MSc in Statistical Science – Prerequisite material

Department of Statistics

University of Oxford

1 Preparing for the MSc

Please be aware that we expect you to arrive in Oxford with a good understanding of basic parts of mathematics, probability and statistics. MSc teaching is mathematically demanding and assumes mathematical maturity, so it is important that you don't neglect the prerequisites.

The probability and statistics material required is covered by the notes on

- Probability [1]
- Statistics [2]
- Simulation & Statistical Programming [3], [4]

copies of which are available at http://www.stats.ox.ac.uk/~laws/msc/.

The MSc will assume you have a good understanding of the prerequisite material and if necessary you should familiarise (or re-familiarise) yourself with it before you arrive so that you are ready when teaching starts. A good level of fluency with this material is important. For example you should be able to complete exercises on this material, not just know the definitions. At the above web-link there are copies of exercises associated with the Probability, Statistics, and Simulation & Statistical Programming notes for you to look at and practise on.

The main mathematical topics required are linear algebra/matrices plus calculus, including some multivariate calculus. These mathematical prerequisites will be assumed, they will not be revised during the MSc.

The following sections give the relevant topics and suggest some places you could look for coverage of the relevant material.

2 Prerequisite material

You could start by looking for the prerequsite material in your undergraduate lecture notes, textbooks, etc. Some online notes, and some good books, that cover this material are given below, but there are many alternatives that you could use instead.

2.1 Mathematics

Matrix algebra – transpose, symmetric, rank, inverse, orthogonal, trace
Matrix operations – multiplication, solution of linear system of equations
Eigenvalues, eigenvectors
Eigendecomposition and singular value decomposition of a matrix
Differentiation – including partial differentiation, Hessian matrix
Integration – including multiple integrals, and Jacobians and change of variables in multiple integrals
Taylor series expansion
Difference equations

The linear algebra notes [5] cover most of the required linear algebra.

Jordan and Smith [12] covers the material at an appropriate level. The relevant material is commonly included in books on engineering mathematics: good examples are Kreyszig [13] and Stroud [18].

2.2 Probability

Probability spaces
Random variables – discrete and continuous
Distributions, expectation, variance, covariance, independence
Joint distributions, conditional distributions, method of change of variables using Jacobian
Multivariate normal distribution – basic properties
Moment generating functions
Convergence of random variables, weak and strong law of large numbers, central limit theorem
Basic properties of discrete-time Markov chains and Poisson processes

The Probability notes [1] cover the required material at an appropriate level, as do Chapters 1–6 of Grimmett and Stirzaker [11]. An alternative introductory probability book is Ross [17]. An alternative source for the material on Markov chains is Chapter 1 of Norris [14].

The books by Rice [15] and DeGroot and Schervish [9] cover most of this material.

2.3 Statistics

Basic exploratory plots, e.g. histograms, boxplots, Q-Q plots
Maximum likelihood estimation
Properties of estimators – unbiasedness, consistency, mean squared error
Delta method, asymptotic normality of maximum likelihood estimator
Confidence intervals – exact intervals, approximate intervals using large sample theory
Hypothesis testing, types of error, including t-tests (basic, paired, two sample)
Likelihood ratio tests, asymptotic distribution of likelihood ratio statistic, applications to contingency tables, χ^2 goodness-of-fit tests
Basic single and multiple linear regression
Basic Bayesian statistics – conjugate prior and posterior, maximum a posteriori and expected posterior estimates, credible/highest posterior density intervals

The Statistics notes [2] cover the required material at an appropriate level. The books by Rice [15] and DeGroot and Schervish [9] provide good coverage of the required material.

Alternative references which are very suitable for MSc preparatory reading on statistics are Casella and Berger [7], Davison [8], Faraway [10], Wasserman [19]. The first five chapters of [19] also provide a summary of much of the prerequisite material on probability.

2.4 Simulation & Statistical Programming

Simulation methods: inversion, transformation, rejection. Importance sampling. Basic aspects of Markov Chain Monte Carlo.
The basics of programming, in either R or Python.

Note: although simulation, and statistical programming in R, is taught in the MSc, it is covered at a fast pace and therefore some previous experience is a distinct advantage and expected. If you have not used R before, then you should gain some familiarity with it before the start of the MSc. If you have used Python but not R, we are assuming you would be able to make the transition to R quickly – you may still want to gain some experience with R before the start of the MSc.

The above material is covered in the Simulation notes [3] and the introduction to R and programming in the Statistical Programming notes [4]. Books that cover this material at an appropriate level are Braun and Murdoch [6], and Robert and Casella [16].

See also the introduction to R:
https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf

You could also consider:
DataCamp tutorial https://www.datacamp.com/courses/free-introduction-to-r
Coursera R programming course https://www.coursera.org/learn/r-programming

References

[1] Probability notes and exercises, J. Martin. [copy; original]

[2] Statistics notes and exercises, N. Laws. [copy; original]

[3] Simulation notes and exercises, G. Nicholls. [copy; original]

[4] Statistical Programming notes and exercises, R. Evans. [copy; original]

[5] Linear Algebra Review and Reference, Z Kolter. [copy; original]

[6] W. J. Braun and D. J. Murdoch. *A First Course in Statistical Programming with R*. CUP, 2007.

[7] G. Casella and R. L. Berger. *Statistical Inference*, 2nd edition. Duxbury, 2008.

[8] A. C. Davison. *Statistical Models*. CUP, 2003.

[9] M. H. DeGroot and M. J. Schervish. *Probability and Statistics*, 4th edition. Pearson, 2012.

[10] J. J. Faraway. *Linear Models with R*, 2nd edition. CRC Press, 2015.

[11] G. R. Grimmett and D. R. Stirzaker. *Probability and Random Processes*, 3rd edition. OUP, 2001.

[12] D. W. Jordan and P. Smith. *Mathematical Techniques*, 4th edition. OUP, 2008.

[13] E. Kreyszig. *Advanced Engineering Mathematics*, 10th edition. Wiley, 2011.

[14] J. R. Norris. *Markov Chains*. CUP, 1997.

[15] J. A. Rice. *Mathematical Statistics and Data Analysis*, 3rd edition. Duxbury, 2006.

[16] C. P. Robert and G. Casella. *Introducing Monte Carlo Methods with R*. Springer, 2010.

[17] S. Ross. *A First Course in Probability*, 9th edition. Pearson, 2013.

[18] K. A. Stroud. *Engineering Mathematics*, 7th edition. Macmillan, 2013.

[19] L. Wasserman. *All of Statistics*. Springer, 2004.