

Survival analysis, longitudinal analysis and causal inference

Odd O. Aalen

Received: 1 December 2009 / Accepted: 1 December 2009 / Published online: 11 December 2009
© Springer Science+Business Media, LLC 2009

Causal inference has become a hot topic in statistics. The term “causal” is often added to statistical concepts, like in “causal estimation”, meaning that one performs the analysis according to certain specific methods, e.g. based on counterfactual methodology. However, the fact of the matter is that there is no magic bullet for drawing safe causal conclusions from statistics. Whichever method one uses, one will typically have to make some untestable assumptions (unless one should be in the happy position of analyzing a randomized double-blind trial). Hence, there should be an opening for various attempts to say something sensible about the causal implications of the data one has available, also those outside the main stream.

In this issue we focus on data collected over time, and the methods presented try to do justice to the time aspect. Causal inference in a strict sense is not the issue of all the papers, some focus on longitudinal and survival analysis trying to understand the dynamics of a process.

Granger causality is the subject of the paper by Michael Eichler and Vanessa Didelez. The concept of Granger causality arose in econometrics, but has never really gained much foothold in main stream statistics. Nevertheless, it is a very useful concept to which much more attention should be paid. Granger causality is a notion based on prediction and in their paper Eichler and Didelez investigate to which extent it can be connected to the intervention based concept of causality which is more common in statistics. In the next paper, Anne Gégout-Petit and Daniel Commenges assume a somewhat more abstract point of view, considering stochastic processes and extending the notion of local independence to a more general concept of influence for stochastic processes. Local independence is connected to Granger causality, and is a concept of much interest in event history analysis and other fields.

O. O. Aalen (✉)
University of Oslo, Oslo, Norway
e-mail: o.o.aalen@medisin.uio.no

Inverse probability weighting has become one of the major tools for making causal inference. In their paper Maarten Bekaert, Stijn Vansteelandt and Karl Mertens use this to study subdistributions for competing risks. Specifically, they consider the issue of assessing mortality attributable to nosocomial infections in the intensive care unit. The authors give a careful analysis using a marginal structural model. This type of model is also used in the paper by Jessica Gerald Young, Miguel A. Hernán, Sally Picciotto and James M. Robins. Here they study the effect of a time-varying exposure on survival, and how one shall adjust for time-dependent confounders. A marginal structural Cox model is compared to two models implementing various types of g-estimation.

Torben Martinussen also considers event time data, but applying a different approach from the above. The focus is on understanding direct and indirect effects by applying a dynamic path analysis, the idea being to extend ideas from path analysis to a time-dependent setting. The aim is to understand to which extent a treatment effect is mediated through an internal time-dependent covariate. Martinussen applies an additive hazard model as an alternative to the Cox model, thereby achieving a more detailed analysis of the changes in effect over time.

Although survival analysis and longitudinal data analysis are often perceived as two different areas, there is actually a natural connection between these fields since they both analyze data collected over time. Such a connection is pointed out in the paper by Daniel Farewell and Robin Henderson, where they argue strongly for closer contact between the fields. They present an application to the analysis of disease progression of patients with psoriatic arthritis. A related issue is discussed in the paper by Entisar Elgmati, Daniel Farewell and Robin Henderson. Here they look at a martingale residual diagnostic which is applicable to both longitudinal and recurrent event data. This is based on the recent extension of martingale methods to longitudinal data in which two of these authors have been involved. New dynamic models have resulted from this and the paper gives an easy-to-use method for checking model fit. Two applications are given: a study of recurrent occurrences of diarrhoea, and a longitudinal study of mental health.

A mixture of longitudinal and survival data is also present in the paper by Weiliang Qiu and Bernard Rosner. They study nutritional data and apply a cumulative average approach to study the relationship between disease incidence and repeated measurements of exposure. The emphasis is on the issue of measurement error, and how to correct for such error. This is an omnipresent problem for all methods trying to say something about risk factors for disease and highly relevant also in a causal setting.