TCORS 2.0
University of Michigan & Georgetown University

Center for the Assessment of Tobacco Regulations [CAsToR]

WORKSHOP (IN-PERSON):
AN INTRODUCTION TO THE ANALYSIS OF TRENDS AND TRANSITIONS FOR TOBACCO AND PUBLIC HEALTH RESEARCH

Location:
Georgetown University
REISS Science Building (REI 282 computer lab)
1551 Tondorf Rd
Washington, D.C.

Schedule:
Monday, August 15th, 2022, 9:00am – 4:45pm
Tuesday, August 16th, 2022, 9:00am – 3:30pm

Contacts:
Please direct all inquiries to: CAsToR-Workshops@umich.edu

Workshop purpose: This in-person introductory workshop will teach about methods for modeling trends and transition analysis with a focus on tobacco and public health research. The main contribution of the workshop would be to help non-statisticians to learn about methods for modeling trends and transition analysis and their utility: its use, practical application, and how to use software. The focus will be on concepts, ideas, and practical examples & applications. Participants will get practice running example codes on tobacco-related datasets.

Target audience: Tobacco researchers, non-statisticians, non-modelers with a background in epidemiology, public health or other health sciences. Federal partners interested in learning more about analytical approaches covered in this workshop.

Prerequisites: Students should have basic statistical or epidemiology knowledge, as well as familiarity with using R for basic coding.

IT and software: The workshop will take place in a computer lab equipped with 25 working stations, PCs and necessary software installed. If you choose to bring your own laptop, please ensure that the required software and packages needed for each hands-on module are properly installed (see each module syllabus for software details).

Workshop Schedule

***** *This schedule is subject to change*

	Module	Instructors
Monday, August 15, 2022		
9:00 - 9:30am	**Introduction to trend analyses and time-series approaches for modeling**	David Mendez (conceptual)
9:30 – 10:15am	**Joinpoint regression**	Huann-Sheng Chen (conceptual)
10:15 – 10:30am	Break	
10:30am – 12:00pm	**Joinpoint regression**	Jihyoun Jeon (application)
12:00 – 1:00pm	Lunch break	
1:00 – 2:30pm	**Age-period-cohort modeling**	Ted Holford (conceptual)
2:30 – 2:45pm	Break	
2:45 – 4:45pm	**Standard longitudinal data analysis** *(includes a 10-minute break between conceptual and hands-on)*	Steven Cook (conceptual and application)
Tuesday, August 16, 2022		
9:00 – 9:30am	**Introduction to compartmental modeling**	David Mendez (conceptual)
9:30 – 10:30am	**Markov modeling of transitions, Part 1: Multistate transition modeling**	Andrew Brouwer (conceptual)
10:30 – 10:45am	Break	
10:45-11:45am	**Markov modeling of transitions, Part 1: Multistate transition modeling**	Andrew Brouwer (application)
11:45am – 12:45pm	Lunch break	
12:45 – 1:45pm	**Markov modeling of transitions, Part 2: Latent transition analysis**	Ritesh Mistry (conceptual)
1:45 – 2:00pm	Break	
2:00 – 2:30pm	**Bayesian estimation and the Kalman Filter**	David Mendez (conceptual)
2:30 – 3:30pm	**Bayesian estimation and the Kalman Filter**	Thuy Le (application)

Workshop Syllabus

Module 1: Introduction to trend analyses and time-series approaches for modeling (conceptual, 30 min.)

Instructor: David Mendez (University of Michigan)

Learning objectives: This module will:

1) Expose you to different approaches to time series modeling.
2) Help you understand the underlying assumptions behind different time series modeling.

Module 2a: Joinpoint regression (conceptual, 45 min.)

Instructor: Huann-Sheng Chen (National Cancer Institute)

Learning objectives: This module will teach you about:

1) The use of Joinpoint modeling to analyze trends
2) Selection of joinpoints, estimation, and interpretation
3) Annual percent change, average annual percent change, and their confidence intervals.
4) Model selection procedures.
5) Clustering of similar trends.
6) Jump model for data with coding change.

Suggested Readings:

1. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications in cancer rates. Stat Med. 2000; **19**(3): 335- 351. correction: 2001;**20(4)**:655.

2. Chen HS, Zeichner S, Anderson R, Espey D, Kim HJ, Feuer EJ. Joinpoint-jump model in trend analysis with applications to coding changes in health statistics. J Off Stat. 2020; **36**: 49- 62.

3. Kim HJ, Luo J, Chen HS, et al. Improved confidence interval for average annual percent change in trend analysis. Stat Med. 2017; **36**: 3059- 3074.

4. Kim HJ, Chen HS, Midthune D, et al. Data driven choice of a model selection method in joinpoint regression. Journal of Applied Statistics. 2022.

5. Kim HJ, Luo J, Kim J, Chen HS, Feuer EJ. Clustering of trend data using joinpoint regression models. Stat Med. 2014; **33**: 4087- 4103.

Module 2b: Joinpoint regression (application, 1 h 30 min.)

Instructor: Jihyoun Jeon (University of Michigan)

Learning objectives: The hands-on part of this module will teach you how to:

1) Perform a trend analysis using the Joinpoint Trend Analysis software developed by the NCI.
2) Identify trend change points which break the trend into distinct periods.
3) Compare trends between different subpopulations.

Suggested Readings:

1. Joinpoint Trend Analysis Software & cited reference - Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. Stat Med 2000;19:335-51 (correction: 2001;20:655)

2. Meza R, Jimenez-Mendoza E, Levy DT. Trends in Tobacco Use Among Adolescents by Grade, Sex, and Race, 1991-2019. JAMA Netw Open. 2020 Dec 1;3(12):e2027465.

3. Salvatore M, Jeon J, Meza R. Changing trends in liver cancer incidence by race/ethnicity and sex in the US: 1992-2016. Cancer Causes Control. 2019 Dec;30(12):1377-1388.

4. Meza R, Meernik C, Jeon J, Cote ML. Lung cancer incidence trends by gender, race and histology in the United States, 1973-2010. PLoS One. 2015 Mar 30;10(3):e0121323.

Required software: Joinpoint Trend Analysis Software (https://surveillance.cancer.gov/joinpoint/). Note that this software is only available for Windows OS. Mac users need to have Windows OS via VMware (e.g., Parallels) to install this software. (The software will be installed on the workstations to be used during the workshop).

The Joinpoint analysis that we will run during this workshop is quite straightforward, so there are no required prerequisites or tutorials for the Joinpoint portion of this workshop.

Module 3. Age-period-cohort modeling (conceptual, 1h 30 min.)

Instructor: Ted Holford (Yale University)

Learning objectives: This module will teach you how to:

1) Express temporal trends from an age, period, and cohort perspective.
2) Parameterize time in a statistical model.
3) Partition temporal trends into linear and curvature effects.
4) Model trends of ever smoking prevalence.
5) Model trends in smoking initiation and cessation probabilities.
6) Combine the contributions of initiation and cessation to quantify smoking exposure.
7) Quantify the effect of smoking on population health.

Required Readings:

1. Holford, T.R. Understanding the effects of age, period and cohort on incidence and mortality rates. Annual Reviews of Public Health 12: 425-457, 1991.

Suggested Readings:

1. Holford, T.R., Levy, D.T., McKay, L.A., Clarke, L., Racine, B., Meza, R., Jeon, J., Feuer, E.J. Patterns of Birth Cohort–Specific Smoking Histories, 1965–2009. American Journal of Preventive Medicine. 46: e31-e37, 2014.

2. Holford, T.R., Meza, R., Warner, K.E., Meernik, C., Jeon, J., Moolgavkar, S.H., Levy, D.T. Tobacco control and the reduction in smoking-related premature deaths in the United States, 1964-2012. Journal of the American Medical Association. 311: 164-171, 2014.

Module 4. Standard longitudinal data analysis (conceptual + application, 2 h.)

Instructor: Steve Cook (University of Michigan)

Learning objectives: The hands-on part of the module will teach you how to:

1) Create censor and duration variables.
2) Restructure data to create an unbalanced person-period data set.
3) Estimate discrete time models.
4) Incorporate replicate weights.

Suggested Readings:

1. Allison, P. D. (2014). Event history and survival analysis: Regression for longitudinal event data (Vol. 46). SAGE publications.

2. Allison, P. D. (2010). Survival analysis using SAS: a practical guide. SAS Institute.

3. Jenkins, S. P. (2005). Survival analysis. Unpublished manuscript, Institute for Social and Economic Research, University of Essex, Colchester, UK, 42, pages 54-56.

4. Singer, J. D., Willett, J. B. (2003). Applied longitudinal data analysis: Modeling change and event occurrence. Oxford university press.

5. Singer, J. D., & Willett, J. B. (1993). It's about time: Using discrete-time survival analysis to study duration and the timing of events. Journal of educational statistics, 18(2), 155-195.

6. Wheaton, B,. & Young, M. (2020). Generalizing the regression model: Techniques for longitudinal and contextual analysis. Sage Publications.

Required software: STATA (https://www.stata.com/). (This software will be installed on your workstation during the workshop).

It is not required that workshop participants have a background in Stata, as it will be relatively simple to take the logic of Dr. Cook's practical example and then to apply it to the programming software of your choice. However, if you are interested in learning more about Stata, DR. Cook has recommended several resources; that list is available here.

Module 5. Introduction to compartmental modeling (conceptual, 30 min)

Instructor: David Mendez (University of Michigan)

Learning objectives: This module will help you:

1) Understand the basic concepts behind compartmental models and their advantages and disadvantages over individual-based (micro) models.
2) Understand a basic taxonomy and key properties of macro-dynamic models.

Suggested Readings:

1) Brauer F, Castillo-Chávez C (2001). Mathematical Models in Population Biology and Epidemiology. NY: Springer. ISBN 0-387-98902-1.

2) Towers, Sherry. Introduction to Compartmental Modeling

3) Mendez D, Warner KE, Courant PN. "Has Smoking Cessation Ceased? Expected Trends in the Prevalence of Smoking in the United States." American Journal of Epidemiology, 1998;148(3):249-58. doi:10.1093/oxfordjournals.aje.a009632

Module 6. Markov modeling of transitions, Part 1: Multistate transition modeling (conceptual + application, 2 h.)

Instructor: Andrew Brouwer (University of Michigan)

Learning objectives: This module will teach you how to:

1) Understand why transitions are important to many research questions for tobacco control.
2) Specify, implement, and interpret a multistate transition applied to longitudinal data of tobacco product use.
3) Graphically summarize transition probabilities and hazard ratios.
4) Apply a multistate transition model to the Population Assessment of Tobacco and Health (PATH) Study.

Required Readings:

1. Brouwer et al. (2022). Transitions between cigarette, ENDS and dual use in adults in the PATH study (waves 1–4): multistate transition modelling accounting for complex survey design. Tobacco Control.

Suggested Readings:

1. If you are not already familiar with R, or if you need a refresher, please use this tutorial. An Introduction to R. Sections 2-10.

2. Jackson. (2019). Multi-state modelling with R: the msm package.

Required software: R (v4.0 or later) and RStudio. Packages: msm, minqa, expm, dplyr, numDeriv, ggplot2, reshape2 (This software will be installed on your workstation during the workshop).

Module 7. Markov modeling of transitions, Part 2: Latent transition analysis (conceptual, 1h).

Instructor: Ritesh Mistry (University of Michigan)

Learning objectives: This module will teach you how to:

1) Understand the conceptual underpinnings of latent transition analysis and its value to tobacco control research.
2) Describe the key analysis steps involved in conducting latent transition analysis.
3) Interpret the parameters and results generated from latent transition analysis.

Suggested Readings:

1) Collins, L. M., & Lanza, S. T. (2009). Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences (Vol. 718). John Wiley & Sons.

2) Lanza, S. T., Patrick, M. E., & Maggs, J. L. (2010). Latent transition analysis: Benefits of a latent variable approach to modeling transitions in substance use. Journal of drug issues, 40(1), 93-120.

3) Lanza, S. T., Dziak, J. J., Huang, L., Wagner, A., & Collins, L. M. (2015). Proc LCA & Proc LTA users' guide (Version 1.3. 2). University Park: The Methodology Center, Penn State.

4) Huh, J., & Leventhal, A. M. (2016). Progression of poly-tobacco product use patterns in adolescents. American Journal of preventive medicine, 51(4), 513-517.

Module 8a. Bayesian estimation and the Kalman Filter (conceptual, 30 min)

Instructor: David Mendez (University of Michigan)

Learning objectives: This module will teach you how to:

1) Understand the basic principles of Bayesian estimation.
2) Understand how Bayesian estimation relates to structural time series modeling.
3) Understand the mechanics of the Kalman filter and its applications to structural time series modeling.

Suggested Readings:

1) Catlin, Donald E. (1989). Estimation, Control, and the Discrete Kalman Filter, Springer-Verlag.

2) Harvey, Andrew. (2002). Forecasting, Structural Time Series Models and the Kalman Filter. Wilmott. 2002. 10.1002/wilm.42820020109.

Module 8b. Bayesian estimation and the Kalman Filter (application, 1h.)

Instructor: Thuy Le (University of Michigan)

Learning objectives: This module will teach you how to:

1) Apply the Kalman filter to estimate the trajectory of the smoking cessation rate from smoking prevalence observations over time.
2) Implement and interpret the results of a Kalman filter estimation in R.

Required Readings:

1) Mendez D, Warner KE, Courant PN. "Has Smoking Cessation Ceased? Expected Trends in the Prevalence of Smoking in the United States." American Journal of Epidemiology, 1998;148(3):249-58. doi:10.1093/oxfordjournals.aje.a009632

2) Welch G & Bishop G. "An Introduction to the Kalman Filter." University of North Carolina at Chapel Hill, 2001 by ACM, Inc.

Suggested Readings:

1) Simon, Dan (2006). Optimal state estimation: Kalman, H infinity, and nonlinear approaches. John Wiley & Sons.

Required software: RStudio (This will be installed on your workstation at the workshop).