Preface and Introduction

Po-Ning Chen

Institute of Communications Engineering

National Chiao-Tung University

Hsin Chu, Taiwan 30010

Introduction

- This book will cover some advanced topics in information theory by primarily focusing on systems with arbitrary statistical memory.
- The mathematical background on which these topics are based can be found in Appendices A and B in [2].

[2] F. Alajaji and P.-N. Chen. *An Introduction to Single-User Information Theory*, Springer, July 2018.

Notations

- For a discrete random variable X, we use P_X to denote its distribution.
- The probability of the elementary event $[X = x]$ is denoted by

$$\text{either} \quad \Pr[X = x] \quad \text{or} \quad P_X(x).$$

- Similarly, the probability of a set characterized by an inequality, such as $f(x) < a$, is expressed by

$$\text{either} \quad P_X\{x \in \mathcal{X} \ : \ f(x) < a\} \quad \text{or} \quad \Pr\left[f(X) < a\right].$$

- In the second expression, $f(X)$ is a new random variable defined through X and a function $f(\cdot)$.
- Obviously, the above expressions can be applied to any legitimate function $f(\cdot)$ defined over $\mathcal{X}$, including any probability function $P_{\hat{X}}(\cdot)$ (or $\log P_{\hat{X}}(x)$) of a random variable $\hat{X}$.
- Therefore, the next two expressions denote the probability of $f(x) = P_{\hat{X}}(x) < a$ evaluated under distribution P_X:

$$P_X\{x \in \mathcal{X} \ : \ f(x) < a\} = P_X\left\{x \in \mathcal{X} \ : \ P_{\hat{X}}(x) < a\right\}$$

and

$$\Pr\left[f(X) < a\right] = \Pr\left[P_{\hat{X}}(X) < a\right].$$

Notations

- As a result, if we write

$$P_{X,Y}\left\{(x,y)\in\mathcal{X}\times\mathcal{Y}\ :\ \log\frac{P_{\hat{X},\hat{Y}}(x,y)}{P_{\hat{X}}(x)P_{\hat{Y}}(y)}<a\right\}$$
$$=\ \Pr\left[\log\frac{P_{\hat{X},\hat{Y}}(X,Y)}{P_{\hat{X}}(X)P_{\hat{Y}}(Y)}<a\right],$$

we mean that we have defined a new function

$$f(x,y):=\log\frac{P_{\hat{X},\hat{Y}}(x,y)}{P_{\hat{X}}(x)P_{\hat{Y}}(y)}$$

in terms of the joint distribution $P_{\hat{X},\hat{Y}}$ and its two marginal distributions, and that we are interested in the probability of $f(X,Y)<a$ where X and Y have joint distribution $P_{X,Y}$.