

MSc Statistics Summer Projects

Instructions for candidates

1 Summary

The Statistics Project is an integral part of the course requirement for Masters degrees. It is an important opportunity to study in depth a topic of specific interest to you, and for you to apply knowledge and skills gained during the MSc programme. Your study should involve the application of one or more of the topics covered in the MSc. The project should be logically structured, well-researched and clearly written. There is a page limit of 50 singled sided sheets of A4; see §3.1. The project accounts for 25% of the MSc and you should allocate your time and effort accordingly. The deadline for submission is Friday 14th September 2007; see §4.

2 Supervisor's role

The project is intended to be an independent piece of work. Supervisors can be expected to offer advice on reading, guidance on the research approach and project structure, and some comments on the early draft sections. Supervisors cannot be expected to correct spelling and grammar, or to comment on a complete draft of the thesis. Supervisors will not always be available outside termtime – you should make sure you know when your supervisor is going to be available for face to face meetings and that you have their contact details for periods that they are not in the department.

3 Content and Approach

The project is an extended piece of written work. You must show your ability to organise your material clearly and logically. A statistically well informed reader (not necessarily a specialist in your chosen subject) should be able to follow your reasoning. You must demonstrate a thorough knowledge of the academic and professional literature relevant to the research topic, and a critical awareness of the contribution of different researchers. It is important that the project explains clearly the methods you have applied and the reasons for your choice of approach. You should show a clear understanding of any shortcomings of your study, in relation to both the methods employed and the quality of the data used. The examiners will pay close attention to the way in which you present your results. As trained statisticians, you are expected to be able to use tables, plots and graphs to present complete yet succinct summaries of your results. Your thesis should also provide the conclusions of your study and possibly suggestions for further work.

3.1 Layout

The elements described below are required parts of the thesis. Marking is done anonymously; please do not include your name on any part of the thesis. An example layout is provided on the MSc Statistics website.

- **Title page:** this should include “Department of Statistics 2007”, the title of the thesis, your candidate number, and the statement “Submitted for the Master of Science, London School of Economics, University of London”.
- **Contents:** This is made up of three elements each giving page numbers:
 - **Table of contents:** an ordered list of the sections and subsections of the project (include the bibliography and any appendices),
 - **List of figures:** an ordered list of diagrams,
 - **List of tables:** an ordered list of tables.
- **Summary:** a summary of the work no longer than 300 words.
- **Introduction:** introductory section that should include the aims of the study and a brief outline of the main sections of the thesis.
- **Main sections:** including a review of the literature, description of methods used and the results of the study.
- **Conclusion:** conclusions that can be drawn from your work, possible extensions and further work.
- **Bibliography:** plagiarism is viewed as a serious offence and you should take great care to cite all of the sources used in your project. More details are given in §3.3.

The inclusion of appendices is not encouraged. However, if necessary, appendices should go after the bibliography. Appendices may be used for technical material that does not belong with the bulk of the thesis or for describing the contents of attached computer disks.

3.2 Length and format of thesis

The thesis must be **maximum 50 singled sided A4 pages** including the bibliography and any appendices (but not including title, contents and summary page). Anything beyond the 50th page will be ignored. The page limit is set ensure that **only relevant material** is included. In particular, tables and diagrams should be included only when they are referred to somewhere in the text. Any devices used to try to get around the page limit will result in marks being deducted; please do not use weird fonts or shrunken diagrams.

The thesis should be word processed on A4 paper using only one-side of the page. The font size must be at least 11pt. The line spacing should be at least 1.5. The left-hand margin should

be 3.5cm and the right-hand margin 2.0cm. Pages should be numbered; use roman numerals (i, ii,...) for the title page, contents page and summary; use arabic numerals (1, 2,...) for the remainder of the thesis starting with 1 on the first page of the introduction. Sections should be numbered consecutively, starting with **1. Introduction**. Start new sections at the top of a new page. Tables and diagrams should be numbered. Pencilled diagrams are not acceptable. Do not use footnotes. Take care over style, grammar, punctuation and general presentation. In summary;

- title, contents, summary + maximum 50 single sided A4 pages,
- font size minimum 11pt,
- line spacing minimum 1.5,
- left-hand margin 3.5cm, right-hand margin 2cm,
- title, contents, summary – numbered using i, ii,...,
- remainder of thesis – numbered using 1, 2,...,
- number sections, new sections start on new pages,
- number tables and diagrams,
- no footnotes.

3.3 Plagiarism, references and bibliography

Plagiarism is taking someone else's work or ideas and passing them off as your own (adapted from Concise Oxford Dictionary definition). This arises in course work as sections of text lifted from books or internet sources and submitted as the student's own work. Plagiarism is a very serious offence that is easy to detect. Plagiarism cases are dealt with by the School's Assessment Misconduct Panel. A likely penalty is automatic failure of the project; for more details see <http://www.lse.ac.uk/resources/schoolRegulations>. You are strongly encouraged to read widely and assimilate ideas from as many sources as possible. However, when you use other people's work you must give a proper reference.

For **references** use the 'Harvard' or bracket system within the text. For example:

- "De Jong and Shephard (1995) put forward...",
- "Several authors (De Jong, 1989; Kohn and Ansley, 1989; Koopman, 93) make use of a..."
- "Harvey (1993, p.25) says, 'Whether or not a mixed process is stationary depends solely on its autoregressive part', but..."

In the **bibliography**, list works cited in alphabetical order by author. Give the author's surname, initials, year of publication, and title of work. For books, give the place of publication

and publisher; for journal articles give the name of the journal, volume and pages; for articles in published collections, give the name of the editor, title of book, place of publication and publisher; for online sources give the source organisation, the URL and the date accessed. The appropriate use of italics and bold is important. For example (in order book, journal article, from collection, online source):

- Harvey, A. C. (1993). *Time Series Models* (2nd edition). London : Harvester-Wheatsheaf.
- De Jong, P. and Shephard, N. (1995). The simulation smoother for time series models. *Biometrika* **82**, pp339-350.
- Bradshaw, J. and Deacon, A. (1986). Social Security, in P. Wilding (ed.), *In Defence of the Welfare State*. Manchester: University Press.
- Penzer, J. (2006) *MSc Statistics Projects – Instructions for candidates* [Internet] London School of Economics, London. Available from:(<http://stats.lse.ac.uk/penzer/>) [Accessed 18th September 2006]

Note, for online sources referencing styles are not very well established. Above I have used author, date (year in which item was last updated), title, place at which the item is published (or affiliation of the author), the URL and the day on which the information was accessed. Be careful to use only reputable sources – there is a lot of nonsense out there.

3.4 Computer programs and output

You may put a substantial amount of time and effort into writing computer programs or manipulating data. The code you have written is unlikely to be of interest to the reader. However, if you feel that it is important that the examiners see electronic material, you should include it on a disk. Think carefully about what will enhance the readers understanding of what you have done. The examiners will not appreciate receiving a disk full of files that are irrelevant or incomprehensible. If you decide to include a disk, its contents and any instructions for using the contents should be carefully described in an appendix. Do not include program listings in your thesis.

4 Submission and Assessment

The project should be submitted as soon as it is complete. The deadline for submission is **Friday 14th September 2007** . Extensions will not normally be granted; see §5. The requirement for submission are given below,

1. 2 bound copies of the thesis with any disks securely attached (ring binding in soft covers).
2. A copy of the title page and summary with your name written on (this is for the administrators to keep track of your thesis).

3. A completed copy of the plagiarism statement. This declaration is required for course work submitted as part of the formal assessment of a degree. It is a statement to the effect that you have read and understood the School's rules on assessment offences (Graduate Handbook) and that the work submitted is your own apart from properly referenced material.

Your project will be marked by two members of staff, one of whom will normally be your supervisor, and then read by the external examiner. Marks are awarded for the overall structure of the project as well as for the quality of the research carried out. The extent to which you use your own (good) ideas and initiative will also be taken into account. The examination results will be announced after the Examiners' Meeting in July. The project mark and final degree result will be announced after the project marks have been confirmed by the External Examiner (usually mid-November).

5 Penalties for late submission

The School regulations for late submission are given below.

1. If a student believes that s/he has good cause not to meet the deadline (for example illness), s/he should first discuss the matter with the course teacher and seek a formal extension from the Chair of the Examination Board. Normally extensions should only be granted where there is a good reason backed by supporting evidence (for example a medical certificate).
2. Any extension should be confirmed in writing to the student.
3. If a student fails to submit by the set deadline (or extended deadline as appropriate) the following penalties will apply:
 - for each day the submission is overdue a deduction of 10% will be made from the mark it would otherwise have obtained.

6 Organising your time

Work on your project will be interrupted by the exams. The exams should be your main focus from Easter until they are over (usually mid-June). That leaves Lent term and the summer to work on your project. Below is a list of some of the things you should be working on during Lent term.

1. **References:** Your focus should be on quality rather than quantity. A few good references is infinitely preferable to a lot of semi-relevant material. You should concentrate on the key contributions that have led to the current state of knowledge in the field. When you come to write up your literature review, tracing the flow of ideas in a rough chronological order is often a useful device.

2. **Aims:** You should work out a list of four or five aims for your project. These may take the form of questions to answer, properties to establish, hypotheses to test or theorems to prove/find counterexamples for. Your aims will evolve as you work on your project; you should not feel bound by them. However, aims are very useful if you start to feel lost.
3. **Software:** Unfortunately the department cannot provide computing facilities or software. You should use software that is available to you on your own machine or via the School's network.
 - Data analysis projects: decide on the data that you are going to use and get these data into a format suitable for the software that you have chosen. You may need to address issues such as missing values and outliers.
 - Simulation projects: it is very easy to get sidetracked. For each simulation, establish the questions you would like to answer before you start. Ensure that the experiment you are planning can adequately address these questions.
4. **By Easter:** you should have –
 - (a) written literature review and bibliography,
 - (b) the appropriate software and the knowledge of how to use it,
 - (c) your data in an appropriate format OR a clear idea of the simulations that you are going to perform,
 - (d) a skeleton structure of chapters for your thesis,
 - (e) a clear idea of what you are going to do over the summer.

7 Some suggestions about writing up

You should try to write up as you are going along. This does not mean spending a long time getting each section in a perfect finished form. However, you should keep a careful record of your progress. This is particularly important when working with computers; it is easy to generate a lot of files.

- When you do data analysis or perform a simulation experiment, write down what you did (method), what the results were and the names/locations of the relevant computer files.
- When you find a good reference add it to a master list of reference immediately.

Finally, no matter how long you spend on data analysis or simulations, it is the thesis that you hand in that will be assessed. You should make sure you leave sufficient time to do a good job of writing up. If your thesis is well written, the examiners will view your work more favourably. Here are some points to bear in mind.

- The thesis is a piece of scientific writing; it does not have to be entirely humourless but it is best to avoid the use of superlatives, exclamation marks and flowery language.

- Your aim is straightforward communication of what you have achieved to a statistically well informed reader; avoid the use of complicated or unnecessarily formal language.
- Avoid the use of vague language. Using definite terms will give your writing greater authority; for example, 'is' instead of 'can be' and 'will' instead of 'could'.
- Avoid using abbreviations eg. ie. etc...

If you would like to read more about getting your message across, try the article at <http://www.amstat.org/publications/jcgs/sci.pdf>.