Teaching Medical Statistics: What are Lectures For?

Nick Fieller
Department of Probability & Statistics
University of Sheffield, UK
11th May 2005

Introductions
- Who?
- Where from?
- What past & future involvement in Medical Statistics teaching?
  - & to whom?
- Why?
  - What are you looking for in this workshop?

Outline
- Background:
  - Sheffield’s courses: recent & past history
  - Why such detailed notes?
- How is the material used?
  - What else do the students see?
- Development of content
- Current content
  - Linking themes and notable omissions

Key fact:
- Sheffield’s Medical Statistics course is to
  - Level 3 BSc
  - Level 3 or level 4 MMath
  - MSc
- MSc includes
  - residential students who attend lectures
  - non-residential distant learners who don’t
- Single course to all groups
  - or, rather, single lecture series to all
  - assessment can be different for the groups

Undergraduates:
- ‘B’ level course
  - Bachelor level (c.f. ‘M’ level Master)
- Level 3 BSc
  - almost compulsory
  - Level 3 or level 4 MMath
  - use up one of 2 ‘unrestricted choices’
- MSc
  - residential students
  - almost compulsory
  - non-residential distant learners
  - compulsory (in final year, i.e. year 2 or year 3)
- Ideally should teach these groups separately

Course also given as part of taught component of a PhD in
Epidemiology at School of Public Health, Tampere
- Started in 1996 as a course on regular Mathematics programme and then incorporated into new PhD
  - 25 lectures in 2 weeks
  - printed notes etc essential
  - similar problems of teaching different groups
- Inherited Sheffield teaching in 1998 on [sudden] early retirement of predecessor
  - since known I had an electronic version of similar course
Some conflict

- MSc ↔ MMath/BSc
  - & MMath ↔ BSc
    - (level / mathematical sophistication)
- MSc take complete package of statistics courses: aim is to provide a thorough training as a professional statistician
- MMath/BSc take a pick & mix selection of all maths subjects (pure, statistics & applied)
  - attendance is voluntary
    - & we need the student numbers

Particular problem with MSc

- Residents attend lectures
- Non-residents don’t
- Course should be equivalent for both
- What are the lectures for?
  - Cannot disadvantage DLs
  - Should not waste time of RSs
    - Lectures are compulsory (Uni policy)
  - Cannot give material to only one group & not to the other
    - (close collaboration & communication between them)
- Inevitably this affects the way I teach

How did we get into this position??

- Somebody thought it was a good idea at the time
- But a brief history explains some things about the way this course is constructed

Background

- The Sheffield MSc in Statistics
  - Long established (mid 60s)
  - Broad based
    - leads to professional statistical career in any branch of application
      - (pharmaceutical, government, industry,...)
  - Successful
    - Jobs
    - PhD

Why Distance Learning?

- New EPSRC scheme of funding prompted a forward look
- Patterns of training changing
  - Career moves into Statistics
  - Skills updates
- Student priorities changing
  - indebtedness

The Aim:

Set up a DL Arm to the MSc that can be followed by part-time study
- 2 or 3 years
- will be of high quality –
  - fully equivalent to existing MSc
    - cannot afford to devalue our current MSc
- can be developed and sustained with available resources
  - some help from MTP funding available
Organizational Consequences
- Adaptation of existing courses
  - Lecture courses
  - Practical/professional skills courses
- Evolution not revolution
- Basic delivery technology
  - WebCT
  - Distance Learning Manager
  - Caitlin
- Teaching & Learning Support Unit Advice
  - University expertise tapped & highly valuable

Needs of Distant Learners
- Time, determination, equipment
- Encouragement
  - contact with department
  - contact with other students
  - weekly reports to Caitlin
- Route map
- Supplements to lecture material
  - amplification
  - reinforcement: five-finger exercises
- Feedback

Extras for Distant Learners
- Time, determination, equipment
- Encouragement
  - contact with department
  - contact with other students
  - weekly reports to Caitlin
- Route map
- Supplements to lecture material
  - amplification
  - reinforcement: five-finger exercises
- Feedback

Obliged to extend this good teaching practice to all MSc students & BSc

Why had we not been doing this already?

Our Solution
- Residential weeks
- Regular contact
- Course guide
- Discussion board
- Extended course materials
Our Solution (ctd)

- Use web for delivery to DLs
  - Material can be corrected, updated, extended
- Use technology supported by University
  - WebCT (Web Course Tools)
  - May not be ideal but technical support available
  - Allows us to concentrate on course content
- Don’t attempt to use Web as a teaching medium in its own right
  - RSs must have access to Web material (but maybe not BSc/MMath)

Clarification

- fully equivalent to existing MSc
  - cannot afford to devalue our current MSc
- DL version of MSc course must be same as existing course to residents
  - Same material
  - Same exercises
  - Same deadlines
  - Same pace of material
  - Can’t ‘read notes when they like’, keep in step
  - Same exams

Immediate Problems

- Lecture notes not self-contained for DL
  - Expand notes to include verbal explanations & additional diagrams & computer transcripts given in lectures
  - Provide simple task sheets each week
- Web pages for DLs useful for Residents but UGs have no access to these
  - I provided duplicate web pages for UGs, (data sets, solutions, …..)
  - some others did not do this

N.B. lecture notes for complete course distributed at beginning

- Alternative models:
  - Distribute lecture by lecture at start of lecture
    - Similarly for DLs (time release)
  - Fill-in-the-blanks notes
    - Students have to keep awake & fill in key numbers, proofs etc
- Overall, variety in department is good
  - No single model universally preferred

Introduced problems

- Danger of over-teaching
  - Too much expected by students
    - e.g. formal solutions expected for Task Sheets such as verify result 2.3 (‘easy to show that.’)
    - refuse firmly & resolutely
  - Students less tolerant of slips & misprints
    - & they can be detected!
    - no avoidance of public apologies

- What are lectures for??
  - if web material is self-contained?
  - Demonstrate more examples
  - Provide summary notes for lecture display
    - very substantial task
  - Must keep students occupied in the lecture
    - Study guide gives advice [ignored?]
    - Displaying copies of printed notes is disaster
    - PowerPoint slides must add value
    - Computer demonstrations must be extra
      - However…
**However**
- Must not be too valuable since DLs are not at lectures
- In lectures students will annotate notes from slides
- Put copies of slides on web for DLs
  - but not distributed in advance
- Sometimes produce extra slides
  - e.g. record of ad hoc software demo
  - annotate and tidied up
  - e.g. in answer to query by hand – can scan and put into .pdf and onto web

**Tablet PC (with OCR)???
- Weekly bulletin on discussion board
  - Outline material covered
  - Examples covered
  - Key points made from examples
- Computer demonstrations ok but must be used sparingly
  - students find it easy to lose thread
  - keep computer sessions short
    - 5 to 10 minutes max, perhaps 2 or 3 times per lecture
- PowerPoint animation can be used effectively – also should be used sparingly

**Extra teaching features:**
**Discussion Lists**
- more open discussion available to all
  - necessary to discourage direct emails & personal queries
  - reply only to discussion lists
- weekly news summaries for each course
  - material covered, queries raised
  - sometimes repeated verbally in lectures

**WebCT**
- Decision made early to use WebCT as medium for course delivery to DLs
- Web based system for
  - organizing/storing files e.g.
    - web pages
    - postscript or pdf documents
    - data sets
  - discussion lists, email,....
  - submitting assignments and maintaining records
  - &c., &c. .......
- Teaching & Learning Support Unit can provide technical help on WebCT
  - We discovered it is not bug free

---

These are copies of the OHP transparencies from the lecture on Monday 21/10/02. There are a couple of deceptively simple results relating to principal matrices.

First, I pointed out that the algebra involved on Slide 4, 6a (ii) is not something I would expect a reasonable human to do unseen and that putting the selection into the record guarantees that it won’t arise in an Open Book examination. It is included here because when it comes to the more...
### Example of Query
- Note ‘pseudo-TeX’ $e_{2i}$ for $e_{2j}$ in message
- reasonably ok even for more technical courses
- replies $e$ and $e_{2i}$

### Dealing with direct email query
- Feedback on exercises / assessed project
  - indicates detail of additional teaching input as consequence of distant learners
  - maybe students print message as substitute for listening & understanding

### Organization of Course
- Two Components
  - Survival Analysis
    - analyzing data on length of lifetimes, e.g. times of remission in leukaemia
  - Clinical trials
    - Experiments on human (and animal) subjects
    - Ethical issues, efficient use of subjects, etc
- Two lectures per week
  - no separate tutorial sessions
  - discussion of examples within lectures
- Full course on Survival Analysis first followed by Clinical Trials
  - Finnish version is alternating lectures since 2 or 3 lectures per day

### Approximately 10 lectures each
- Lectures built around PowerPoint slides
- Printed notes distributed in advance
  - Students expected to annotate
    - Leave plenty of white space in notes!
  - (many students have no blank paper)
- Task sheets distributed weekly
  - solutions in batches every 2 or 3 weeks
  - copies placed on web page
  - handout .pdf versions of PowerPoint slides placed on web page after ~every chapter

### Organization of course material
- Two sets of lecture notes (Survival & Clinical)
- Survival Chapters 1 – 4 main part of course
- Appendix 0: background maths
  - Maximum Likelihood Estimation
    - (but used only in a couple of places)
- Appendix 1 use of computer packages
  - SAS, SPSS, Minitab, S-PLUS
- Clinical Chapters 1 – 9 (~ 1 per lecture)

### Task Sheets & Exercises
- Task sheets:
  - ~ each week
  - simple quick short exercises/reading
  - reinforce / consolidate lecture material
- Exercises:
  - 3 sets during semester in weeks 5, 8, 10
  - Work submitted within 2 weeks will be marked and returned
  - See Study Guide
  - recommendations on time to spend

### Course web page
- http://www.shef.ac.uk/nickfieller
  - Click on Teaching & then on PAS6012/PAS361 Medical Statistics
- Lecture notes, task sheets, solutions & data sets available here after distribution in lectures
  - (I don’t keep back copies)
Summary so far:—

Courses have several objectives

- MSc — part of integrated programme in applied statistics
- BSc — optional course, should not be too mathematical (i.e. difficult)
- MMath — also optional but needs to be demonstrably ‘useful’ to attract customers [and some ‘nice’ maths?]

So, what do put into course to meet these various objectives?

Undergraduates:

- All entrants registered for MMath
  - take Pure, Statistics and Applied in year 1
  - take 2 or 3 of these in year 2
  - transfer to BSc if level 2 < 55% or ‘voluntarily’

- level 3 BSc are typically ‘less mathematical’
  - may not have taken any level 2 statistics
  - (~ 10 students)

- course should have minimal technical content
  - t-tests, χ²-tests, simple regression OK
    - but not likelihood ratio tests etc

MMath can take maximum of only TWO B-level courses over years 3 and 4, so some

- Choose Medical Statistics because of topic
  - signalled as major employer of statisticians to level 2 and 3 students
  - or because of reputation as less mathematical
  - or timetable fits (Tuesday and Friday at 11:00)
  - e.g. taken by many Dual Hons Maths/Accounting!
  - or like idea of ‘medical’

So, need course with

- not too much mathematics
- not too much technical statistics
- enough to give realistic route into employment

BUT

- must meet needs of a coherent MSc programme in Applied Statistics
- assessment can be different

- note that route into employment in medical statistics is almost essentially after an MSc
  - certainly in pharmaceutical companies

ASSESSMENT

- separate examinations for UGs and MSc
  - typically only slightly different
  - maybe different weightings in marking

- MSc [and soon MMath] have extended project for 30% of overall assessment

- currently large data set on time to failure of knee implants
  - many covariates / prognostic factors
  - Cox proportional hazards regression
  - not very satisfactory
  - parametric modelling worthwhile

CONTENT (derivation)

- Initially inherited from originators of course (Richard Kay, Ian Dunsmore)

  - based on books by Pocock, Altman, …

  - Some modification from course

    Statistics for Medicine
    - given by Statistical Services Unit since ~1990 aimed at medical researchers
    - (I teach on this occasionally)

    - Occasional additions / examples from regular
    Statistics Clinic run by SSU for Sheffield Central teaching Hospitals
    - e.g. sample size questions
Teaching Medical Statistics: What Are lectures For?

11/5/2005

Some input on content from tutees going to interviews
- I maintain a [restricted] web page of questions
Some input from comments from past graduates in pharmaceutical jobs
- relative risks, odds ratios inserted last year
- next year measures off agreement (Bland-Altman plots etc) to be included
Some input from other sources
- e.g. chapter on multiplicity from a ICH 9 workshop

CONTENT (Clinical Trials topics)
- BSc motivated
  - History (streptomycin trial & Salk polio vaccine)
  - Drug development, Phases I, II, III & IV etc
  - Placebo effect, single, double blinding
  - Ethics, Treaty of Helsinki, ethical committees
  - Protocols, protocol deviations
  - Basic designs: parallel groups, in series, factorial, sequential
- All of these are non-mathematical.
  - All yield common interview questions
  - None make good exam questions
  - NB: all Sheffield exams are ‘open-book’

Clinical trials topics for exam questions
- Cross-over trials
  - tests for carry-over, period & treatment effects
    - only use t-tests or χ²-tests (i.e. 1st year material)
    - (perhaps extend to non-parametric Mann-Whitney tests, not seen otherwise)
  - exam questions for MSc can be extended
    - e.g. benefit of covariates
    - e.g. unequal numbers in groups and CIs
    - e.g. drop-outs
- Randomization lists
  - common sense skills, partly covered in level 1
    - minimization/dynamic allocation methods
    - more substantial for MSc

r × c tables
- Mantel-Haenszel test (& Simpson’s paradox)
  - not mathematically demanding
  - plenty of straightforward exam questions
    - in total population drug A better than drug B
    - but B better than A for subgroup of men
    - & B better than A in women
  - which drug do you recommend?
- Odds ratios (retrospective studies)
- Relative risks (prospective studies)
- Matched pairs (McNemar’s test)

Sample size / power calculations
- covered informally in level 1 course
- extend to one and two sample problems of means and proportions
  - perhaps do two of the four cases in detail, leaving others for examining (e.g. MSc)
  - plenty of routine questions with allowing for drop-outs
    - e.g. “if required to recruit for specified power at difference of d, allowing for 15% drop outs?”
    - but drop out really 20% so what power?
    - what if only record ‘improve/not improve’?
  - plenty of software S+, nQuery, SPSS, ....
    - (which produce different answers!!)

Clinical trials topics for exam questions (ct²)
leads to logistic regression models
- very informal approach, no theory
  - theory covered in Generalized Linear Models
    - M-level course (& compulsory for MSc)
  - interpretation of coefficients in computer output
    - “|estimate/s.e.(estimate)| > 2 ⇒ effect”
    - confidence intervals for odds ratios
      - (e.g. at specific combinations of covariates)
    - idea of interactions
  - NB no diagnostics, no model building
    - but can be covered for MSc in projects or separate exercises
Survival Analysis topics
- Discursive:
  - censoring — intuitive ideas, information contained in censored values
  - life tables
  - (tedious? not good for exams but necessary)
- Examinable topics:
  - Kaplan-Meier estimation
  - log-rank tests

Examinable topics:
- Kaplan-Meier estimation
- log-rank tests

Survival analysis topics for exam questions (ctd.)
- exponential models
  - estimation of parameter in censored case by maximum likelihood (ok for MSc)
  - not accessible to BSc (not many MMath)
  - since not previously seen likelihoods (!!!)
  - nor any technique of estimation
  - provide appendix (for those who want to know)
- one and two sample tests
  - appealing to general ML theory
  - i.e. recipe-based for UGs

Topics for the well-educated Applied Statistician
- i.e. topics which arise especially in medical statistics and have general applicability
- Multiplicity
  - subgroup analyses (M & F, old & young etc)
  - illustrate with Richard Peto’s Signs of Zodiac
  - multiple endpoints (systolic and diastolic b.p.)
  - commend multivariate analyses
    - (not accessible to BSc)
- interim analyses
  - paradox example (ex 6.4.3)

Bjørn Andersen (1990)
Methodological errors in medical research
- Preface declares the author's wish that there will never be a 2nd edition
  - now out of print and 1 buyer waiting on Amazon
- Long chapter provided in notes (35+ pages)
  - (from a workshop on ICH 9 Points to Consider)
- Difficult to convince some students that problem is not confined to medical statistics
  - “Oh, I decided not to use multiplicity here…”
  - (after testing English, Scottish, Welsh, Irish and UK pound coins for left-handedness separately)
Educational topics (cont'd)

- Randomization
  - digress to randomization tests
- Protocols
  - planning analysis before data collection
  - sample size calculation
  - handling drop-outs
    - ('intention to treat' / 'per protocol')
- Experimental Design
  - for many students this all they encounter
- Regression models
  - good to see unity in modelling even though underlying mathematics is very different

Summary (update)

- Clinical Trials
  - Lots of discursive topics —
    - good for general discussion, interviews etc, not easy to examine
  - plenty of worthwhile topics requiring little technical statistical background
    - easy to examine
- Survival analysis
  - more technical
    - highly technical if done 'properly'
- Both areas provide good motivation for generally applicable methodology

Preference is for Clinical Trials section to be first, Survival second

- logistic regression 'easier to explain' than Cox
- more familiar with other statistical techniques
  - maybe courses on linear models, inference, S+
- However, for organisational reasons have reversed with survival analysis first
  - allows MSc project to be set in week 5
    - need to have MSc assessed projects evenly spaced
    - also number & timings of example sheets

Notable Omissions

- epidemiological [observational] studies
  - case-control studies
  - cross-sectional studies
  - cohort studies
- meta analysis (good book by Anne Whitehead)
  - forest plots
  - funnel plots
- agreement studies
  - kappa, Bland-Altman plots
- population studies
  - age standardization etc

Other organisational details

- computing originally based on Minitab
  - and SPSS for Cox regression
    - no Cox regression in Minitab
    - good for parametric models
- (Finnish version entirely on SPSS)
- now converted to S+
  - S+ introduced in 1st year in 2002
  - (so level 4 started with Minitab)
  - conversion not very smooth
    - Eventi & Rabe-Heskth and Venables & Ripley
      - good for syntax and facilities
    - menus very limited
    - but has allowed extension to parametric models
  - R might be preferable


- ~1995 first use 'interpret Minitab output' questions in examinations
- ~1996 introduction of extended coursework project for MSc courses
  - 30% of credit for course
    - (max permitted by University)
  - used in Medical Statistics from 1998
Current content

NB 20 contact hours for lectures & examples
- Guidelines are to spend further 180 hours self-study
  - task sheets
  - exercises
  - lecture notes
  - exam revision
- Includes 1 reading week with no lectures
- i.e. ~4 hours extra per week + revision
  - rarely done

AstraZeneca Statistical Challenge

AstraZeneca plc have invited Sheffield to provide a team of five undergraduate students interested in Medical Statistics for the above event at their site at Mereside, Alderley Park (nr Macclesfield). The day will include talks about careers in Pharmaceutical Statistics and a site tour. Open to all UGs (including Erasmus visitors). Travel expenses will be paid by AZ.

Date: Tuesday 25th November (9:30-4:30)
email me by 2:00pm Tuesday 14th October if interested

Announcements

RSS Local Group meeting

Thurs 3rd 5:00. Sheffield Hallam, Room 7138 Stoddart Building (~opposite Novotel)
Tea 4:30, Wine &c. 5:45, Deli Bar

PRACTISING STATISTICIANS
James Godfrey, MMR Food & Drink Research Worldwide Limited
Mark Pilling, Statistical Services Unit, Sheffield University
Lucy Spooner, Ehsbrann Discovery
John Spooner, SAS

Interim Discussions & Working Groups

Working Groups
- each group to identify 3 points for discussion
  - preferably with initial comments
- possible topics:
  - discursive vs examinable topics
  - assessment generally
  - computing (but see later)

Working Groups: Assessment
- What are the objectives of assessment?
- What ways can these be achieved?
- Does it depend on target group?
- Should account be taken of custom & practice in other modules?
- How can assessment be fair to all?
- A O B??

Miscellaneous practicalities

- Computer packages
  - S-plus
    - can handle most analyses required but not from menus
  - R
    - better for expert user (more esoteric libraries)
  - Minitab
    - no Cox regression
  - SPSS
    - no command line programming
  - SAS
    - Industry standard
Computer packages

- StatsDirect
  - Excel add-in
  - cheap
    - (£99 for 3 years including upgrades)
  - very comprehensive
    - meta analyses etc

Other packages??

- Stata
- MLWin
- Genstat
- ???
- ???

Teaching to non-specialists

- SSU course Statistics for Medicine
  - commercial course, run 5 times a year
  - 3 modules Monday midday – Friday
    - Basic Statistical Concepts (1½ days, 10 CME points)
    - Statistics for the Initiated (2 days, 13 CME points)
    - Designing a Clinical Study (1 day, 6 CME points)
  - most people take modules separately

Teaching to non-specialists

- SSU course Statistics for Medicine
  - commercial course, run 5 times a year
  - 3 modules Monday midday – Friday
    - Basic Statistical Concepts (1½ days, 10 CME points)
    - Statistics for the Initiated (2 days, 13 CME points)
    - Designing a Clinical Study (1 day, 6 CME points)
  - most people take modules separately

Finnish version

- ~ 15 PhD Epidemiology students
  - a few with good mathematics
  - most with little mathematics but reasonable statistics
  - some level 3 Mathematics students
  - downplay mathematical side
    - skip slides on LRTs & derivation of Cox regression etc

Summary & Conclusions & Future

Topics for discussion

- ……
- ……
- ……

Future:

- What do you think will be key changes?
- By when?
- How can we improve knowledge & teaching?
- ……