7.3 Conditional Intensity Models

A simple non-stationary Poisson process could be modeled by a conditional intensity function such as

$$\lambda(t) = \alpha + \beta t$$

so that there is a trend over time governed by the parameter β. This model does not depend on any history, but it is clear that if $\beta > 0$ then as time increases, the rate of events also increases. One can also imagine including other trend terms like periodic functions, however, one must be careful to ensure that $\lambda(t)$ is always positive, as a negative value of the conditional intensity would not be meaningful.

Ensuring that the conditional intensity is always positive can be trickey later on when dealing with parameter estimation. One way to ensure positivity is to specify a log-linear model for $\lambda(t \mid H_t)$. Using the linear trend model above as an example, we could specify

$$\log \lambda(t) = \alpha + \beta t.$$

Clearly, the interpretations of α and β have changed, as the components of the model now multiply each other. But such an interpretation may be a feature, if the goal is to model the *risk* of events occuring. It may be more intuitive to think of risks multiplying as opposed to adding together.

One feature of a non-stationary Poisson-type of model is that given values of the parameters, we can draw the conditional intensity function without seeing any data. That is because the function itself is not dependent on any data (i.e. the history process H_t). With the cluster models described below, it is not possible to draw the conditional intensity function without the data.

7.3.1 Cluster Models

Cluster models typicaly describe processes where the occurrence of an event increases the probability of an event occurring nearby in time. For example, a main shock earthquake increases the likelihood of an aftershock occurring soon afterwards. Cluster models are sometimes called "self-exciting" because their intensity increases with each occurrence of an event. A more formal way to state this is that for two non-overlapping intervals (a, b) and (b, c), we have $\mathrm{Cov}(N(a,b), N(b,c)) > 0$, which is in contrast to the assumptions of a Poisson process.

Cluster point process models depend on the history and will typically modulate themselves based on the what has occurred in the recent past. A general class of cluster models is described by the Hawkes process, which specifies a conditional intensity of the form

$$\lambda(t \mid H_t) = \mu + \int_0^t g(t-u)\, N(du) = \mu + \sum_{t_i < t} g(t - t_i).$$

Here, we can see that the risk at time t is dependent on the entire history of the process. Each event in the past (i.e. $t_i < t$) contributes to the risk at time t by an amount $g(t - t_i)$. Typically, we might want to structure the function g such that points in the distant past have less influence than more recent points. Such a function might look like

$$g(s) = \sum_{k=1}^{K} \phi_k s^{k-1} e^{-\alpha s}$$

where the parameters ϕ_k and $\alpha > 0$ are unknown (but K is pre-specified). Here, as s grows large, the exponential term dominates and eventually tapers to zero. Therefore, points in the distant past will only contribute negligible amounts to the current risk. In the earthquake literature, g is sometimes referred to as the *trigger density*.

The interpretation of this kind of model varies with the application, of which their are many. In earthquake settings, the parameter μ is thought of as the "background rate" or the rate at which main shock earthquakes occur. Then the trigger density indicates the increased rate over the background rate at which aftershocks occur. In infectious disease settings, μ might indicate the background rate of infection and then the trigger density could describe the rate at which other individuals subsequently become infected.

This type of conditional intensity model cannot be draw in advance because its values depend on the occurrence of events in the past. Therefore, for every realization of the process, the conditional intensity function will look different, even for the same sets of parameter values.

7.3.2 ETAS Model

Other options are available for the trigger density in a Hawkes-type cluster process model. For example, the model

$$g(s) = \frac{\kappa}{(s+\phi)^\theta}$$

corresponds to the modified Omori law describing earthqake aftershock frequncy. The use of this form of g in a Hawkes type conditional intensity model is sometimes referred to as the Epidemic-Type Aftershock Sequence (ETAS) model. In addition to earthquakes, this model has been used in infectious desease settings. The key difference with this trigger density and the previous one is the power-law decay in the clustering rate, as opposed to an exponential decay in the model above.

7.3.3 Self-Correcting Models

Another type of model is a kind of “stress-release” model where we can imagine the risk of an event increasing over time until an event occurs, at which point the “stress” in the system instantaneously decreases by a fixed amount. These models are “self-correcting” in the sense that they adjust their conditional intensity each time an event occurs. In contrast with self-exciting models, we have for two intervals (a, b) and (b, c), that $\text{Cov}(N(a,b), N(b,c)) < 0$.

Such a model might have the form

$$\lambda(t \mid H_t) = \alpha + \beta t - \nu N[0, t)$$

where $\nu \geq 0$. Here, the risk increases linearly with time, but decreases by the amount ν each time an event occurs.