Mathematics and Statistics Undergraduate Handbook Supplement to the Handbook

Honour School of Mathematics and Statistics
Syllabus and Synopses for Part C 2018–2019
for examination in 2019

Contents

SC8 is now Topics in Computational Biology. Two new courses SC9 Interacting Particle Systems and SC10 Algorithmic Foundations of Learning have been added for 2018/2019.

List of Mathematics units updated.
Number of units to be taken updated.

Every effort is made to ensure that the list of courses offered is accurate at the time of going online. However, students are advised to check the up-to-date version of this document on the Department of Statistics website.
Notice of misprints or errors of any kind, and suggestions for improvements in this booklet should be addressed to the Academic Administrator in the Department of Statistics, academic.administrator@stats.ox.ac.uk.

Updated January 2019   v.4

1   Honour School of Mathematics and Statistics

1.1   Units

See the current edition of the Examination Regulations at
http://www.admin.ox.ac.uk/examregs/2015-16/hsomathandstat/studentview/
for the full regulations governing these examinations. The examination conventions can be
found at http://www.stats.ox.ac.uk/current_students/bammath/examinations

In Part C,
- each candidate shall offer a minimum of six units and a maximum of eight units from the schedule of units for Part C
- and each candidate shall also offer **a dissertation** on a statistics project (equivalent of 2 units).

At least **two** units should be from the schedule of 'Statistics' units.
The USMs for the dissertation and the best six units will count for the final classification.

Units from the schedule of 'Mathematics Department units' for Part C of the Honour School of
Mathematics are also available – see Section 3.

This booklet describes the units available in Part C. Information about dissertations/ statistics
projects will be available on the Department of Statistics website at
http://www.stats.ox.ac.uk/current_students/bammath/projects

All of the units described in this booklet are "M-level".

Students are asked to register for the options they intend to take by the end of week 10, Trinity
Term 2018 using the Mathematical Institute course management portal.
https://courses.maths.ox.ac.uk/ . Students may alter the options they have registered for after
this but it is helpful if their registration is as accurate as possible. Students will then be asked to
sign up for classes at the start of Michaelmas Term 2018. Students who register for a course or
courses for which there is a quota should consider registering for an additional course (by way
of a "reserve choice") in case they do not receive a place on the course with the quota.

Every effort will be made when timetabling lectures to ensure that mathematics lectures do not
clash. However, because of the large number of options this may sometimes be unavoidable.

1.2   Language Classes

If spaces are available, Mathematics and Statistics students are also invited to apply to take
classes in a foreign language.   In 2018-2019, French and German language classes will be
offered. Students' performance in these classes will not contribute to the degree classification in
Mathematics and Statistics. However, successful completion of the course, may be recorded on
a student's transcript. See   https://www1.maths.ox.ac.uk/members/students/undergraduate-
courses/teaching-and-learning/part-bc-students#PartC_Options
 for further information.

1.3   **Part C courses in future years**

In any year, most courses available in Part C that year will normally also be available in Part C the following year. However, sometimes new options will be added or existing options may cease to run. The list of courses that will be available in Part C in any year will be published by the end of the preceding Trinity Term.

1.4   **Course list by term**

The list of 2018-2019 Part C courses by term is:

Michaelmas Term

SC1 Stochastic Models in Mathematical Genetics
SC2 Probability and Statistics for Network Analysis
SC6 Graphical Models
SC9 Interacting Particle Systems
SC10 Algorithmic Foundations of Learning

Hilary Term

SC4 Advanced Topics in Statistical Machine Learning
SC5 Advanced Simulation Methods
SC7 Bayes Methods
SC8 Topics in Computational Biology
C8.4 Probabilistic Combinatorics.

2.   Statistics Units

2.1   SC1 Stochastic Models in Mathematical Genetics – 16 MT

Level: M-level
Method of Assessment: written examination
Weight: Unit

Recommended Prerequisites
Part A A8 Probability.
SB3.1 (formerly SB3a) Applied Probability would be helpful.

Aims & Objectives
The aim of the lectures is to introduce modern stochastic models in mathematical population genetics and give examples of real world applications of these models. Stochastic and graph theoretic properties of coalescent and genealogical trees are studied in the first eight lectures. Diffusion processes and extensions to model additional key biological phenomena are studied in the second eight lectures.

Synopsis
Evolutionary models in Mathematical Genetics:
The Wright-Fisher model. The Genealogical Markov chain describing the number ancestors back in time of a collection of DNA sequences.

The Coalescent process describing the stochastic behaviour of the ancestral tree of a collection of DNA sequences. Mutations on ancestral lineages in a coalescent tree. Models with a variable population size.

The frequency spectrum and age of a mutation. Ewens' sampling formula for the probability distribution of the allele configuration of DNA sequences in a sample in the infinitely-many-alleles model. Hoppe's urn model for the infinitely-many-alleles model.

The infinitely-many-sites model of mutations on DNA sequences. Gene trees as perfect phylogenies describing the mutation history of a sample of DNA sequences. Graph theoretic constructions and characterizations of gene trees from DNA sequence variation. Gusfield's construction algorithm of a tree from DNA sequences. Examples of gene trees from data.

Modelling biological forces in Population Genetics: Recombination. The effect of recombination on genealogies. Detecting recombination events under the infinitely-many-sites model. Hudson's algorithm. Haplotype bounds on recombination events. Modelling recombination in the Wright-Fisher model. The coalescent process with recombination: the ancestral recombination graph. Properties of the ancestral recombination graph.

Introduction to diffusion theory. Tracking mutations forward in time in the Wright-Fisher model. Modelling the frequency of a neutral mutation in the population via a diffusion process limit. The generator of a diffusion process with two allelic types. The probability of fixation of a mutation. Genic selection. Extension of results from neutral to selection case. Behaviour of selected mutations.

Reading
R. Durrett, *Probability Models for DNA Sequence Evolution*, Springer, 2008
A.Etheridge, Some Mathematical Models from Population Genetics. Ecole d'Eté de Probabilities de Saint-Flour XXXIX-2009, Lecture Notes in Mathematics, 2012
W. J. Ewens, *Mathematical Population Genetics,* 2nd Ed, Springer, 2004
J. R. Norris, *Markov Chains*, Cambridge University Press, 1999
 M. Slatkin and M. Veuille, *Modern Developments in Theoretical Population Genetics*, Oxford Biology, 2002
S. Tavaré and O. Zeitouni, *Lectures on Probability Theory and Statistics, Ecole d'Eté de Probabilities de Saint-Flour XXXI - 2001*, Lecture Notes in Mathematics 1837, Springer, 2004

2.2 **SC2 Probability and Statistics for Network Analysis** – 16 MT

Level: M-level
Method of Assessment: Written examination
Weight: Unit

For this course, 2 lectures and 2 intercollegiate classes are replaced by 2 practical classes. (The total time for this course is the same as for other Part C courses.)

Recommended prerequisites
Part A A8 Probability and A9 Statistics

Aims and Objectives
Many data come in the form of networks, for example friendship data and protein-protein interaction data. As the data usually cannot be modelled using simple independence assumptions, their statistical analysis provides many challenges. The course will give an introduction to the main problems and the main statistical techniques used in this field. The techniques are applicable to a wide range of complex problems. The statistical analysis benefits from insights which stem from probabilistic modelling, and the course will combine both aspects.

Synopsis
Exploratory analysis of networks. The need for network summaries. Degree distribution, clustering coefficient, shortest path length. Motifs.

Probabilistic models: Bernoulli random graphs, geometric random graphs, preferential attachment models, small world networks, inhomogeneous random graphs, exponential random graphs.

Small subgraphs: Stein's method for normal and Poisson approximation. Branching process approximations, threshold behaviour, shortest path between two vertices.

Statistical analysis of networks: Sampling from networks. Parameter estimation for models. Inference from networks: vertex characteristics and missing edges. Nonparametric graph comparison: subgraph counts, subsampling schemes, MCMC methods. A brief look at community detection. Examples: protein interaction networks, social ego-networks.

Reading
R. Durrett, *Random Graph Dynamics*, Cambridge University Press,2007
P. Grindrod, Mathematical Underpinnings of Analytics, Oxford University Press, 2015
R.v.d. Hofstad, Random Graphs and Complex Networks, Manuscript available at
http://www.win.tue.nl/~rhofstad/
E.D Kolaczyk and G. Csádi, *Statistical Analysis of Network Data with R*, Springer, 2014
M. Newman, *Networks: An Introduction*. Oxford University Press, 2010

2.3   **SC4 Advanced Topics in Statistical Machine Learning** – 16 HT

Level: M-level
Methods of Assessment: written examination.
Weight: Unit

Recommended prerequisites
 The course requires a good level of mathematical maturity. Students are expected to be familiar with core concepts in statistics (regression models, bias-variance tradeoff, Bayesian inference), probability (multivariate distributions, conditioning) and linear algebra (matrix-vector operations, eigenvalues and eigenvectors). Previous exposure to machine learning (empirical risk minimisation, dimensionality reduction, overfitting, regularisation) is highly recommended. Students would also benefit from being familiar with the material covered in the following courses offered in the Statistics department: SB2.1 (formerly SB2a) Foundations of Statistical Inference and in SB2.2 (formerly SB2b) Statistical Machine Learning.

Aims and Objectives
Machine learning is widely used across many scientific and engineering disciplines, to construct methods to find interesting patterns and to predict accurately in large datasets.
This course introduces several widely used data machine learning techniques and describes their underpinning statistical principles and properties. The course studies both unsupervised and supervised learning and several advanced topics are covered in detail, including some state-of-the-art machine learning techniques. The course will also cover computational considerations of machine learning algorithms and how they can scale to large datasets.

Synopsis
Convex optimisation and support vector machines. Loss functions. Empirical risk minimisation. Kernel methods and reproducing kernel Hilbert spaces. Representer theorem. Representation of probabilities in RKHS.
Nonlinear dimensionality reduction: kernel PCA, spectral clustering.
Probabilistic and Bayesian machine learning: mixture modelling, information theoretic fundamentals, EM algorithm, Probabilistic PCA. Variational Bayes. Laplace Approximation.
Collaborative filtering models, probabilistic matrix factorisation.
Gaussian processes for regression and classification. Bayesian optimisation.
(+Latent Dirichlet allocation [if time allows])

Software
Knowledge of Python is not required for this course, but some examples may be done in Python.

Students interested in learning Python are referred to the following free University IT online courses, which should ideally be taken before the beginning of this course:
https://courses.it.ox.ac.uk/detail/LY020
https://courses.it.ox.ac.uk/detail/LY021

Reading
C. Bishop, Pattern Recognition and Machine Learning, Springer,2007
K. Murphy, Machine Learning: a Probabilistic Perspective, MIT Press, 2012

Further Reading
T. Hastie, R. Tibshirani, J Friedman, Elements of Statistical Learning, Springer, 2009
Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011,
http://scikit-learn.org/stable/tutorial/

2.4   **SC5 Advanced Simulation Methods** - 16 HT

Level: M-level
Methods of Assessment: This course is assessed by written examination.
Weight: Unit

Recommended Prerequisites
The course requires a good level of mathematical maturity as well as some statistical intuition and background knowledge to motivate the course. Students are expected to be   familiar with core concepts from probability (conditional probability, conditional densities, properties of conditional expectations, basic inequalities such as Markov's, Chebyshev's and Cauchy-Schwarz's, modes of convergence), basic limit theorems from probability in particular the strong law of large numbers and the central limit theorem, Markov chains, aperiodicity, irreducibility, stationary distributions, reversibility and convergence. Most of these concepts are covered in courses offered in the Statistics department, in particular prelims probability, A8 probability and SB3.1 (formerly SB3a) Applied Probability.
Familiarity with basic Monte Carlo methods will be helpful, as for example covered in A12 Simulation and Statistical Programming.
Some familiarity with concepts from Bayesian inference such as posterior distributions will be useful in order to understand the motivation behind the material of the course.

Aims and Objectives
The aim of the lectures is to introduce modern simulation methods.
This course concentrates on Markov chain Monte Carlo (MCMC) methods and Sequential Monte Carlo (SMC) methods. Examples of applications of these methods to complex inference problems will be given.

Synopsis
Classical methods: inversion, rejection, composition.

Importance sampling.

MCMC methods: elements of discrete-time general state-space Markov chains theory, Metropolis-Hastings algorithm.

Advanced MCMC methods: Gibbs sampling, slice sampling, tempering/annealing, Hamiltonian (or Hybrid) Monte Carlo, pseudo-marginal MCMC.

Sequential importance sampling.

SMC methods: nonlinear filtering.

Reading
C.P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd edition, Springer-Verlag, 2004

Further reading
J.S. Liu, *Monte Carlo Strategies in Scientific Computing*, Springer-Verlag, 2001

2.5 **SC6 Graphical Models** – 16 MT

Level: M-level
Method of Assessment: Written examination
Weight: Unit

Recommended Prerequisites
The basics of Markov chains (in particular, conditional independence) from Part A Probability is assumed. Likelihood theory, contingency tables, and likelihood-ratio tests are also important; this is covered in Part A Statistics. Knowledge of exponential families and linear models , as covered in SB2.1 (formerly SB2a) Foundations of Statistical Inference and SB1.1 (formerly SB1a) Applied Statistics, would be useful, but is not essential.

Aims and Objectives
This course will give an overview of the use of graphical models as a tool for statistical inference. Graphical models relate the structure of a graph to the structure of a multivariate probability distribution, usually via a factorisation of the distribution or conditional independence constraints.

This has two broad uses: first, conditional independence can provide vast savings in computational effort, both in terms of the representation of large multivariate models and in performing inference with them; this makes graphical models very popular for dealing with big data problems. Second, conditional independence can be used as a tool to discover hidden structure in data, such as that relating to the direction of causality or to unobserved processes. As such, graphical models are widely used in genetics, medicine, epidemiology, statistical physics, economics, the social sciences and elsewhere.

Students will develop an understanding of the use of conditional independence and graphical structures for dealing with multivariate statistical models. They will appreciate how this is applied to causal modelling, and to computation in large-scale statistical problems.

Synopsis
Independence, conditional independence, graphoid axioms
Exponential families, mean and canonical parameterisations, moment matching; contingency tables, log-linear models.
Undirected graphs, cliques, paths; factorisation and Markov properties, Hammersley-Clifford Theorem (statement only)
Trees, cycles, chords, decomposability, triangulation. Maximum likelihood in decomposable models, iterative proportional fitting.
The multivariate Gaussian distribution and Gaussian graphical models.
Directed acyclic graphs, factorisation. Paths, d-separation, moralisation. Ancestral sets and sub-models. Decomposable models as intersection of directed and undirected models.
Running intersection property, Junction trees; message passing, computation of marginal and conditional probabilities, introduction of evidence. Gibbs sampling.
Causal models, structural equations, interventions, constraint-based learning, faithfulness. The trek rule and back-door adjustment.

Reading
S.L. Lauritzen, *Graphical Models*, Oxford University Press, 1996
T. Hastie, R Tibshirani and J.Friedman, Elements of Statistical Learning, 2nd edition, Springer 2009 (available for free at https://statweb/stanford.edu/~tibs/ElemStatLearn/)
D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*, MIT Press, 2009
J. Pearl, *Causality,* 3rd edition, Cambridge University Press, 2013
M.J. Wainwright and M.I. Jordan, Graphical Models, Exponential Families, and Variational Inference, *Foundations and Trends in Machine Learning*, 2008 (available for free at https://people.eecs.berkeley.edu/~wainwrig/Papers/WaiJor08_FTML.pdf)

2.6   **SC7 Bayes Methods** – 16 HT

Level: M-level
Method of Assessment: Written examination
Weight: Unit

Recommended prerequisites
SB2.1 (formerly SB2a) Foundations of Statistical Inference is desirable, of which 6 lectures on Bayesian inference, decision theory and hypothesis testing with loss functions are assumed knowledge. A12 Simulation and Statistical Programming desirable.

Synopsis
Theory: Decision-theoretic foundations, Savage axioms. Prior elicitation, exchangeability. Bayesian Non-Parametric (BNP) methods, the Dirichlet process and the Chinese Restaurant Process. Asymptotics, information criteria and the Bernstein-von Mises approximation.

Computational methods: Bayesian inference via MCMC; Estimation of marginal likelihood; Approximate Bayesian Computation and intractable likelihoods; reversible jump MCMC.

Case Studies: extend understanding of prior elicitation, BNP methods and asymptotics through a small number of substantial examples. Examples to further illustrate building statistical models, model choice, model averaging and model assessment, and the use of Monte Carlo methods for inference.

Reading
C.P. Robert*,The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*, 2nd edition, Springer, 2001

Further Reading
A. Gelman et al*, Bayesian Data Analysis*, 3rd edition, Boca Raton Florida: CRC Press, 2014
P Hoff*, A First Course in Bayesian Statistical Methods,* Springer, 2010
DeGroot, Morris H., *Optimal Statistical Decisions.* Wiley Classics Library. 2004.

2.7 **SC8 Topics in Computational Biology** – 16 HT

Level: M-level
Method of Assessment: Mini-project
Weight: Unit

Recommended Prerequisites:
The lectures attempt to be self-contained but clearly knowledge algorithms, combinatorics and probability theory - A8 Probability and SB3.1 (formerly SB3a) Probability- would be a help. The course requires a good level of mathematical maturity.

Aims & Objectives
Modern molecular biology generates large amounts of data, such as sequences, structures and expression data, that needs different forms of statistical analysis and modelling to be properly interpreted. This course focuses on four topics within this vast area: Molecular Dynamics, Molecule Enumeration, Comparative Biology and Overview of Computational Biology and Computational Neurosciences.

Synopsis:
Overview of Computational Biology and Computational Neuroscience - Computational Biology is a very large and diverse field: Basically all the fields of biology where computation has started to be essential. Computational Neuroscience is in massive growth, but has a history going back to the 1940s with publications like McCullogh and Pitts (1943) paper on neural networks. The present progress is driven by progress on three fronts: (i) Experimental data on brains, nerve systems and individual neurons, (ii) increased success in designing artificial neural networks with an increasing variety in architectures with applications in Deep Learning/AI and (iii) the simulate very complex models as models of biological neural networks.

Molecular Dynamics (MD) – MD is another huge application area that describes the dynamics of molecules with few to thousands of atoms, for very short time periods like microseconds to nanoseconds. MD is bound to continue to grow for decades, and stochastic methods are central in exploring a large configuration space. The lectures are in Hamiltonian Dynamics; the

canonical distribution and stochastic differential distributions, the Langevin model for Brownian motion and comparison of MD trajectories.

Molecule Enumeration – How many molecules are possible with a given number of atoms, from the set of carbon/nitrogen/phosphorus/oxygen/sulphur (CNPOS)? This question is central in drug design and has many statistical problems embedded. Exhaustive enumeration is at present limited to molecules with 18 CNPOS atoms, and including one more atom expands the numbers about a factor of 10-100 at this point. But there are many other possible avenue such as sampling or exploring a subspace generated by an initial set of molecules and a set of reactions. There are many advanced issues in counting molecules such Polya Counting and imposing constraints making molecular graphs embeddable in 3D.

Comparative Biology – Phylogenetics and comparative genomics have been the important areas of the last 15-20 years as a consequence of the growth to sequence data. However, there are other levels of data and biological organisation that are as least as interesting: protein structures, networks, shapes, movements.   The lectures include models of evolution of these data types; the so-called COMPARATIVE MODEL; and simultaneous modelling of several levels.

Reading:
The teaching material from 2018 would be useful to browse, but the 2019 will have some change in syllabus and improvements:
https://heingroupoxford.com/learning-resources/topics-in-computational-biology/

Further Reading
M. Steel, Phylogeny: Discrete and Random Processes in *Evolution*, chapt 1-2, SIAM Press (2003).
T. Schlick, *Molecular Modeling and Simulation*. Chapt 13-14, Springer (2010).
M. Meringer "Structure Enumeration and Sampling" chapt. 8 in *Handbook of Chemoinformatics Algorithms* (eds Faulon) (2010). Chapman and Hall.
B.C. O'Meara Evolutionary Inferences from Phylogenies: A Review of Methods, *Annu. Rev. Ecol. Evol. Syst.* 2012. 43:267–85

2.8   **SC9 Interacting Particle Systems** – 16 MT

Level: M-level
Method of Assessment: Written examination
Weight: Unit

Recommended Prerequisites
Discrete and continuous time Markov process on countable state space, as covered for example in Part A A8 Probability and Part B SB3.1 (formerly SB3a) Applied Probability.

Aims and Objectives
The aim is to introduce fundamental probabilistic and combinatorial tools, as well as key models, in the theory of discrete disordered systems. We will examine the large-scale behaviour of systems containing many interacting components, subject to some random noise. Models of this type have a wealth of applications in statistical physics, biology and beyond, and we will see

several key examples in the course. Many of the tools we will discuss are also of independent theoretical interest, and have far reaching applications. For example, we will study the amount of time it takes for this process to reach the stationary distribution (mixing time). This concept is also important in many statistical applications, such as studying the run time of MCMC methods.

Synopsis

- Percolation and phase transitions.
- Uniform spanning trees, loop-erased random walks, and Wilson's algorithm.
- Random walks on graphs and electrical networks (discrete potential theory).
- Important models: Stochastic Ising model (Glauber dynamics), Random-cluster model, Contact process, Exclusion process, Hard-core model.
- Important tools: Monotonicity, coupling, duality, FKG inequality.
- Gibbs measures and a discussion of phase transitions in this context.
- Mixing times and the spectral gap.

Reading

G. Grimmett, *Probability on Graphs: random processes on graphs and lattices*, Cambridge University Press, 2010; 2017 (2nd edition).
B. Bollobás, O. Riordan, *Percolation*, Cambridge University Press, 2006.
T. Liggett, *Continuous time Markov Processes: an introduction*, American Mathematical Society, 2010.
D. A. Levin, Y. Peres, E. L. Wilmer, *Markov Chains and Mixing Times*, American Mathematical Society, 2009.
P. G. Doyle, J. L. Snell, *Random Walks and Electric Networks*, Mathematical Association of America, 1984.

2.9 **SC10 Algorithmic Foundations of Learning** – 16 MT
Level: M-level
Method of Assessment: Written examination
Weight: Unit

Recommended Prerequisites
The course requires a good level of mathematical maturity. Students are expected to be familiar with core concepts in probability (basic properties of probabilities, such as union bounds, and of conditional expectations, such as the tower property; basic inequalities such as Markov's and Jensen's), statistics (confidence intervals, hypothesis testing), and linear algebra (matrix-vector operations, eigenvalues and eigenvectors; basic inequalities, such as Cauchy-Schwartz's and Hölder's). Previous exposure to machine learning (empirical risk minimisation, overfitting, regularisation) is recommended.
Students would benefit from being familiar with the material covered in SB2.1 (formerly SB2a) Foundations of Statistical Inference (in particular, Decision Theory) and in SB2.2 (formerly SB2b) Statistical Machine Learning.

Aims and objectives

The course is meant to provide a rigorous theoretical account of the main ideas underlying machine learning, and to offer a principled framework to understand the algorithmic paradigms being used, involving tools from probability, statistics, and optimisation in high-dimension.

Synopsis

- Statistical learning framework for prediction, estimation, and online learning
- Probability:
  - -Maximal inequalities
  - -Rademacher and Gaussian complexities
  - -Elements of VC theory
  - -Covering and packing numbers
  - - Chaining
  - -Concentration inequalities
- Statistics:
  - -Bayes decision rules.
  - -Empirical risk minimisation.
  - -Error decomposition: generalisation, optimisation and approximation.
  - -Learning via uniform convergence, margin bounds, and algorithmic stability.
  - - Regularisation: structural (constraints and penalisation) and implicit (algorithmic)
  - - Convex loss surrogates.
  - -Slow and fast rates
  - -Minimax lower bounds and hypothesis testing
- Optimisation:
  - -Elements of convex theory.
  - -Oracle model. Gradient descent. Mirror descent.
  - -Stochastic oracle model. Stochastic gradient descent. Stochastic mirror descent. Variance reduction techniques.
  - - Approximate Message Passing
  - - Online optimisation
- Examples:
  - -Linear predictors, including Boosting
  - -Non-linear predictors, including Vector Support Machines and Neural Networks.
  - -High-dimensional estimators for sparse and low-rank problems, including Lasso.
  - -Online learning, including multi-armed bandit problems and algorithms.

Reading

Shai Shalev-Shwartz and Shai Ben-David. Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press, 2014.
Ramon van Handel. Probability in High Dimension. Lecture notes available online (http://www.princeton.edu/~rvan/APC550.pdf), 2016.
Sébastien Bubeck. Convex Optimization: Algorithms and Complexity. Foundations and Trends in Machine Learning, 2015.
Tor Lattimore and Csaba Szepesvari. Bandit algorithms. Book available online (http://downloads/tor-lattimore.com/book.pdf), 2008

2.10   **C8.4 Probabilistic Combinatorics -** 16 HT

Level: M-level
Method of Assessment: Written examination.
Weight: Unit

Recommended Prerequisites:
Part B Graph Theory and Part A A8 Probability. C8.3 Combinatorics is not an essential prerequisite for this course, though it is a natural companion for it.

Overview
Probabilistic combinatorics is a very active field of mathematics, with connections to other areas such as computer science and statistical physics. Probabilistic methods are essential for the study of random discrete structures and for the analysis of algorithms, but they can also provide a powerful and beautiful approach for answering deterministic questions. The aim of this course is to introduce some fundamental probabilistic tools and present a few applications.

Learning Outcomes
The student will have developed an appreciation of probabilistic methods in discrete mathematics.

Synopsis
First-moment method, with applications to Ramsey numbers, and to graphs of high girth and high chromatic number.
Second-moment method, threshold functions for random graphs.
Lovász Local Lemma, with applications to two-colourings of hypergraphs, and to Ramsey numbers.
Chernoff bounds, concentration of measure, Janson's inequality.
Branching processes and the phase transition in random graphs.
Clique and chromatic numbers of random graphs.

Reading
N. Alon and J.H. Spencer, *The Probabilistic Method*, 3rd edition, Wiley, 2008

Further Reading:
B. Bollobás, *Random Graphs*, 2nd edition, Cambridge University Press, 2001
 M. Habib, C. McDiarmid, J. Ramirez-Alfonsin, B. Reed, ed., *Probabilistic Methods for Algorithmic Discrete Mathematics,* Springer, 1998
S. Janson, T. Luczak and A. Rucinski, *Random Graphs,* John Wiley and Sons, 2000
M. Mitzenmacher and E. Upfal, *Probability and Computing: Randomized Algorithms and Probabilistic Analysis,* Cambridge University Press, New York (NY), 2005
M. Molloy and B. Reed, *Graph Colouring and the Probabilistic Method,* Springer, 2002
R. Motwani and P. Raghavan, *Randomized Algorithms,* Cambridge University Press, 1995

3   Mathematics units

The Mathematics units that students may take are drawn from Part C of the Honour School of Mathematics. For full details of these units, see the Syllabus and Synopses for Part C of the Honour School of Mathematics, which are available on the web at
https://www1.maths.ox.ac.uk/members/students/undergraduate-courses/teaching-and-learning/handbooks-synopses

The Mathematics units that are available are as follows:

Unit	Term
C1.1: Model Theory	16 MT
C1.2: Godel's Incompleteness Theorems	16 HT
C1.3: Analytic Topology	16 MT
C1.4: Axiomatic Set Theory	16 HT
C2.1: Lie Algebras	16 MT
C2.2: Homological Algebra	16 MT
C2.3: Representation Theory of Semisimple Lie Algebras	16 HT
C2.4: Infinite Groups	16 HT
C2.5: Non-Commutative Rings	16 HT
C2.6: Introduction to Schemes	16 HT
C2.7: Category Theory	16 MT
C3.1: Algebraic Topology	16 MT
C3.2: Geometric Group Theory	16 MT
C3.3: Differentiable Manifolds	16 MT
C3.4: Algebraic Geometry	16 MT
C3.5: Lie Groups	16 HT
C3.7: Elliptic Curves	16 HT
C3.8: Analytic Number Theory	16 HT
C3.9: Computational Algebraic Topology	16 HT
C3.10 Additive and Combinatorial Number Theory	16 HT
C4.1: Further Functional Analysis	16 MT
C4.3:Functional Analytic Methods for PDEs	16 MT
C4.4 Hyperbolic Equations	16 HT
C4.6: Fixed Point Methods for Nonlinear PDEs	16 HT
C4.8: Complex Analysis: Conformal Maps and Geometry	16 MT
C5.1: Solid Mechanics	16 MT
C5.2: Elasticity and Plasticity	16 HT
C5.4: Networks	16 HT
C5.5: Perturbation Methods	16 MT
C5.6: Applied Complex Variables	16 HT
C5.7: Topics in Fluid Mechanics	16 MT
C5.9 Mathematical Mechanical Biology	16 HT
C5.10 Mathematics and Data Science for Development	16 HT
C5.11: Mathematical Geoscience	16 MT
C5.12: Mathematical Physiology	16 MT
C6.1: Numerical Linear Algebra	16 MT
C6.2: Continuous Optimisation	16 HT
C6.3: Approximation of Functions	16 MT
C6.4: Finite Element Methods for PDEs	16 HT
C6.5 Theories of Deep Learning	16 HT

C7.1 Theoretical Physics   24MT/16HT
C7.4 Introduction to Quantum Information   16 HT
C7.5: General Relativity I   16 MT
C7.6: General Relativity II   16 HT
C8.1: Stochastic Differential Equations   16 MT
C8.2: Stochastic Analysis and PDEs   16 HT
C8.3: Combinatorics   16 MT