

Chapter 12

Group Assignments

12.1 Group Assignment 1

- Prove Lemma [1.14](#)
- Prove Lemma [2.8](#)
- Prove property 4 of Theorem [2.18](#)
- Solve Exercise [2.59](#)
- Prove the "tower property" (Theorem [2.60](#)) for a discrete random variable taking a finite number of values.

12.2 Group Assignment 2

- Prove Corollary [3.7](#).
- Prove Lemma [3.15](#), properties 1-4.
- Solve Exercise [3.16](#)
- Solve Exercise [4.7](#)
- Prove Theorem [4.9](#) with all details, basically referring to all the properties of the indicator function used, the monotonicity of measures etc.

12.3 Group Assignment 3

- Solve Exercise [5.20](#)
- Solve Exercise [6.11](#)

- Solve Exercise [6.19](#)
- Solve Exercise [7.12](#)
- Solve Exercise [7.17](#)

Bibliography

- [BIHo] Blum, A., Hopcroft, J., and Kannan, R. *Foundations of data science*. Cambridge University Press, 2020.
- [BLM] Boucheron, S., Lugosi, G., Massart, P. *Concentration inequalities: a nonasymptotic theory of independence*, Oxford University Press, 2013.
- [PTPR] Devroye, L., Györfi, L. and Lugosi, G., *A Probabilistic Theory of Pattern Recognition*, Springer 1997
- [HD] T. E. Hull and A. R. Dobell, *Random Number Generators*, SIAM Review 1962 4:3, 230-254
- [K] Knuth, Donald (1997). Seminumerical Algorithms. The Art of Computer Programming. 2 (3rd ed.). Reading, MA: Addison-Wesley Professional. pp. 10–26.
- [ReWa] Reiß, M., and Wahl, M. *Nonasymptotic upper bounds for the reconstruction error of PCA*. The Annals of Statistics 48.2 (2020): 1098-1123.
- [SLT] Vapnik, V. , *Statistical Learning Theory*, Wiley, 1998.
- [WW] M. Wainwright, *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*, 2019.
- [W] Wasserman, L. All of Statistics. Springer Texts in Statistics, Springer, 2004.

Index

- σ -algebra, 4
- d -dimensional ball, 141
- d -dimensional sphere, 141
- t -step transition matrix, 103
- n-product experiment, 2
- “complexity”, 122
- homogeneous, 104
- Accept-Reject Sampler, 96
- Addition Rule, 4
- adjacency matrix, 109
- aperiodic state, 107
- asymptotically consistent, 81
- asymptotically consistent estimator, 82
- asymptotically unbiased, 79
- Bayes classification rule, 72
- Bennett’s inequality, 139
- Bonferroni correction, 53
- Boole’s inequality, 6
- Borel sigma-algebra, 8
- Chebychev’s inequality, 49
- communicates, 106
- conditional probability, 9
- conditional probability mass / density function, 42
- congruential generator, 92
- continuous, 20
- Convergence Almost Surely, 61
- Convergence in L^p , 62
- Convergence in Distribution, 58
- Convergence in Probability, 61
- cumulative distribution function, 15
- cylinder set, 9
- data space, 78
- decision function, 71
- decision rule, 71
- decreasing, 25
- degree, 109
- directed edge, 109
- directed graph, 109
- discrete, 18
- discrete (or discrete-time) stochastic process, 99
- distribution function, 15, 27
- edge, 109
- edge set, 109
- empirical measure, 125
- empirical risk minimizer, 120, 136
- events, 1
- Events in Probability Model, 8
- expectation, 30
- expected value, 30
- experiment, 1, 7
- first moment, 30
- first singular value, 156
- first singular vector, 156
- Google’s random surfer model, 110
- Graph, 109

- graph theory definitions, 109
- Greedy Algorithm, 158
- half-spaces, 8
- Hoeffdings inequality, 51
- Hoeffdings lemma, 49
- homogeneous, 101, 102
- in-edges, 109
- inclusion-exclusion principle, 6
- increasing, 25
- independent, 4, 12
- Indicator Function, 17
- integral, 77
- intercommunicates, 106
- irreducible, 106
- Johnson Lindenstrauss, 154
- joint cumulative distribution function, 32
- joint distribution function, 32
- joint probability density function, 35
- joint probability mass function, 33
- left singular vectors, 161
- Linear Classifiers, 122
- marginal distribution, 32, 34, 35
- marginal probability mass function, 34, 35
- Markov chain, 101, 102
- Markov's inequality, 48
- Matrix Bernstein inequality, 165
- mean, 30
- mean squared error, 81
- multigraph, 109
- mutually exclusive, 1
- neighbourhood, 109
- neighbours, 109
- non-parametric model, 66
- norm, 38
- normalized Gaussian, 142
- occured, 1
- one-to-one and monotone, 25
- out-edges, 109
- outcomes, 1
- pair-wise disjoint, 1
- parameter, 77
- parameter space, 77
- parametric model, 66
- period, 107
- periodic, 107
- Point estimation, 77
- point estimator, 78
- possible return times, 107
- posterior probability of, 12
- power set, 8
- prior probability of, 12
- probability density function (PDF), 20
- probability mass function, 18
- probability measure, 5, 14
- probability model, 7, 8
- probability triple, 5, 7
- property graphs, 109
- pseudorandom, 91
- random mapping representation, 105
- Random Projection, 153
- random surf, 110
- Random Variable (RV), 15, 31
- random walk on graphs, 108
- Real-world Interpretation, 8
- reducible, 106
- reversible Markov chain, 108
- rotationally symmetric, 149
- sample mean, 80
- sample space, 1
- sampling distribution, 79
- Semigroup, 104
- sigma-algebra, 4
- sigma-field, 4
- simple event, 1

Singular Value Decomposition,
155

Something Happens, 3

spherical Gaussian, 142

standard deviation, 30

standard error, 80

standard Gaussian, 22

state space, 101, 102

stationary distribution, 107

statistic, 78

statistical model, 66

stochastic process, 99

sub-exponential, 54

sub-Gaussian, 53

SVD, 155

testing dataset, 120, 136

testing set, 120, 136

training dataset, 120, 136

transition matrix, 102

trial, 2

UCEM, 126

unbiased, 79

undirected edge, 109

undirected graph, 109

uniform at random from the unit
ball, 148

uniform at random from the unit
sphere, 148

uniform pseudorandom number
generator, 91

union bound, 7

unit ball, 141, 145

unit sphere, 141

validation data, 121

variance, 30

VC-dimension, 130

vertex set, 109

vertices, 109

weighted, 109