POLS571 - Longitudinal Data Analysis
October 18, 2001
Time Series Models for Event Counts

1 Introduction

Modeling time series of event counts is tricky. In the first place, all the reasons that one can't use Gaussian errors for event counts in a cross-sectional context (e.g. negative predicted counts, poor distributional fit, etc.) all apply to time-series as well. Beyond these problems, however, there are also more specific issues related to the time-series aspect of the model. Think about it for a minute...

- Event counts are positive, and discrete; that means that the extent to which they can be autoregressive is limited.
- First-order dependence (e.g., along the lines of $Y_t = \rho Y_{t-1} + ...$ imply that the value of the count will grow without bound, or go to zero, over time; most of the time, this isn't realistic.
- Taking first differences also isn't a very attractive option, primarily for distributional reasons. While we know the distribution of a variable that is the difference of (e.g.) two *independent* Poisson processes, its much harder to say anything about such a difference when the two aren't independent (as is almost always the case with time-series data).

If the counts are large enough (typically, above 25 or so) then treating them as normally-distributed isn't such a terrible thing.

2 Three Approaches

2.1 PEWMA

PEWMA stands for "Poisson Exponentially Reweighted Moving Average". It is the creation of Pat Brandt and John Williams (*AJPS* paper). It is basically a Poisson model that captures temporal dynamics through a state-space formulation. Specifically, the model estimates a time-dependent average of the mean of the event count process, that is discounted exponentially (hence the name).

The model has two parts. The first is the *measurement equation*, which is simply:

$$Pr(Y_t|\mu_t) = \frac{\mu_t^{Y_t} exp(-\mu_t)}{Y_t!} \tag{1}$$

which is a standard Poisson distribution, with μ the mean arrival rate for events. This mean is set equal to:

$$\mu_t = \mu_{t-1}^* exp(X_t \beta) \tag{2}$$

That is, the covariates impact the rate as a multiplicative function on μ_{t-1}^*, where μ_{t-1}^* is a gamma-distributed conjugate prior, calculated from the previous $t-1$ observations. The transition equation is then defined as:

$$\mu_t = exp(r_t)\mu_{t-1}\eta_t \tag{3}$$

where η_t is beta-distributed with parameter ω and r_t describes the growth of the series (if any).

If all this sounds complicated, it is. But, its also intuitive, as just a Poisson process in which the mean evolves over time, according to a combination of an exponential weighting of past values of the mean and the current values of the covariates. Just remember that:

- β gives the effect(s) of the covariate(s) on the mean level of the count μ_t. The effect of those covariates enters exponentially, just as in the standard Poisson model.

- ω is the weighting parameter, which determines the degree of persistence in the series. Smaller values of ω mean less discounting, and correspondingly higher persistence/temporal dependence. In the limit, $\omega = 1$ yields a process with a constant mean.

- The transition equation (3) can be rewritten as:

$$ln\mu_t - ln\mu_{t-1} = rt + ln\eta_t \tag{4}$$

which means that the period-to-period change in the (log of the) mean rate is a function of the growth term and an "error" (η_t). This in turn means that:

- r_t describes the growth in the series; $r_t = 0$ means that the series isn't growing over time.
- is the proportional stochastic shift ("shock") in μ_t between $t-1$ and t.

2.2 Characteristics

PEWMA:

- Is appropriate for series which are *persistent* – that is, in which the event count process changes slowly over time.
- Is a data generating process that yields nonstationary series with medium- to long-term dependence and persistence, but also differenced series which are stationary.
- Is estimable in GAUSS using a routine written by Pat & Co.
- Is almost certainly "better" (from both a statistical and a substantive perspective) than OLS, Poisson with a lagged dependent variable, or other commonly-used routines.

3 PAR

PAR(p) stands for *Poisson Autoregression of order p*. It is also due to Brandt and Williams (*Political Analysis* 2001). The model is:

$$E(Y_t|Y_0, Y_1, ...Y_{t-1}, X_0, X_1, ...X_{t-1}) = \sum_{i=1}^{p} \rho_i Y_{t-i} + \lambda \qquad (5)$$

which is a general form of a mean-stationary time-series model (i.e., without any particular distributional assumptions about Y). To get the PAR model, we start by assuming that the conditional values of Y are drawn from a Poisson distribution:

$$Pr(Y_t|\mu_t) = \frac{\mu_t^{Y_t} exp(-\mu_t)}{Y_t!} \quad (6)$$

and that μ_t is the conditional mean of the AR process given above (see Brandt and Williams for details). This is the measurement equation; the state equation's density is a gamma prior (as in the PEWMA model):

$$Pr(\mu_t|Y_0, Y_1, ...Y_{t-1}, X_0, X_1, ...X_{t-1}) = \Gamma(\sigma_{t-1}\mu_{t-1}, \sigma_{t-1}) \quad (7)$$

where μ is the conditional expectation of Y_t (that is, conditional on all past values of Y and X) and σ is its conditional variance. As with PEWMA, the model is estimated using a Kalman-style filter. Brandt and Williams (2001) gives the likelihood for the PAR(p) model.

3.1 Details

The PAR model differs from the PEWMA in a number of significant ways:

- The PAR dgp yields a mean-reverting (stationary) series, rather than a nonstationary one.

- The instantaneous effect of a one-unit change in X differs from the Poisson (and PEWMA) models. In the latter, its simply $exp(X_t\beta)$. For PAR, however, the impact also depends on the value of the autoregressive parameter: $(1 - \sum_{i=1}^{p} \rho_i)exp(X_t\beta)$. The difference is due to the fact that the PAR model accounts for the dynamic changes in the influence of the covariates over time, in a way that standard Poisson models don't.

- In general, the PAR model has the same stationarity restrictions on the ρs as does a linear time-series model (e.g., $|\rho| < 1$, in the AR(1) case).

- Note that both PAR(p) and PEWMA nest the Poisson model as a special case (the former when $\rho = 0$, the latter when $\omega = 0$.

- In general, one would choose the PEWMA model when the ACF shows a long-range dependence in the counts, and choose PAR(p) when the ACF tends to "die out" more quickly.

- GAUSS code for estimating both PEWMA and PAR(p) is available at Pat's website: http://www.psci.unt.edu/ brandt/pests/pests.htm

4 Autoregressive Poisson

The autoregressive Poisson model we're talking about here is one developed by Schwartz, J., Spix, C., Touloumi, G., et al. (1996). I haven't been able to get my hands on these papers yet, but here's what I know:

- The models allow for autoregression and overdispersion (which is nice – one thing that happens in time-series models of event counts is that temporal contagion yields, and looks exactly like, overdispersion).
- It can be estimated in Stata (using the `-arpois-` command).