Preventive Veterinary Medicine 113 (2014) 304–312

Contents lists available at ScienceDirect

Preventive Veterinary Medicine

journal homepage: www.elsevier.com/locate/prevetmed

The analysis—Hierarchical models: Past, present and future

Henrik Stryhn[a,*], Jette Christensen[b]

[a] Centre for Veterinary Epidemiological Research, Atlantic Veterinary College, University of Prince Edward Island, Charlottetown, PE C1A 4P3, Canada

[b] Canadian Food Inspection Agency, Epidemiology and Surveillance Section, Atlantic Veterinary College, Department of Health Management, 550 University Avenue, Charlottetown, Prince Edward Island C1A 4P3, Canada

ARTICLE INFO

Article history:
Received 13 March 2013
Received in revised form
10 September 2013
Accepted 1 October 2013

Keywords:
Hierarchical data structure
Random-effects model
Survival analysis
Non-proportional hazards
Multi-level
Bayesian modelling

ABSTRACT

This paper discusses statistical modelling for data with a hierarchical structure, and distinguishes in this context between three different meanings of the term hierarchical model: to account for clustering, to investigate variability and separate predictive equations at different hierarchical levels (multi-level analysis), and in a Bayesian framework to involve multiple layers of data or prior information. Within each of these areas, the paper reviews both past developments and the present state, and offers indications of future directions. In a worked example, previously reported data on piglet lameness are reanalyzed with multi-level methodology for survival analysis, leading to new insights into the data structure and predictor effects. In our view, hierarchical models of all three types discussed have much to offer for data analysis in veterinary epidemiology and other disciplines.

1. Introduction

In everyday language, a hierarchy may be understood as "a system in which people are put at various levels or ranks according to their importance" (Cambridge International Dictionary of English, 1995). In this definition we may replace "people" by other—concrete or abstract—items, and "importance" can also be understood broadly as relating to a certain order in which we view the items, e.g. in taxonomy. Given such a broad definition and usage of the term hierarchy, it is hardly surprising that the term "hierarchical (statistical) model" is used in a variety of different contexts and meanings. Our main focus here is on hierarchical data structures in which the "subjects" (experimental or measurement units) are organized in groups that can be described by, or depicted in, a hierarchy. The organization into groups may follow from the physical location of the subjects or from the circumstances of the recording of data. Typical examples from veterinary studies involve records on animals housed in farms or treated at veterinary clinics or hospitals. Data hierarchies can comprise multiple levels, e.g. by considering sub-units within farms such as pens, or by considering further grouping of farms into regions. Longitudinal data may be seen as a special case of a hierarchical data structure, with repeated measures taken on subjects, where special consideration is needed for potential autocorrelation within subjects over time. In the usual nomenclature, the lowest hierarchical level corresponds to the unit of the observations (or measurements); for example, level 1 may correspond to animals and level 2 to farms. As already indicated, a complex data structure can also be comprised of several unrelated hierarchies pertaining to different characteristics of the subjects; a typical example is that in addition to the location of animals their origin forms another relevant hierarchy. We refer to Dohoo et al. (2009, Chapter 20) for further examples and illustrations of hierarchical and related data structures.

Traditionally the importance of hierarchical data structures has been linked to the violation of independence assumptions involved in classical statistical models, such

* Corresponding author. Tel.: +1 902 894 2847; fax: +1 902 566 0823.
E-mail address: hstryhn@upei.ca (H. Stryhn).

0167-5877/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.prevetmed.2013.10.001

as linear (regression) models. It is intuitively obvious that subjects linked by their presence at the same hierarchical level, e.g. animals in the same farm, may no longer give rise to observations that can be assumed independent. The phrase "animals in the same farm are more similar than animals in different farms" will have made its way through innumerable classrooms. The term clustering is also commonly used to express the notion that similarity between observations sharing certain hierarchical levels leads to clusters in the data. This should not be confused with cluster analysis, which aims at identifying clusters in the data without referring to known structures. Notwithstanding the validity of the assertion (the violation of assumptions), and the need to adjust the modelling approach to account for it, this particular consequence of a hierarchical data structure has to some extent overshadowed the new potentials offered by complex data structures. Modelling and exploring dependence (or correlation) structure may be perceived as more complicated than building predictive equations from a set of predictor variables. It does however offer different kinds of insights into the factors affecting the variability of outcomes of interest across a population. Multi-level modelling utilizes the decomposition of variability across the hierarchical levels to study the impact of predictor variables through separate modelling equations at each level of the hierarchy.

Even after restricting the coverage of hierarchical models to involve modelling derived from particular data structures, we will for the present discussion further distinguish between 3 major domains or scopes of modelling.

(i) Models to account for hierarchical data structure, or clustering (Section 2).
(ii) Simultaneous modelling at multiple hierarchical levels, or multi-level analysis (Section 3).
(iii) Bayesian hierarchical modelling, with multiple layers of equations and assumptions (Section 4).

The objective of the paper is to provide insight into the current state of statistical theory and applications for each of these domains of hierarchical models, with particular focus on their application to veterinary epidemiology. For this purpose we briefly outline historical developments and explain the fundamentals of the modelling, without reproducing any of the detailed expositions from the current literature. Instead, we indicate some recent and ongoing developments by examples from our practice. This paper is based on a presentation given at the 2012 Calvin W. Schwabe Symposium honouring the lifetime achievement in veterinary epidemiology and preventive medicine of Dr. Ian R. Dohoo.

2. Hierarchical models I: to account for clustering

The most commonly used and most versatile method to account for dependence derived from hierarchical data structures consists in including random effects in the statistical model. As a general rule each hierarchical level above the lowest (observation) level should be represented by a set of random effects. Random effects are latent (unobserved) variables with assumed distributions, most commonly Gaussian (normal) distributions, that reflect the variability in the population the random effects represent (e.g. a population of farms). As the random effects are shared by all observations within the same unit of a hierarchical level (e.g., all animals in a given farm), they induce a dependence between such observations. Accordingly, statistical inference based on a random-effects model accounts for clustering in the data.

Random effects in linear models date back at least to work in the 1940s on genetics and animal breeding. A substantial body of statistical theory was developed in the 1950s and 1960s to estimate variance components (the variances of the different random effects), as described in a subsequent review (Robinson, 1991). The assumed Gaussian distributions in linear models with random effects made the calculations manageable with limited computing power, while random-effects models for non-normal data started to appear in theory and applications in the 1980s. This development made use of the formulation of a generalized linear model (GLM) framework, on which the alternative generalized estimating equations (GEE) methodology was also based. Software implementations gradually became available in the 1990s, starting with models involving two hierarchical levels. A seminal paper introducing the new methods for dealing with clustered data to veterinary epidemiology was McDermott et al. (1994), later followed by Dohoo et al. (2001).

The inclusion of random effects in GLMs as well as GEE estimation are by now mainstream approaches, used routinely in veterinary epidemiology. The applied researcher can choose from a multitude of textbooks and implementations in statistical software, although some differences still exist in estimation methods and flexibility. Further developments in recent years have been confined to "difficult situations" (e.g. due to large data sets with many hierarchical levels or to non-standard data structures) and to models beyond the standard GLM framework. We specifically discuss challenges and developments in two commonly encountered areas: time-to-event (also known as "survival") outcomes, and "composite" outcomes that may be viewed as consisting of two (or more) distinct components. Evidently the question of adjusting for hierarchical structure may be asked (and also has been addressed in the literature) for a wide range of analytical settings, from diagnostic-test evaluation and agreement to multivariate analysis.

2.1. Models for composite outcomes

Here we are concerned not with a truly multivariate outcome but with univariate distributions formed by adjoining outcomes of different types. The most common example we have in mind is a quantitative outcome, either a count or a measurement, that can also take values at zero, at a lower detection threshold or at an upper detection threshold. Count distributions such as the Poisson or negative binomial include zero as a legitimate value, but trouble arises if the zeros occur more frequently in the data than predicted from these distributions, a phenomenon broadly referred to as zero-inflation. For continuous outcomes, the presence of measurements equal to a lower detection