

# Event-History Analysis: Regression for Longitudinal Event Data

7.5 credits, Spring 2019

## Syllabus

Last Updated: March 14, 2019

### Instructors

Sven Drefahl ([sven.drefahl@sociology.su.se](mailto:sven.drefahl@sociology.su.se))

Eleonora Mussino ([eleonora.mussino@sociology.su.se](mailto:eleonora.mussino@sociology.su.se))

Matthew Wallace ([matthew.wallace@sociology.su.se](mailto:matthew.wallace@sociology.su.se))

### Entry Requirements

*Bachelor's degree* or equivalent, and *English 6* or equivalent, and *Quantitative Methods in the Social Sciences I – Introduction to Regression Analysis* or equivalent.

### Course description

This course is an introduction to event-history analysis (also known as survival analysis, hazard regression, intensity regression, or duration data analysis) and is given at the advanced (Masters / PhD level). Duration data is commonly used to address many research questions in demography, social sciences, and epidemiology. Examples of such questions are: Which factors influence how long people live, how long they stay unemployed, or when do they start a family? This course introduces the techniques for analyzing such questions and data and covers univariate and basic multivariate (regression) methods for analysis of duration (event-history) data. Students also learn data management skills that are specific to conducting event-history analysis in Stata.

### Expected learning results

By the end of the course, students should be able to:

- Describe the basic concepts of event-history analysis
- Understand the link between event-history analysis, basic demographic methods and regression analysis
- Recognize the type of research questions for which event history analysis would be a suitable method
- Interpret studies that have used basic event-history methods
- Reflect on the assumptions, problems and limitations of event-history methods

Using *Stata*:

- Transform data into the basic data layout of event history analysis
- Analyze time-dependent univariate and multivariate relationships
- Specify appropriate regression models using time-constant and time-varying explanatory variables
- Interpret results obtained and communicate them to experts and non-experts alike

### **Teaching**

The course is given half-time over a 9 week period. Coursework and examination consist of lectures, research output from demographic studies, and computer-based exercises. The exercises are done with the latest Stata statistical package using data from the European Social Survey (round 3). Students will receive feedback after each computer exercise.

### **Examination**

Examination is based on active participation including a short study presentation, computer exercises, and a take-home exam. Students are graded according to **15** separate evaluations (specified below), and each is graded Fail (0 points), Pass (1 point), Good (2 points).

**Participation (1. – 2.)** is evaluated by the student's attendance in lectures and computer exercises, including discussion of the assigned readings (**1.**) and a brief oral presentation of a particular research question that can be addressed with event-history analysis (**2.**).

**Each of the five computer exercises** is evaluated (**3. – 7.**). The computer exercises should include proper solutions to the assigned problems and clear presentations of the Stata-syntax ("do-files") and the output.

The **take-home exam (8. – 15.)** consists of a small independent study using event-history analysis with data provided by the instructors. The following aspects are evaluated:

- 8.** Argument for research question and choice of data and method
- 9.** Data description, manipulation and variable construction)
- 10.** Stata-syntax ("do-files") that are clear and easy to follow
- 11.** Descriptive analyses
- 12.** Appropriate model specification for multivariate analysis
- 13.** Execution of multivariate analysis
- 14.** Presentation of results
- 15.** Interpretation of results (including limitations)

The maximum number of points a student can attain is 30. In addition, extraordinary performance in any of the aspects can be rewarded with up to 2 extra points that can compensate for any shortcomings.

The final course grade is based on the following criteria:

A (Excellent) = 28-30 points

B (Very good) = 25-27 points

C (Good) = 22-24 points

D (Satisfactory) = 19-21 points

E (Sufficient) = 15-18 points

Fx (Insufficient) = Fail for one or two of the aspects specified above

F (Fail) = Fail more than two

Examination according to this syllabus can be completed up to three semesters after it expires.

### **Plagiarism, cheating and unauthorized cooperation**

As part of your responsibility as a student, you have to be familiar with the rules for examination. Detailed information is available at the institution's and Stockholm University's website [www.su.se/regelboken](http://www.su.se/regelboken). Teachers are obliged to report suspicion of cheating and plagiarism to the Director of Studies and the Disciplinary Board. An example of plagiarism is to formally or almost verbatimly write a text (also applies to single sentences) and without indicating where this comes from. This also applies to texts you have previously written (self-plagiarism). Having study groups together is encouraged, but when it comes to exam assignments, you must be careful to work for yourself (unless otherwise stated) in order not to risk it being counted as unauthorized cooperation.

### **Literature**

The course book and other suggested books are listed here, while all other readings for each session are listed in the course schedule below. Most readings can be accessed online from an SU computer. Readings that are not available online can be found in the course compendium (.pdf file available from Mondo).

### **Course book**

- Blossfeld, H-P, Golsch, K. & Rohwer, G. 2007. Event History Analysis Using Stata. Lawrence Erlbaum. (hereafter, BGR)

### **Other suggested books:**

- Cleves, M., Gutierrez, R.G., Gould, W. & Marchenko, Y.V. 2010. An Introduction to Survival Analysis Using Stata. Stata Press.
- Kleinbaum, David G./Klein, Mitchel (2005): Survival Analysis: A Self-Learning Text. 2nd Edition. New York: Springer.

### **Other reading material**

(Fulltext access for some items might be limited to SU computers)

1. Allison, P. (1984) Event History Analysis. Sage, pages 14-22  
(In course compendium)

2. Andersson, G. (1998). Trends in marriage formation in Sweden 1971-1993. *European Journal of Population* 14(2): 157-178.  
Fulltext pdf: <http://rdcu.be/Fh4G>
3. Andersson, G. and D. Philipov (2002). Life-table representations of family dynamics in Sweden, Hungary, and 14 other FFS countries. *Demographic Research* 7(4): 67-144.  
Fulltext pdf: <http://dx.doi.org/10.4054/DemRes.2002.7.4>
4. Andersson, G. and S. Drefahl (2017). Long-distance Migration and Mortality in Sweden: Testing the Salmon Bias and Healthy Migrant Hypotheses. *Population Space and Place* 23(4)  
Fulltext pdf: <http://doi.org/10.1002/psp.2032>
5. Breslow, N.E. and N. E. Day (1980). *Statistical Methods in Cancer Research, Volume 1 - The Analysis of Case-Control Studies*. Lyon: International Agency for Research on Cancer. Selected pages of Sections 2.1 through 2.7.  
(In course compendium)
6. Cleves, M., Gutierrez, R.G., Gould, W. & Marchenko, Y.V. 2010. *An Introduction to Survival Analysis Using Stata*. Stata Press, pages 365-391  
(In course compendium)
7. Coviello, V. and M. Boggess (2004). Cumulative incidence estimation in the presence of competing risks. *The Stata Journal* 4(2): 103-112. Fulltext pdf: [http://ageconsearch.umn.edu/bitstream/116230/2/sjart\\_st0059.pdf](http://ageconsearch.umn.edu/bitstream/116230/2/sjart_st0059.pdf)
8. Drefahl, S. (2010) How Does the Age Gap Between Spouses Affect Their Survival? *Demography*, 47(2): p. 313-326.  
Fulltext pdf: <http://dx.doi.org/10.1353/dem.0.0106>
9. Dykstra, P. and L. van Wissen (1999). Introduction: The life course approach as an interdisciplinary framework for population studies. In: van Wissen, L.J.G. and Dykstra, P.A. (eds.) *Population Issues: An Interdisciplinary Focus*. New York: Kluwer Academic/ Plenum Publishers, 1-14.  
Fulltext pdf: [http://dx.doi.org/10.1007/978-94-011-4389-9\\_1](http://dx.doi.org/10.1007/978-94-011-4389-9_1)
10. Härkönen, J. (2005). Divorce risk factors across Finnish marriage cohorts, 1954-1989. *Yearbook of Population Research in Finland*, 41, 151-164  
Fulltext pdf: <http://journal.fi/fypr/article/view/45019/11297>
11. Härkönen, J. (2014) A note on interactions in event-history models  
(In course compendium)
12. Hoem, J.M. and M. Kreyenfeld (2006). Anticipatory analysis and its alternatives in life-course research. Part 1: Education and first childbearing. *Demographic Research*, 15:16, 461-484  
Fulltext pdf: <http://dx.doi.org/10.4054/DemRes.2006.15.16>
13. Hoem, J.M. (2014). The dangers of conditioning on the time of occurrence of one demographic event in the analysis of another. *Population Studies*, 68(2): 151-159  
Fulltext pdf: <http://dx.doi.org/10.1080/00324728.2013.843019>
14. Hosmer, D.W. and S. Lemeshow (2008) *Applied Survival Analysis: Regression Modeling of Time to Event Data*, 2<sup>nd</sup> edition (Wiley Series in Probability and

- Statistics), pages 132 – 136  
(In course compendium)
15. Jaccard, J. (2001). Interactions in Logistic Regression, especially pages 1-2, 12-23  
(In course compendium)  
Fulltext also available from SU computers: <http://srmo.sagepub.com/view/interaction-effects-in-logistic-regression/SAGE.xml?rskey=XPc5Xj&row=1>
  16. Kan, M. (2012). Ethnic-specific Reproductive Behavior in Independent Kazakhstan. *Stockholm Research Reports in Demography* 2012:15  
Fulltext pdf: [http://www.suda.su.se/polopoly\\_fs/1.295444.1473149506!/menu/standard/file/SRRD\\_2012\\_15.pdf](http://www.suda.su.se/polopoly_fs/1.295444.1473149506!/menu/standard/file/SRRD_2012_15.pdf)
  17. Kulu, H., Hannemann, T., Pailhé, A., Neels, K., Krapf, S., González-Ferrer, A., & Andersson, G. (2017). Fertility by birth order among the descendants of immigrants in selected European countries. *Population and Development Review*, 43(1), 31-60.  
Fulltext pdf: <https://doi.org/10.1111/padr.12037>
  18. Kleinbaum, David G./Klein, Mitchel (2005): *Survival Analysis: A Self-Learning Text*. 2nd Edition. New York: Springer, 4-8  
(In course compendium)
  19. Kreyenfeld, M. and Andersson, G. (2014) Socioeconomic differences in the unemployment and fertility nexus: Evidence from Denmark and Germany. *Advances in Life Course Research*,  
Fulltext pdf: <http://dx.doi.org/10.1016/j.alcr.2014.01.007>
  20. Mussino, E. and S. Strozza (2012) The fertility of immigrants after arrival: The Italian case. *Demographic Research*, 26(4): 99-130  
Fulltext pdf: <http://doi.org/10.4054/DemRes.2012.26.4>
  21. Ramirez, F.O., Souzal, Y., & Shanahan, S. (1997). The changing logic of political citizenship: cross-national acquisition of women's suffrage rights, 1890 to 1990. *American Sociological Review*, 62(5), 735-745.  
Fulltext pdf: <http://www.jstor.org/stable/2657357>
  22. Wallace, M., & Kulu, H. (2015). Mortality among immigrants in England and Wales by major causes of death, 1971–2012: a longitudinal analysis of register-based data. *Social Science & Medicine*, 147, 209-221.  
Fulltext pdf: <https://doi.org/10.1016/j.socscimed.2015.10.060>

## **Schedule: Event-History Analysis: Regression for Longitudinal Event Data, 7.5 ECTS-credits, Spring 2019**

<b>Meeting</b>	<b>Date, Time, Room</b>	<b>Topic</b>	<b>Reading Course Book</b>	<b>Other Readings</b>
1	03/28/19, Thursday 10-13 in B389	Introduction	BGR Ch1: 1-12 BGR Ch2: 38-57	Kleinbaum, p.4-8; Dykstra (1999)
2	03/29/19, Friday 10-13 in B389	Stata and Data		
3	04/01/19, Monday 09-16 in B389	Rates, duration, survival, hazard, and cumulative functions	BGR Ch1: 31-37 BGR Ch3: 58-85	Andersson & Philipov (2002); Kan (2012)
4	04/17/19, Wednesday 09-16 in B389	Event-history models: Exponential and piecewise exponential	BGR Ch4: 87-96 BGR Ch5: 116-127	Ramirez (1997); Mussino & Strozza (2012)
5	04/24/19, Wednesday 09-16 in B389	Time-varying covariates	BGR Ch6: 128-143	Kulu et al. (2017)
6	04/29/19, Monday 09-16 in B389	Interactions	BGR Ch6: 144-152	Jaccard (2001) p. 1-2, 12-23; Härkönen (2014); Härkönen (2005); Andersson (1998)
7	05/08/19, Wednesday 10-12 in B389	Model-Fit and Postestimation	BGR Ch5: 97-101	Hoem & Kreyenfeld (2006); Hosmer & Lemeshow (2008); Hoem (1996 & 2014)
8	05/08/19, Wednesday 13-15 in B389	More on parametric survival models	BGR Ch7: 182-215	Wallace & Kulu (2015)
9	05/22/19, Wednesday 09-15 in B389	Cox regression & Discrete-time Analysis	BGR Ch9: 223-246	Drefahl (2010); Andersson & Drefahl (2017)
10	05/22/19, Wednesday 15-16 in B389	Course Summary		
	<b>09/06/19</b>	<b>Deadline Take-Home Exam</b>		