Inference and Computational Methods for Regression Modelling in Multiple Time Series of Counts with Random Effects

William T.M. Dunsmuir

The University of New South Wales

Australia

Abstract. This talk is concerned with modelling multiple time series of counts in which shared regression effects need to be tested and in which there are random effects and serial dependence. Let Y_{jt} be the observation at time $t = 1, \ldots, n$ on the jth time series where $j = 1, \ldots J$. Typically n is much larger than J in the settings we focus on. Health data applications used to motivate this work have values of these ranging from $n = 72$ to 1500 and $J = 5, 10, 50$. For each j assume that, given the state process $\{W_{jt}\}$, the Y_{jt} are independent with exponential family distribution $f(y_{jt}|\{W_{jt}\}) = \exp\{[y_{jt}W_{jt} - b(W_{jt})] / a_{jt}(\phi) + c(y_{jt})\}$. The state process is assumed to be of the form $W_{jt} = r_t^T U_j + x_{0,t}^T \beta^{(0)} + x_{j,t}^T \beta^{(j)} + \alpha_{jt}$ in which there are q random effects $U \sim \mathrm{N}(0, \Sigma_U)$ where Σ_U is a $q \times q$ covariance matrix with associated covariates $r_t = (r_{1t}, \ldots, r_{qt})^T$, there are fixed effect covariates $x_{0,t}$ common to each individual series and particular covariates $x_{j,t}$ relevant to the j-th series and that the process α_{jt} either follows a transition process $\alpha_{jt} = \sum_{l=1}^{\infty} \gamma_l^{(j)}(\tau^{(j)}) e_{j,t-l}$ with $e_{j,s} = (y_{j,s} - \mu_{j,s}) / s(\mu_{j,s})$ for some normalizing scale function $s(\cdot)$ or is an unobserved stationary Gaussian time series with parameters θ_j. Methods of inference based on the likelihood will be considered. The primary focus will be on the transition specification of α_{jt}. In this case the likelihood over all series can be readily built up using existing software of the author for fitting transition models to single time series. The talk will describe how this is done and illustrate the methods on an example of assessing the impact, on monthly counts of single vehicle night time fatalities, of lowering the legal blood alcohol level in 17 US states. Estimation for the latent process specification α_{jt} is computationally more difficult requiring as it does approximation of large (n) dimensional integrals. Available computational methods for this specification will be

discussed. For either specification of α_{jt} theory of inference is underdeveloped at this stage and what is known will be reviewed.