

# ECSE 6530 - Information Theory

## Part I - Foundations of Information Theory

### 1- Information Measures: Entropy and Divergence

1. Entropy: axiomatic characterization
2. History of entropy
3. Entropy: Han's inequality and Shearer's Lemma
4. Divergence
5. Differential entropy

### 2 - Information Measures: Mutual Information

1. Divergence: Main inequality
2. Conditional divergence
3. Mutual information
4. Conditional mutual information and conditional independence
5. Strong data-processing inequalities

### 3 - Sufficient Statistics: Continuity of Divergence and Mutual Information

1. Sufficient statistics and data processing
2. Geometric interpretation of mutual information
3. Variational characterizations of divergence: Donsker-Varadhan
4. Variational characterizations of divergence: Gelfand-Yaglom-Perez
5. Continuity of divergence. Dependence on  $\sigma$ -algebra
6. Variational characterizations and continuity of mutual information

## Part II - Applications of Information Theory to Statistics

### 5 - Statistical Decision Theory : Fundamentals

1. Introduction
2. Minimax risk
3. Bayes risk
4. Minimax risk of Gaussian location models
5. Sample complexity
6. Nonparametric extension
7. Non-quadratic loss

### 6 - Hypothesis testing: Binary Models

1. Binary hypothesis testing
2. Neyman-Pearson formulation
3. Likelihood ratio tests
4. Achievable and converse bounds
5. Asymptotes
6. Information projection and large deviation

### 7 - $f$ -divergence: Information inequalities

1.  $f$ -divergences
2. Total variations
3. Inequality between  $f$ -divergences
4. Variational representation
5. Fisher information
6. Hammersley-Chapman-Robbins (HCR) lower bound
7. Cramér-Rao (CR) lower bound
8. Bayesian CR lower bound

## Part III - Applications of Information Theory to High-dimensional Statistics

### 8 - Unstructured Estimation: High-dimensional regime

1. Exact minimax risk for Gaussian location model
2. Le Cam's method
3. Two-point method
4. Mutual information method
5. Fano's method
6. Density estimation

### 9 - Structured Estimation: High-dimensional regime

1. Denoising sparse vectors
2. Denoising lower bounds
3. Denoising upper bounds
4. Minimax rates for sparse linear regression

### 10 - Functional Estimation: Composite Hypothesis testing

1. Functional estimation and testing
2. Lower bounds
3. LeCam's method
4. Upper bounds
5. Approximation-theoretic methods
6. Adaptation and Aggregation

### 11 - Miscellaneous: Other topics in machine/statistical learning

1. Maximum entropy principle
2. Maximum conditional entropy principle
3. Duality to maximum likelihood
4. Method of types and applications to hypothesis testing
5. Estimation of information measures