

---

# Contents

|          |   |     |
|----------|---|-----|
| <b>1</b> | <b>Universal lossless data compression</b>                              | 5   |
| 1.1      | A link between coding and estimation                                    | 5   |
| 1.2      | Universal coding and mixture codes                                      | 13  |
| 1.3      | Lower bounds for the minimax compression rate                           | 20  |
| 1.4      | Mixtures of i.i.d. coding distributions                                 | 25  |
| 1.5      | Double mixtures and adaptive compression                                | 33  |
|          | <b>Appendix</b>   | 49  |
| 1.6      | Fano's lemma  | 49  |
| 1.7      | Decomposition of the Kullback divergence function                       | 50  |
| <b>2</b> | <b>Links between data compression and statistical estimation</b>        | 55  |
| 2.1      | Estimating a conditional distribution                                   | 55  |
| 2.2      | Least square regression   | 56  |
| 2.3      | Pattern recognition   | 58  |
| <b>3</b> | <b>Non cumulated mean risk</b>  | 71  |
| 3.1      | The progressive mixture rule  | 71  |
| 3.2      | Estimating a Bernoulli random variable                                  | 76  |
| 3.3      | Adaptive histograms   | 78  |
| 3.4      | Some remarks on approximate Monte-Carlo computations                    | 80  |
| 3.5      | Selection and aggregation : a toy example pointing out some differences | 81  |
| 3.6      | Least square regression   | 83  |
| 3.7      | Adaptive regression estimation in Besov spaces                          | 89  |
| <b>4</b> | <b>Gibbs estimators</b>   | 97  |
| 4.1      | General framework   | 97  |
| 4.2      | Dichotomic histograms   | 103 |
| 4.3      | Mathematical framework for density estimation                           | 113 |
| 4.4      | Main oracle inequality  | 117 |

|          |   |            |
|----------|---|------------|
| 4.5      | Checking the accuracy of the bounds on the Gaussian shift model . . . . . | 120        |
| 4.6      | Application to adaptive classification . . . . .                          | 123        |
| 4.7      | Two stage adaptive least square regression . . . . .                      | 131        |
| 4.8      | One stage piecewise constant regression . . . . .                         | 136        |
| 4.9      | Some abstract inference problem . . . . .                                 | 144        |
| 4.10     | Another type of bound . . . . .   | 153        |
| <b>5</b> | <b>Randomized estimators and empirical complexity</b> . . . . .           | <b>155</b> |
| 5.1      | A pseudo-Bayesian approach to adaptive inference . . . . .                | 155        |
| 5.2      | A randomized rule for pattern recognition . . . . .                       | 158        |
| 5.3      | Generalizations of theorem 5.2.3 . . . . .                                | 165        |
| 5.4      | The non-ambiguous case . . . . .  | 167        |
| 5.5      | Empirical complexity bounds for the Gibbs estimator . . . . .             | 173        |
| 5.6      | Non randomized classification rules . . . . .                             | 176        |
| 5.7      | Application to classification trees . . . . .                             | 177        |
| 5.8      | The regression setting . . . . .  | 181        |
| 5.9      | Links with penalized least square regression . . . . .                    | 186        |
| 5.10     | Some elementary bounds . . . . .  | 193        |
| 5.11     | Some refinements about the linear regression case . . . . .               | 194        |
| <b>6</b> | <b>Deviation inequalities</b> . . . . .                                   | <b>199</b> |
| 6.1      | Bounded range functionals of independent variables . . . . .              | 200        |
| 6.2      | Extension to unbounded ranges . . . . .                                   | 206        |
| 6.3      | Generalization to Markov chains . . . . .                                 | 210        |
| <b>7</b> | <b>Markov chains with exponential transitions</b> . . . . .               | <b>223</b> |
| 7.1      | Model definition . . . . .  | 223        |
| 7.2      | The reduction principle . . . . .   | 225        |
| 7.3      | Excursion from a domain . . . . .   | 230        |
| 7.4      | Fast reduction algorithm . . . . .  | 235        |
| 7.5      | Elevation function and cycle decomposition . . . . .                      | 237        |
| 7.6      | Mean hitting times and ordered reduction . . . . .                        | 244        |
| 7.7      | Convergence speeds . . . . .  | 249        |
| 7.8      | Generalized simulated annealing algorithm . . . . .                       | 255        |
|          | <b>References</b> . . . . .   | <b>261</b> |
|          | <b>Index</b> . . . . .  | <b>267</b> |
|          | <b>List of participants</b> . . . . .                                     | <b>271</b> |
|          | <b>List of short lectures</b> . . . . .                                   | <b>273</b> |