POISSON DYNAMIC MODELS. AN APPLICATION TO INTERURBAN ACCIDENTS IN SPAIN

Camino González Fernández, ETSII-UPM, Spain
Jose Mira McWilliams, INSIA-UPM, Spain
Blanca Arenas Ramírez* INSIA-UPM, Spain
Claudia Saenz Sagástegui, ETSII-UPM, Spain

Introduction

In this study we apply count data models to four integer-valued time series related to accidentality in Spanish roads applying both the frequentist and Bayesian approaches.

The time series are: number of fatalities, number of fatal accidents, number of killed or seriously injured (KSI) and number of accidents with KSI.

The model structure is Poisson regression with first order autoregressive errors.

The purpose of the paper is first to sort out the explanatory variables by relevance and second to carry out a prediction exercise for validation

Model equations

Indirect exposure

$$E[y_t] = \mu_t = \exp[\beta_0 + \sum_{j=1}^{k} \beta_j x_{jt} + \rho y_{t-1}]$$

$$E[y_t] = \mu_t = \exp[\beta_0 + \sum_{j=1}^{k} \beta_j x_{jt} + \rho[y_{t-1} - \exp(\sum_{j=1}^{k} \beta_j x_{jt})]$$

Direct exposure

$$E[y_t] = \mu_t = \exp[\log(v_t) + \beta_0 + \sum_{j=1}^{k} \beta_j x_{jt} + \rho y_{t-1}]$$

$$E[y_t] = \mu_t = \exp[\log(v_t) + \beta_0 + \sum_{j=1}^{k} \beta_j x_{jt} + \rho(y_{t-1} - \exp(\log(v_{t-1}) + \beta_0 + \sum_{j=1}^{k} \beta_j x_{jt-1}]$$

Estimation and prediction

Data:
The empirical analysis in this study was carried out using the monthly data on road accidents in Spain. The data covers the period 2000-2007.

Features of modelling estimation and and prediction:

- The four times series are modeled independently, i.e, without taking into account possible cross-correlations.
- Two types of observation-driven models are applied: Lagged Poisson Models (LPM) and Corrected Lagged Poisson Models (CLPM)
- The explanatory variables can be clustered in nine categories: exposure, economic situation, climatology, driver statistics, law and enforcement, road network, fleet characteristics, calendar, driver behavior.
- Two different forms of measuring exposure were incorporated: direct and indirect.
- Free software is used for both cases: R under the frequentist approach and WINBUGS under the Bayesian approach.
- Bayesian estimation through WINBUGS implements the Markov Chain Monte Carlo methodology.
- The goodness of fit measures used for model selection are the pseudo-R^2 and the AIC under the frequentist approach and the deviance information criteria (DIC) under the Bayesian approach.
-

Results

INDIRECT EXPOSURE

ERRORS IN % FROM R RESULTS -LPM				
	Accidents		Victims	
	Fatal	KSI	Fatalities	SI
Jan-09	3.145	6.047	4.05	0.68
Feb-09	18.33	4.033	2.87	4.27
March-09	10.29	3.53	11.15	8.69
April -09	28.45	4.01	6.43	13.48
May-09	2.26	10.21	5.08	7.70
June-09	8.17	11.05	2.86	5.31
July-09	11.73	13.55	3.11	9.78
August-09	6.86	6.72	6.23	23.01
Sep-09	9.66	8.55	11.22	24.37
Oct-09	23.19	15.26	2.69	34.11
Nov-09	14.72	9.57	2.07	20.18
Dic-09	6.83	5.19	1.39	23.81
MAPE	9.95	8.14	4.93	14.56

DIRECT EXPOSURE

ERRORS IN % FROM R RESULTS -LPM				
	Accidents		Victims	
	Fatal	KSI	Fatalities	SI
Jan-09	1.258	13.55	16.22	1.80
Feb-09	10.00	10.56	14.83	0.09
March-09	8.088	8.98	20.60	10.04
April -09	23.57	8.21	14.55	12.87
May-09	1.85	0.45	33.05	5.44
June-09	0.00	2.47	31.15	0.61
July-09	14.28	2.04	45.14	7.18
August-09	8.82	4.37	34.07	17.43
Sep-09	3.98	1.86	55.12	16.00
Oct-09	14.94	7.63	42.15	21.18
Nov-09	9.20	2.12	43.00	8.07
Dic-09	1.86	3.19	47.22	14.02
MAPE	10.19	5.85	35.59	8.17

Conclusions

- The effects of the explanatory variables on the responses resulted, for most models, as expected in sign and magnitude.
- Prediction errors were competitive versus alternative models in the literature.
- Frequentist and Bayesian results were very similar.
- Use of different parameterizations for exposure.
- Use of different models for autoregressive errors.
- Successful application of the Bayesian approach is encouraging for future use of powerful Bayesian tools for model selection .

Main references

1. Hardin, J and Hilbe, J., (2007). Generalized Estimating Equations. Chapman and Hall.
2. Davis, R. and Duinsmuir, W. (2003). Observation-driven models for Poisson counts. Biometrika, 90,4, 777-790.
3. Robert, C. P., and Casella, G. (2004). Monte Carlo statistical methods . New York: Springer.
4. Zeiles, A., Kleiber, C. and Jackman, S. (2008). Regression Models for Count data in R. Innsbruck.