Statistics in Medicine (3 units)

Statistics in Medical Practice (M12)

Lecturers from the MRC Biostatistics Unit

This part of the course includes three modules covering a range of statistical methods and their application in three areas of biostatistics.

A. Stochastic Models for Chronic and Infectious Diseases [4 Lectures] **(C. Jackson, A. Presanis & D. De Angelis)**

Continuous-time multi-state and Markov models: properties and quantities of interest, and fitting models to individual disease history data. Applications to modelling the onset and progression of chronic diseases. Multi-state modelling to estimate incidence of infectious diseases from population-level prevalence data. Backcalculation methods for the estimation of incidence of disease with long incubation periods. Dynamic modelling of infectious disease transmission.

B. Causal Inference [4 Lectures] **(S. Burgess)**

It is well known that "correlation is not causation". But how then do you assess causal claims? Is it possible to show that X is a cause of Y? What does it even mean to say that X is a cause of Y? In this module, we introduce definitions of causal concepts, starting with the work of Rubin, Pearl, and Robins, and discuss practical approaches for assessing causal claims from observational data.

C. Design and Analysis of Randomised Trials [4 Lectures] **(M. Pilling, S. Villar, D. Robertson & S. Seaman)**

Sample size estimation for clinical trials; Adaptive and multi-stage designs (including group-sequential designs). Treatment effect estimation following a group-sequential trial. Types of randomisation procedures. Optimal response-adaptive procedures. Handling missing data: classification of missingness mechanisms, maximum likelihood, and multiple imputation.

Statistics in Medicine (3 units)

Analysis of Survival Data (L12)

P. Treasure

This part of the course includes three modules covering the fundamentals of time-to-event analysis with applications to cancer survival.

D. Time-to-Event Analysis [4 Lectures]

'Survival analysis' is generalised to *time-to-event* analysis. The implications of event times which are unknown or in the future (*censored* data) are discussed. Time-to-event distributions are introduced and their parametric (maximum likelihood) and non-parametric (*Kaplan-Meier*) characterisations are described. Methods for comparing two time-to-event distributions (as in a clinical trial of an active treatment versus a placebo) are derived (*log-rank* test).

E. Modelling Hazard [4 Lectures]

The *hazard* function (instantaneous event rate as a function of time) is defined. It is shown how the hazard function can naturally be used to model the effect of explanatory variables (such as age, gender, treatment, blood pressure, tumour location and size...) on the time-to-event distribution (*proportional hazards* modelling). Model checking procedures are introduced with an emphasis on excess event (*Martingale*) plots.

F. Population Cancer Survival Analysis [4 Lectures]

(Provisional topics for 2022:) Analysis of survival data from real-world cancer studies is complicated by patients also being at risk from other causes of death. Methods of dealing with more than one cause of death are presented for the cases (i) the cause of death is known (*competing risk* analysis) and (ii) the cause of death is unknown (*net survival*). The conceptual difficulties inherent in the notion of a cancer survival distribution relevant to a particular calendar year (e.g. 2019) are addressed: *period* survival analysis.

Statistics in Medicine (3 units)

Additional Information

Prerequisites

Undergraduate-level statistics and probability: including analysis and interpretation of data, maximum likelihood estimation, hypothesis testing, basic stochastic processes.

Literature

There are no course books, but relevant medical papers may be made available before some of the lectures for prior reading. A few books to complement the course material are listed below.

1. Armitage P, Berry G, Matthews JNS, *Statistical Methods in Medical Research.* Wiley-Blackwell, 2001. [A good introductory companion to the whole course]
2. van den Hout, A, *Multi-State Survival Models for Interval-Censored Data.* Chapman and Hall, 2016 [Module A]
3. Keeling, M. J., & Rohani, P. *Modeling Infectious Diseases in Humans and Animals.* Princeton University Press, 2008 [Module A]
4. Burgess S, Thompson SG, *Mendelian Randomization: Methods for Using Genetic Variants in Causal Estimation* Chapman and Hall, 2015 [Module B]
5. Senn, S. *Statistical Issues in Drug Development.* 3rd Edition, Wiley, 2021. [Module C]
6. Jennison C, Turnbull B, *Group Sequential Methods with Applications to Clinical Trials.* Chapman and Hall, 2000. [Module C]
7. Rosenberger, William F., and John M. Lachin. *Randomization in clinical trials: theory and practice.* John Wiley & Sons, 2015. [Module C]
8. Cox DR, Oakes D, *Analysis of Survival Data.* Chapman and Hall, 1984 [Modules D, E, F: the classic text]

9. Collett D, *Modelling Survival Data in Medical Research.* CRC Press, 2015 [Modules D, E, F: modern, applied, supports and extends lectures.]

10. Aalen OO, Borgen Ø, Gjessing HK, *Survival and Event History Analysis: A Process Point of View.* Springer, 2008 [Modules D, E, F: excellent modern approach]

Additional support

[Modules A, B, C] A two-hour example class, supported by question sheets and solutions, will be given at the start of Lent term.

[Modules D, E, F] A two-hour example class, supported by question sheets and solutions, will be given in each of the Lent and Easter Terms. A two-hour revision class will be held just before the examination.