Stockholm University

Gebrenegus Ghilagaber
(PhD, Professor)

STOCKHOLM UNIVERSITY
Department of Statistics
Spring term 2023

16 January 2023

Course Description for *Analysis of Survival Data - with Demographic Applications*, 7.5 Higher Education Credits (HEC) at Basic-level, Course Code: ST303G

COURSE CONTENTS

The course provides introduction to models and methods used in the analysis of survival data -with applications in the social sciences. The contents of the course can be applied in many areas like Demography, Economics, Public Health and Medicine, etc. The course also provides introduction to some software or program languages that are appropriate for the analysis of survival data. The topics that will be covered in more details include:

- Special features of survival data such as censoring,
- Functions of survival time,
- Kaplan-Meier and Life-table methods for estimating survival functions,
- Comparison of survival functions,
- Some theoretical distributions for survival time: exponential, Weibull, lognormal, gamma, etc.
- Survival data with covariates: Cox proportional hazards model, parametric models for continuous survival data and models for grouped survival data.
- Models for dichotomous variables: logistic regression.
- Survival data with multiple causes of failure: Competing risks models
- Overview to multilevel models, unobserved heterogeneity, and selection bias

LEARNING GOALS

For successful completion of the course students should be able to:

- describe and explain basic concepts, functions, and distributions for survival data;
- compute and compare survival functions for different groups;
- model associations between survival functions and explanatory variables and test hypotheses concerning models and parameters;
- use one or more computer programs and/or software to analyze continuous, grouped, and binary survival data;
- identify potential limitations of basic survival models and describe alternative advanced models.

TEACHING FORMAT

Teaching consists of 12 lectures (L1 - L12) and 2 computer-sessions (C1-C2) according to separate schedule. With reservations for changes, the following topics will be covered during the respective lectures (see Table below).

Lecture	Contents (from the Literature)
L1	Basic Concepts: Censoring, Survivor function, Density function, Hazard function, & their relationships (Allison, Chapters 1 and 2; Moore, Chapters 1 & 2)
L2	Non-parametric estimation: Product-Limit (Kaplan-Meier) and Life-Table (Actuarial) Methods (Allison, Chapter 3, Moore, Chapter 3)
L3	Non-parametric comparisons of two or more survival curves (Allison Chap. 3, Moore, Chap. 4)
L4	Parametric survival distributions: Exponential, Weibull, Lognormal, Gamma, etc. (Moore Chapter 2, Section 2.4 and Chapter 10).
L5	Models of Regression type: Cox's Proportional Hazards model. (Allison Chapter 5; Moore, Chapters 5 and 6)
L6	Cox's Proportional Hazards model. Model assumptions & diagnostics (Moore, Chap. 7)
L7	Models of Regression type: Parametric models (Allison Chapter 4, Moore, Chapter 10, Ghilagaber, 2018; Liang & Ghilagaber, 2022)
L8	Models of Regression type: Parametric models (contd...)
L9	Models of Regression type: Multiplicative (Log-Linear) models for grouped survival data (Moore, Chapter 12)
L10	Models with dichotomous outcome variable: logistic regression. (Allison, Chapter 7)
L11	Special topics such as competing risks, unobserved heterogeneity and selection bias. (Allison, Chap. 6 & 8; Moore, Chapters 9 & 11, Ghilagaber 2005; Ghilagaber & Wänström, 2015; Munezero & Ghilagaber, 2022; Ghilagaber, Akinyi Lagehäll & Yemane, 2022).
L12	Special topics (contd...), Summary, Revision, Evaluation **Handing out Take-home Examination (to be handed-in by Monday 20 March 2023 at 23:59)**
Tuesday 4 April	Latest date for returning corrected exam-papers
Thursday 20 April	Handing-out re-examination (to those concerned) – to be handed-in by Wednesday 26 April 2023.

The two computer sessions will cover the following topics:

C1	Introduction to some procedures in R (and/or SAS)
C2	More (advanced) procedures in R (and/or SAS)

Lecture notes and other relevant material can be handed-out (or made available in Athena) in connection with the lectures and computer sessions.

COMPULSORY ATTENDANCE

- Lecture 1 where general information about the course is provided
- Computer Session 1

EXAMINATION AND GRADING

Note that students should register for exam/s in Ladok <u>no later than 10 days before the starting date of the examination</u>.

Knowledge control on learning outcomes is assessed through a take-home examination. The home-exam, is to be worked individually, without any help from others, and consists of both theoretical questions and empirical analysis of data. Oral exam can be used to complement the written take-home examination if the need arises. The home exam is preceded by 2-3 optional "preparatory" assignments which may be solved individually or in groups (2-3 persons per group).

The final grade is based solely on the individual home-exam and according to the following criteria:

Grade	Criteria (Requirements)
A	The student has reached the learning goals excellently: This requires at least 55 points (out of 60) in the written exam.
B	The student demonstrates a very good ability in reaching the learning goals: requires 50-54 points (out of 60) in the written exam.
C	The student demonstrates a good ability in reaching the learning goals: requires 45-49 points (out of 60) in the written exam.
D	The student has reached the learning goals in a satisfactory way: requires 38-44 points (out of 60) in the written exam.
E	The student has reached the learning goals adequately: requires 30-37 points (out of 60) in the written exam.
Fx	Inadequate but the student can reach the learning goals after some make-up work: corresponds to 25-29 points (out of 60) in the written exam.
F	Totally inadequate, corresponds to 0-24 points (out of 60) in the written exam.

- A grade of E or higher is required in order to pass the course.
- Students who have attained the grade Fx or F on an examination are entitled to at least four additional examinations to achieve at least grade E as long as the course is offered.
- Neither Fx nor F are passing grades and both demand some form of re-examination.
- Students who have received at least grade E on an examination may not retake another examination in an attempt to achieve a higher grade.
- Students who have received the grade Fx or F on an examination on two occasions by the same examiner have the right to request that a different examiner be appointed to set the grade of the examination on the next possible occasion. The request must be in writing and sent to the head of the department. The examination denotes all compulsory elements of the course.

COURSE LITERATURE AND OTHER TEACHING MATERIALS

- Moore, D. F. (2016), *Applied Survival Analysis Using R*. Springer Verlag. Available online via SU-library
- Allison, P. D. (2010), *Survival Analysis using the SAS System*. 2nd edition, SAS Institute Inc. Cary, NC.

Some additional materials may be suggested during the course (handed out or made available via Athena/Library). Some examples are:

- Ghilagaber, G. (2005), Incompatibility between Hazard and Logistic Regression in Modeling Competing Risks. *Quality and Quantity - International Journal of Methodology*, 39, 37-44.
- Ghilagaber, G., and Wänström, L. (2015), Adjusting for selection bias in assessing the relationship between sibship size and cognitive performance. *Journal of the Royal Statistical Society - Series A*, 178 (# 4), 925-944.
- Ghilagaber, G. (2018), Environmental recidivism in Sweden: distributional shape and effects of sanctions on duration of compliance. *Quality and Quantity - International Journal of Methodology*, 52 (# 2), 869-882.
- Liang, Y., Ghilagaber, G. (2022), Bayesian Survival Analysis with the Extended Generalized Gamma Model: Application to Demographic and Health Survey Data. Pages 287-318 in Chen, D. G., Manda, S. O. M., and Chirwa, T. F. (eds.), *Modern Biostatistical Methods for Evidence-Based Global Health Research*. Emerging Topics in Statistics and Biostatistics. Springer, Cham. https://doi.org/10.1007/978-3-031-11012-2_11
- Munezero, P., Ghilagaber, G. (2022). Dynamic Bayesian Modeling of Educational and Residential Differences in Family Initiation Among Eritrean Men and Women. Pages 319-337 in Chen, D. G., Manda, S. O. M., and Chirwa, T. F. (eds.), *Modern Biostatistical Methods for Evidence-Based Global Health Research*. Emerging Topics in Statistics and Biostatistics. Springer, Cham. https://doi.org/10.1007/978-3-031-11012-2_12
- Ghilagaber, G., Akinyi Lagehäll, A., Yemane, E. (2022). Stratified Multilevel Modelling of Survival Data: Application to Modelling Regional Differences in Transition to Parenthood in Ethiopia. Pages 431-456 in Chen, D. G., Manda, S. O. M., and Chirwa, T. F. (eds.), *Modern Biostatistical Methods for Evidence-Based Global Health Research*. Emerging Topics in Statistics and Biostatistics. Springer, Cham. https://doi.org/10.1007/978-3-031-11012-2_17

COURSE INSTRUCTOR AND EXAMINER

- Gebrenegus Ghilagaber (Professor): Room A4602, Bldg. 4, 6th floor (Albano Campus (Alvanovägen 12), Tel. 08-162983, E-mail: Gebre@stat.su.se
  Consultation hours: In connection with lectures or by appointment.