy
- Fixed Length Codes
- Variable Length Codes
- Shannon Codes
- Shannon-Fano Codes (Fano Codes)
- Huffman Codes

- **3. Channel Coding and Error Correcting Codes**
- Error Detecting and Correcting Codes
- Systematic and Non Systematic Codes, Hamming Distance, Hamming Weight, Hamming Bound
- Linear and Non Linear Block Codes, Hamming Codes, Encoding and Decoding
- Cyclic Codes, Encoding and Decoding and Implementation



References:

- **Lecture Notes**
- **Information Transmission, Modulation, and Noise, M Schwartz, 1990.**
- **Information and Coding Theory, Gareth A. jones and j. Mary jones. Springer, 2000.**
- **Modern Digital and Analog Communication Systems, 4th ed.; B.P. Lathi and Zhi Ding; Oxford University Press; 2009.**
- **Communication Systems Engineering 2nd Ed by John G. Proakis and Masoud Salehi 2002.**

Introduction

What is Data?

Data is defined as a collection of individual facts or statistics. Data can come in the form of text, observations, figures, images, numbers, graphs, or symbols. For example, data might include individual prices, weights, addresses, ages, names, temperatures, dates, or distances. Data is a raw form of knowledge and, on its own, doesn’t carry any significance or purpose. In other words, you have to interpret data for it to have meaning. Data can be simple—and may even seem useless until it is analysed, organized, and interpreted.

Information: an organized, meaningful, and useful clarification of data. It is thus related to data and knowledge, as data represents values attributed to parameters, and knowledge signifies understanding of real things or abstract concepts. Most scientists agree that information theory began in 1948 with Shannon’s famous article. In that paper, he provided answers to the following questions: - What is “information” and how to measure it? - What are the fundamental limits on the storage and the transmission of information?

The Key Differences Between Data vs Information

- ✓ Data is a collection of facts, while information puts those facts into context.
- ✓ While data is raw and unorganized, information is organized.
- ✓ Data, on its own, is meaningless. When it’s analysed and interpreted, it becomes meaningful information.
- ✓ Data does not depend on information; however, information depends on data.
- ✓ Data typically comes in the form of graphs, numbers, figures, or statistics. Information is typically presented through words, language, thoughts, and ideas.
- ✓ Data isn’t sufficient for decision-making, but you can make decisions based on information.

Uncertainty and Information

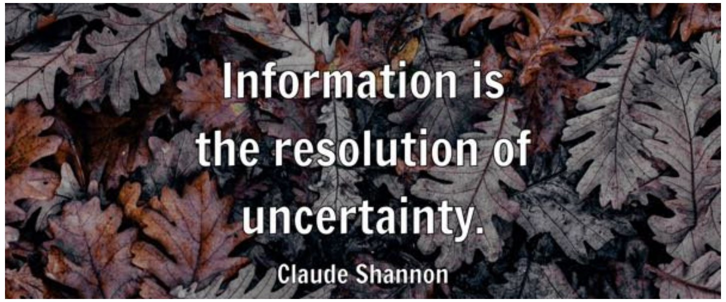
Before we go on to develop a mathematical measure of information, let us develop an intuitive feel for it.

Read following sentences:

(A)Tomorrow, the sun will rise from the East.

(B) The phone will ring in the next one hour.

(C) It will snow in Shirqat this summer.



The three sentences carry different amounts of information.

- In fact, the first sentences hardly carry any information. Everybody knows that the sun rises in the east and probability of this happening again is almost unity.
- Sentences (B) appears to carry more information than sentence (A). The phone may ring, or it may not. There is a finite probability that the phone will ring in the next one hour.
- The last sentence probably made you read it over twice. This is because it has never snowed in Shirqat, and the probability of snowfall is very low.
- It is interesting to note that the amount of information carried by the sentences listed above have something to do with the probability of occurrence of the events started in the sentences. And we observe an inverse relationship. Sentence (A), which talks about an event which has a probability of occurrence very close to 1 carries almost no information.
- Sentence (C), which has very low probability of occurrence, appears to carry a lot of information (made us read it twice to be sure we got the information right!).
- The other interesting thing to note is that the length of the sentence has nothing to do with the amount of information it carries. In fact, sentence (A) is the longest but carries the minimum information.

❖ In this chapter, we will develop a mathematical measure of information:

Probability:- A foundational concept from information is the quantification of the amount of information in things like events, random variables, and distributions. Quantifying the amount of information requires the use of probabilities, hence the relationship of information theory to probability. Probability is the branch of mathematics concerning numerical descriptions of how likely an event is to occur, or how likely it is that a proposition is true. The probability of an event is a number between 0 and 1, where, roughly speaking, 0 indicates impossibility of the event and 1 indicates certainty.

Probability of event X (self or marginal entropy)

If the experiment is repeated (N) times then the probability of the event (X_i) is given by

$$\boxed{P(X) = 1 \lim_{n\to\infty}\left(\frac{n\,(X_i)}{N}\right)}$$

Where: n (X) is the number of times (outcomes) that the event X occurs, N is the total number of trails

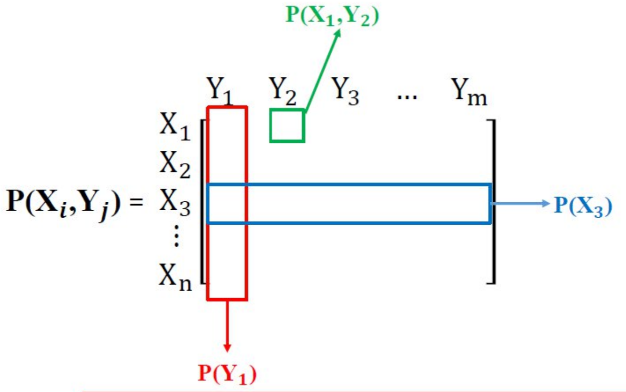
Trial: By a trial, we mean performing a random experiment. Example: (i) Tossing a fair coin, (ii) rolling an unbiased (fair) die.

Note that $0 \le P(X_i) \le 1$ and $\sum_{i=1}^{N} P(X_i) = 1$.

Joint Probability $P(X_i, Y_j)$

Joint probability is a statistical measure that calculates the likelihood of two events occurring together and at the same point in time. Joint probability is the probability of event Y occurring at the same time that event X occurs. If we have two experiments X & Y such that experiment X has mutually exclusive outcomes $X_1, X_2, X_3, \ldots, X_n$ and experiment Y has mutually exclusive outcomes $Y_1, Y_2, Y_3, \ldots, Y_m$, then $P(X_i, Y_j)$ is the joint probability. In addition, $P(X_i, Y_j)$ can be arranged as a matrix.

$$P(X_i, Y_j) = \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \\ \ldots \\ X_n \end{array}\begin{array}{c} \begin{array}{ccccc} Y_1 & Y_2 & Y_3 & \ldots & Y_m \end{array} \\ \begin{bmatrix} \blacksquare & \blacksquare & \blacksquare & & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & \ldots & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & & \blacksquare \end{bmatrix} \end{array}$$



Note that

$$\sum_{j=1}^{m}\sum_{i=1}^{n} \mathbf{P(X}\boldsymbol{i}\mathbf{\,,Y}\boldsymbol{j}\mathbf{)} = 1.$$

$$\sum_{i=1}^{n} \mathbf{P(X}\boldsymbol{i}\mathbf{\,,Y}\boldsymbol{j}\mathbf{)} = P\left(Y_j\right)$$

sum of the ith column

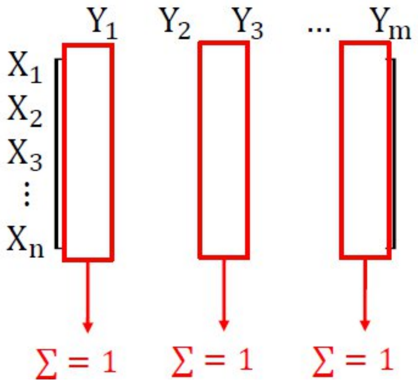
$$\sum_{j=1}^{m} \mathbf{P(X}\boldsymbol{i}\mathbf{\,,Y}\boldsymbol{j}\mathbf{)} = P\left(X_j\right)$$

sum of the ith row

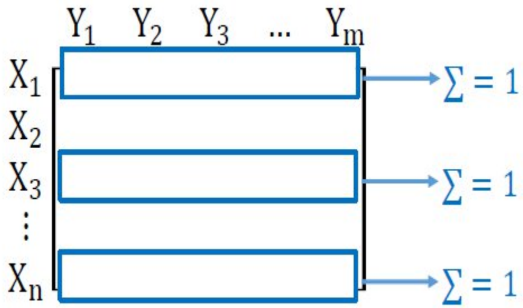
Conditional Probability $P(X_i/Y_j)$, $P(Y_j/X_i)$

Conditional probability is defined as the likelihood of an event or outcome occurring, based on the occurrence of a previous event or outcome. Conditional probability is calculated by multiplying the probability of the preceding event by the updated probability of the succeeding, or conditional, event. Conditional Probability is the probability of some event X, given the occurrence of some other event Y. Conditional probability is written: $P(X_i/Y_j)$ and is read "the probability of X_i, given Y_j". In addition, $P(Y_j/X_i)$ is conditional probability and read is "the probability of Y_j, given X_i.

$P(X_i/Y_j) =$



$P(Y_j/X_i)=$



Note that

$$P(X_i, Y_j) = P(X_i)\ P(Y_j/X_i)$$
$$= P(Y_j)\ P(X_i/Y_j)$$

$$P(X_i, Y_j) = P(Y_j, X_i)$$

$$P(Y_j/X_i) = \frac{P(X_i, Y_j)}{P(X_i)}$$

$$P(X_i/Y_j) = \frac{P(X_i, Y_j)}{P(Y_j)}$$

Note that sum of any column = 1 in P (X/Y)
while the sum of any row = 1 in P(Y/X)

Statistically Independence:-

Statistical independence is a concept in probability theory. Two events X and Y are statistically independent if and only if their joint probability can be factored into their marginal probabilities. The concept can be generalized to more than two events. Two events are independent, statistically independent, or stochastically independent if the occurrence of one does not affect the probability of occurrence of the other. Similarly, two random variables are independent if the realization of one does not affect the probability distribution of the other.

We have $\mathbf{P(X_i, Y_j) = P(X_i)\ P(Y_j/X_i) = P(Y_j)\ P(X_i/Y_j)}$, then when both **X** and **Y** are statistically independent events, we can obtain the following relations.

$\mathbf{P(X_i / Y_j) = P(X_i)}$ and $\mathbf{P(Y_j/X_i) = P(Y_j)}$ and hence $\mathbf{P(Y_j, X_i) = P(X_i)\ P(Y_j)}$

Homework: X & Y have the joint probability matrix

$$\mathbf{P(X_i,Y_j)} = \begin{array}{c} \\ X_1 \\ X_2 \\ X_3 \end{array} \begin{array}{c} \begin{array}{cc} Y_1 & \quad Y_2 \end{array} \\ \begin{bmatrix} 0.1 & 0.25 \\ 0 & 0.2 \\ 0.25 & 0.2 \end{bmatrix} \end{array}$$

, Find $P(X_2)$, $P(Y_1)$, and $P(X_i/Y_j)$ & $P(Y_j/X_i)$ for all i and j values.

Homework: The conditional probability has the matrix

$$P(Y_j/X_i) = \begin{bmatrix} 0.8 & a & 0.05 \\ b & 0.8 & a \\ a & b & 0.8 \end{bmatrix}$$

, Find a, b, and the joint probability if $P(X_1) = P(X_2) = P(X_3)$

Example: X has three outcomes X_1, X_2, X_3 & Y has two outcomes Y_1, Y_2, and the joint probability is

$$\mathbf{P(X_i,Y_j)} = \begin{array}{c c} & \begin{array}{cc} Y_1 & Y_2 \end{array} \\ \begin{array}{c} X_1 \\ X_2 \\ X_3 \end{array} & \begin{bmatrix} 0.2 & 0.1 \\ 0.15 & 0.25 \\ 0.1 & 0.2 \end{bmatrix} \end{array}, \quad \text{Find } P(X_1), P(Y_2), \text{ and } P(X_i/Y_j).$$

Solution:

1) $P(X_1) = 0.2 + 0.1 = 0.3$
2) $P(Y_2) = 0.1 + 0.25 + 0.2 = 0.55$
3) Additional calculations $P(X_2) = 0.15 + 0.25 = 0.4$, $P(X_3) = 0.1 + 0.2 = 0.3$, $P(Y_1) = 0.2 + 0.15 + 0.1 = 0.45$
4) Note that: $P(X) = P(X_1) + P(X_2) + P(X_3) = 1$, also $P(Y) = P(Y_1) + P(Y_2) = 1$
5) $P(X_i/Y_j) = \dfrac{P(X_i,Y_j)}{P(Y_j)} =$

$$P(X_i/Y_j) = \begin{array}{c c} & \begin{array}{cc} Y_1 & Y_2 \end{array} \\ \begin{array}{c} X_1 \\ X_2 \\ X_3 \end{array} & \begin{bmatrix} \frac{0.2}{0.45} & \frac{0.1}{0.55} \\ \frac{0.15}{0.45} & \frac{0.25}{0.55} \\ \frac{0.1}{0.45} & \frac{0.2}{0.55} \end{bmatrix} \end{array}, \quad P(Y_j/X_i) = \frac{P(X_i,Y_j)}{P(X_i)} = \begin{array}{c} X_1 \\ X_2 \\ X_3 \end{array} \begin{bmatrix} \frac{0.2}{0.3} & \frac{0.1}{0.3} \\ \frac{0.15}{0.4} & \frac{0.25}{0.4} \\ \frac{0.1}{0.3} & \frac{0.2}{0.3} \end{bmatrix}$$