HEALTH DATA SCIENCE

Instructor: N.DEMIRIS

Course Code: 61231
Course Type: Compulsory of Course Group 1
Course Level: Graduate (MSc)
Year of Study: A'
Semester: 2nd
ECTS: 7,5
Language: English

Course Description

This course has the following content

Basic concepts in survival analysis, definitions, hazard and survival functions, relationships, parametric methods, likelihood function, Exponential and Weibull Models, applications in R

Non-parametric methods: Kaplan-Meier estimator, Greenwood and Nelson-Aalen estimator, graphical goodness of fit, log rank test.

Regression models, Cox proportional hazards, Survival Analysis theory, counting processes, applications in R

Martingale/Deviance/Schoenfeld residuals. Heterogeneity and frailty models, LASSO and elastic net, hyperparameter selection via cross-validation, applications in glmnet

Non-proportional hazards models, additive hazards, accelerated failure time, proportional odds, competing risks and (non-)identifiability, multi-state models

Prospective and retrospective studies, (non)interventional, AR, RR and OR, equivalence of OR. Screening tests, PPV/NPV and sensitivity/specificity

Clinical trial design and analysis, protocol, sample size calculations, phase I, MTD, 3+3 design, Phase II, safety and efficacy, phase III. Real world vs (and/or) randomised data.

CRM+adaptive designs, Simon 2-stage design, Bayesian and historical/synthetic controls

Meta analysis, systematic reviews, fixed effects, heterogeneity, random effects, publication bias, funnel plots, indirect treatment comparisons and network meta analysis, example applications in health economics using ICER, INB and CEAC.

Evidence synthesis and conflict diagnostics

Introduction to Epidemic models, main results, vaccination and control.

Basic stochastic models, branching processes and coupling, functional LLN and CLT, connections between the different types of model.

Inference for chain binomial models using MCMC. Inference for deterministic models using HMC.

Heterogeneity, multiple age-groups, contact matrices, epidemics among households.

Epidemics on networks

Prerequisites

Probability and Statistical Inference, Computational Statistics

Target Learning Outcomes

At the end of the course students will have knowledge of the basic principles, methods and implementation tools of the main data science techniques that are being used in the analysis of health data.

Recommended Bibliography

No single textbook covers the material presented. There will be lecture notes and course code from books in Biostatistics, Survival Analysis and Epidemic models as well as relevant research papers

Teaching and Learning Activities

Classroom teaching and assignments.

Assessment and Grading Methods

The main course assessment will be based upon assignments that will be prepared, given as reports and presented on an individual basis.