XVI CONVEGNO ITALIANO DEGLI UTENTI DI STATA

Firenze, 26-27 Settembre 2019

CODICE

I-SUG

LUOGO E DATA

Il Convegno si terrà a Firenze
il 26-27 Setembre 2019 presso
Hotel Brunelleschi Firenze
Piazza S. Elisabetta 3

CONTATTI

Monica Gianni
Via Rettangolo, 12-14
67039 Sulmona (AQ)
T. +39 0864 210101

www.tstat.it | www.tstattraining.eu
formazione@tstat.it | training@tstat.eu

9.00 - 9.30 Registrazione dei partecipanti

9.30 - 10.30 I SESSIONE - EXPLOITING THE POTENTIAL OF STATA 16, I

Stata 16 - Under the Hood • Bill Rising, StataCorp.

Stata 16 added a number of new features which could affect how users think of working with Stata. Two of these features are frames and dynamic document extensions. This talk will show the basics of using frames and explain how they can be exploited both for space and speed. I'll also introduce the new dynamic document features together with some example documents.

10.30 - 10.45 Pausa caffè

10.45 - 12.00 II SESSIONE - COMMUNITY CONTRIBUTED, I

Estimation of a latent network via LASSO regression using Stata • Giovanni Cerulli, IRCrES-CNR, Roma.

I present a model and a new Stata routine for estimating a latent (i.e., non-observable) network among N units (e.g., individuals, companies, countries, etc.) using a set of units' characteristics and without knowing any prior linkage among them. In this approach, the units are represented via a set of indicators measuring a specific concept, such as riskiness, knowledge, etc. A standard regression would be unsuited for estimating the linkages as - in this case - the number of observations (the characteristics vector) is much smaller than the number of variables (the units). The Lasso Regression allows one to address high-dimensional settings like this one by allowing for an estimation of a generally sparse matrix of linkages (the network). We will provide a Stata simulation and possibly an application to real data.

Modelling the probability of occurrence of events with the new *stpreg* command • Matteo Bottai, Andrea Discacciati* e Giola Santoni, Karolinska Institutet, Stoccolma.

We introduce the new ***stpreg*** command, to fit flexible parametric models for the event-probability function, a measure of occurrence of an event of interest over time. The event-probability function is defined as the instantaneous probability of an event at a given time point conditional on having survived until that point. Unlike the hazard function, the event-probability function defines the instantaneous probability of the event. This talk describes its properties and interpretation along with convenient methods for modelling the possible effect of covariates on it, including flexible proportional-odds

models and flexible power-probability models, which allow for censored and truncated observations. We compare these with other popular methods and discuss the theoretical and computational aspects of parameter estimation through a real data example.

Simulating Gaussian Stationary Dynamic Panel Data Models in Stata: New Features of *xtarsim* • Giovanni Bruno, Università Commerciale Luigi Bocconi, Milano.

The original Stata command *xtarsim* simulates dynamic panel data models with exogenous regressors and *i.i.d.* errors. I have now extended ***xtarsim*** in order to simulate models with various types of endogenous or predetermined regressors. This new version of ***xtarsim*** also allows MA(1) errors when regressors are exogenous.

12.10 - 13.00 III SESSIONE - EXPLOITING THE POTENTIAL OF STATA 16, II

Nonlinear Dynamic Stochastic General Equilibrium Models • David Schenck - Senior Econometrician, StataCorp.

Dynamic stochastic general equilibrium (DSGE) models are used in macroeconomics for policy analysis and forecasting. A DSGE model consists of a system of equations (usually a nonlinear system of equations) that is derived from economic theory. This presentation illustrates how to easily solve, estimate, and analyse nonlinear DSGEs. We will explore how to obtain policy matrices, transition matrices, and impulse response functions for nonlinear models.

13.00-14.00 Pranzo

14.00 - 14.50 IV SESSIONE - COMMUNITY CONTRIBUTED, II

Measuring Heterogeneity And Efficiency of Firms within the Same Industry: a C++ plugin for Stata for computing the *Zonotope* • Marco Cococcioni*, Università degli Studi di Pisa, Marco Grazzi, Università Cattolica del Sacro Cuore, Milano, Le Li Chuo University, Tokyo e Federico Ponchio, ISTI CNR, Pisa.

In this work we describe a new Stata command ***zonotope*** that, by resorting to a geometry-based approach, enables one to provide a measure of productivity that fully accounts for the existing heterogeneity across firms within the same industry. Further, the method that we propose also enables one to assess the extent of multi-dimensional heterogeneity with applications to production analysis and productivity measurement. Finally, we detail the functioning of the software to perform the related empirical analysis and we discuss the main computational issues encountered in its development.

Calling Python Scripts in Stata: a Power-Law application • Antonio Zinilli, IRCrES-CNR, Roma.

The -***python***- package facilitates integrating Python with Stata 16 by allowing automatic inter-process communication between the two software packages. Here we present a statistical application implemented in Python and called in Stata for discerning and quantifying power-law behaviour in empirical data.

14.50 - 15.40 V SESSIONE - STUDI APPLICATIVI USANDO STATA

The Contribution of Proportional Taxes and Tax-Free Cash Benefits to Income Redistribution over the Period 2005-2018: Evidence from Italy • Stefano Boscolo, Università degli Studi di Modena e Reggio Emilia, Modena.

During the last two decades a growing interest in understanding what determines the redistributive role of tax-benefit systems has been recorded worldwide. For the case of Italy, previous analyses were mainly focused on quantifying the contribution of marginal tax rates, deductions and tax credits to the redistributive capacity of PIT, neglecting the effect on income redistribution of proportional taxes and income sources exempt from taxation such as tax-free cash benefits. The following paper aims to fill this gap by applying two alternative Gini-based decomposition methodologies (Onrubia et al., 2014; Urban, 2014) to the Italian tax-benefit system's redistributive power over the period 2005-2018. The contribution of each tax-benefit instrument is quantified for several baseline policy scenarios which diverge from each other for being representative of different degrees of extension of the tax-benefit system under study.

An application in STATA when investigating the relationship between cancer and dementia • Cecilia Damiano e Rino Bellocco, Università degli Studi di Milano-Bicocca.

Older people are often affected by several comorbid conditions and by an increasing risk of death that arises with aging. Previous studies examining the association of cancer with dementia in older adults have usually used standard approaches without taking into account the competing risk of mortality. However, ignoring mortality may not provide valid estimates of risk of dementia, because cancer is strongly associated with the competing risk of death. The present study considers people over-72 years old from two Swedish population-based longitudinal studies: the Kungsholmen Project (KP) and the Swedish National Study on Aging and Care project conducted in the Kungsholmen district of the city of Stockholm (SNAC-K project). The aim of the study is to analyse the association between cancer and the onset of dementia in the considered older population. The competing risk methodology is used, to illustrate the appropriate statistical methods for competing risks, their correct application and interpretation of the results, having death as the competing event.

15.40 - 16.00 Pausa caffè

16.00 - 16.45 VI SESSIONE - INVITED SPEAKER

A Brief Introduction to Machine Learning • Achim Ahrens - Public Policy Group, ETH Zürich.

In this short lecture I will attempt to demystify the field of Machine Learning and compare it to traditional statistical approaches in economics and social sciences. I discuss relative strengths and weaknesses, and how Machine Learning can facilitate causal inference. The lecture serves as a preamble to the one-day workshop that will take place on the following day.

16.50 - 17.30 VII SESSIONE - REPORT TO USERS WISHES AND GRUMBLES • BILL RISING AND DAVID SCHENCK, STATACORP

La sessione “Wishes and Grumbles” offre ai partecipanti la possibilità di interagire direttamente con la StataCorp: sarà possibile evidenziare problemi o limitazioni del software nonché suggerire eventuali miglioramenti o comandi che potrebbero essere inclusi in Stata.

20.15 Cena Sociale (facoltativa)

Come ogni anno proponiamo una serata insieme al termine del Convegno. Quest’anno la cena sociale si terrà, la sera del 26 Settembre, presso la Trattoria Sant’Agostino, Via Sant’Agostino 23R - Ang Via Maffia https://www.trattoriasantagostino.com.

Il ritrovo sarà direttamente al ristorante alle ore 20.15 (previa conferma di adesione al momento dell’iscrizione).

27 SETTEMBRE 2019 - CORSO DI FORMAZIONE

An Introduction to Machine Learning in Stata

Achim Ahrens - Public Policy Group, ETH Zürich

DESCRIZIONE DEL CORSO

Over the last few years we have experienced an unprecedented explosion in the availability of data relating to social, economic, financial and health-related phenomena. Today researchers, professionals and policy makers have therefore access to enormous datasets (so-called Big Data) containing an abundance of information regarding individuals, companies and institutions.

Machine learning has evolved in response to both the need to analyse extremely large databases and the availability of both sophisticated software and extremely powerful computer capacity. Machine learning, an application of artificial intelligence, offers a relatively new approach to data analysis, which trains systems to automatically learn and improve from experience without being explicitly programmed, relying instead on patterns and inference in the data.

This workshop offers participants an introduction to both machine learning techniques and the commands for Machine Learning recently introduced in Stata 16.

DESTINATARI: Il workshop è di particolare interesse per ricercatori e professionisti in biostatistica, economia, marketing, scienze sociali e sanità pubblica che desiderano acquisire gli strumenti fondamentali per implementare l'approccio di *machine learning* sui così detti *Big Data.*

REQUISITI RICHIESTI: Conoscenza di base di econometria/ statistica e del software Stata.

PROGRAMMA

SESSION 1: EXAMPLES OF MACHINE LEARNING METHODOLOGIES

This opening session focuses on the more popular supervised and unsupervised Machine Learning (ML) techniques, and their implementation in Stata. This session focuses on regression trees, random forests and cluster analysis.

SESSION 2: REGULARIZED REGRESSION

Regularized regression and the Lasso approach play a central role in Machine Learning. This session is therefore devoted to Lasso, Elastic Net and related methodologies. We will demonstrate their application in Stata using both the user written Lassopack commands and Stata 16's new Machine Learning routines.

SESSION 3: CAUSAL INFERENCE WITH MACHINE LEARNING

The primary strength of Machine Learning is prediction. In this session, we illustrate how Lasso and other Machine Learning methodologies can also be used to facilitate causal inference. The workshop concludes by looking at the latest developments in the evolving literature on Machine Learning and causal inference.

INFORMAZIONI GENERALI

Il materiale didattico distribuito include le dispense con la parte teorica, i do file e le banche dati per l'implementazione di tutte le applicazioni empiriche e una licenza temporanea del Software Stata 16 valida per 30 giorni dall'inizio del corso. Data la natura applicata del corso si consiglia l'utilizzo del proprio personal computer per seguire autonomamente le sessioni applicate.

Il numero massimo di iscritti ammessi al Corso di Formazione è 15, ed il termine per presentare la propria richiesta di ammissione è il 15 Settembre 2019.

Per ulteriori informazioni consultare la pagina del convegno https://www.tstat.it/utenti/XVI-convegno-italiano-degli-utenti-di-stata/ oppure contattare la segreteria organizzativa a **formazione@tstat.it**.