# The Multivariate Hawkes Process

## with Inhibition

Isabella Deutsch

University of Edinburgh
& The Alan Turing Institute

ISBA, July 2022

THE UNIVERSITY of EDINBURGH

The Alan Turing Institute

LONDON MATHEMATICAL SOCIETY EST. 1865

# RESEARCH GOAL

We aim to quantify product cannibalisation for existing products in an apparel wholesale data set.

> *Product cannibalisation refers to the decrease in sales of one product due to (the introduction of) a closely related product. [Copulsky, 1976]*

The current focus is on **inference** to detect and understand product cannibalisation for products that already have a sales history.

Very important from a business perspective!

- Find a model
- Implement that model
- Fit that model on real data

# Find A Model

**Data**: event times (plus additional covariates)
Let $N(t)$ be the number of observed events from $0$ to $t$.
**Homogeneous** ($\lambda$ constant) Poisson point process:

$$\mathbb{P}[N(t) = N] = \frac{(\lambda t)^N}{N!} e^{-\lambda t} \tag{1}$$

$$\mathbb{E}[N(t)] = \lambda t \tag{2}$$

**Inhomogeneous** ($\lambda(t)$ variable) Poisson point process:

$$p(t_1 \dots t_N) = \prod_{i=1}^{N} \lambda(t_i) \, e^{-\int_0^\infty \lambda(z) \, dz} \tag{3}$$

$$\mathbb{E}[N(t)] = \int_0^t \lambda(z) \, dz \tag{4}$$

[Daley and Vere-Jones, 2003]

Hawkes processes [Hawkes, 1971] are a class of point processes that are used to model event data when the events can occur in clusters or bursts. They are defined on the interval $[0, T]$ with **conditional intensity function**:

$$\lambda(t) = \mu(t) + \sum_{i:t>t_i} K\, g(t - t_i) \tag{5}$$

Here, $\mu(t)$ can capture seasonality and underlying trends and we use $g(t - t_i) = \beta e^{-\beta(t-t_i)}$ for the self-excitement.

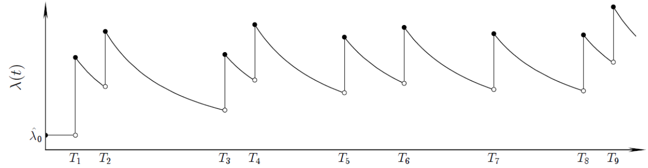


**Figure 1:** Intensity function with self exciting kernel [Rizoiu et al., 2017]

Assume that there are $M$ dimensions with data $Y_1 = (t_{1\,1} \dots t_{1\,N_1}) \dots Y_M = (t_{M\,1} \dots t_{M\,N_M})$. At time $t$ the intensity in dimension $i$ is defined as the sum of the background rate $\mu_i(t)$ and contributions from all dimensions:

$$\lambda_i(t) = \left[ \mu_i(t) + \sum_{j=1}^{M} \sum_{l:t>t_{j\,l}} K_{ji}\, g_{ji}(t - t_{j\,l}) \right]_+ \tag{6}$$

We assume the following form for the influence for all $i, j$: $g_{ij}(x) > 0$ for $x > 0$ and $\int_0^\infty g_{ij}(x)\, dx = 1$. Here, each $K_{ij} < 1$, and we write them as matrix $\mathbf{K} = \{K_{ij}\}$ where $i, j = 1 \dots M$.

We use a multivariate Hawkes Process where each dimension represents one product. This allows us to estimate the 'influence' $K_{ij}$ from an event (sale) of one product $i$ onto each product $j = 1 \dots M$.

A positive influence $K_{ij} > 0$ is called excitation, a negative influence $K_{ij} < 0$ is referred to as inhibition. The latter is interpreted as **product cannibalisation**.

# Implement That Model

To implement this model we needed to overcome a few challenges:

- Ensuring a non-negative intensity $\rightarrow$ link function
- Integrating the intensity $\rightarrow$ numerical approximation
- Checking for stability $\rightarrow$ new criterion

All methodological details can be found in the mathematical draft on ArXiv [Deutsch and Ross, 2022].

# Fit That Model on Real Data

We use a multivariate Hawkes Process where each dimension represents one product. This allows us to estimate the 'influence' $K_{ij}$ from an event (sale) of one product $i$ onto each product $j = 1 \ldots M$.

A positive influence $K_{ij} > 0$ is called excitation, a negative influence $K_{ij} < 0$ is referred to as inhibition. The latter is interpreted as **product cannibalisation**.

We fit the following multivariate Hawkes process in a Bayesian manner. The intensity in dimension $i$ is

$$\lambda_i(t) = \left[ \mu_i(t) + \sum_{j=1}^{M} \sum_{l:t>t_{j\,l}} \underbrace{\{f(\mathbf{K}^*)\}_{ji}}_{\mathbf{K}_{ji}} g_{ji}(t - t_{j\,l}) \right]_+$$

where $f(\mathbf{X}) = I - (\mathbf{X} - I)^{-1}$. For $\mu_i(t)$ we choose a step function with pre-defined change points where each product has an on-season and off-season background rate. Their priors are independent:

$$\mu_{i,\,\text{on}} \sim \mathcal{U}(0, 10) \qquad \text{for } i = 1 \ldots M$$

$$\mu_{i,\,\text{off}} \sim \mathcal{U}(0, 10) \qquad \text{for } i = 1 \ldots M$$

For the influence kernels we utilise the popular exponential kernel $g_{ij}(x) = \beta_{ij}\, exp\,(-\beta_{ij}\, x)$. Here, we assume that all $\beta_{ii} = \beta_{\text{diag}}$ and $\beta_{ij} = \beta_{\text{off}}$ when $i \neq j$.

$$\begin{aligned}\beta_{\text{diag}} &\sim \mathcal{U}(0,3)\\ \beta_{\text{off}} &\sim \mathcal{U}(0,3)\end{aligned}$$

In line with our previous arguments we place priors on the entries of $\mathbf{K}^*$. The estimation (using Stan) is carried out both using Normal priors

$$K^*_{ij} \sim \mathcal{N}(0,1) \qquad\qquad \text{for } i,j = 1 \ldots M$$

and sparsity-inducing horseshoe priors

$$\begin{aligned}\xi_{ij} &\sim \text{Cauchy}(0,1)\\ K^*_{ij} &\sim \mathcal{N}(0,\xi_{ij}) \qquad\qquad \text{for } i,j = 1 \ldots M\end{aligned}$$

| | **Product 1** | **Product 2** | **Product 3** | **Product 4** |
|---|---|---|---|---|
| Main Colour | black | black | white | white |
| Branding | white | minimal | minimal | green |
| Label | none | known | known | known |

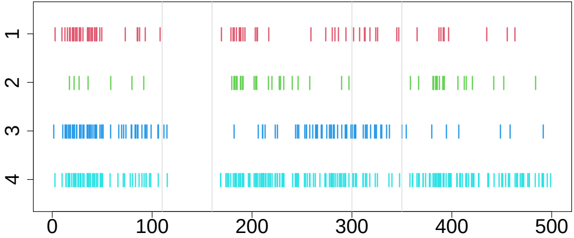


**Figure 2:** Observations

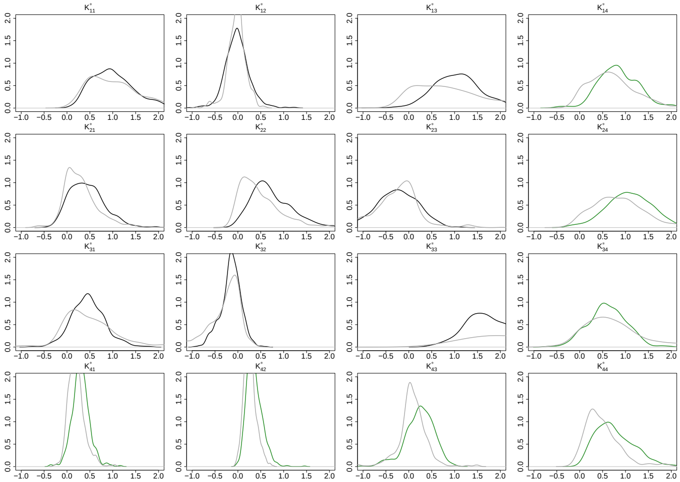


**Figure 3:** Posterior density estimates of $\mathbf{K}^*$ (excitation/inhibition parameter), using independent Normal (black) and horseshoe (grey) priors.

- Most products do not cannibalise each other as the posterior mass is mostly above zero.
- However, Product 2 and Product 3 display product cannibalisation in both directions ($K^*_{23} < 0$ and $K^*_{32} < 0$). Product 2 and 3 come from the same label. Maybe therefore wholesalers make the decision to only order one of the two due to their similar branding and label.
- Product 4 has the most sale events. It differs from the other products as it is the only one featuring a colour (green branding on the heel). This very popular style seems unaffected by product cannibalisation from other products.

# Summary

**Methodological Advances**
to make the implementation of a multivariate Hawkes Process with inhibition easier.

**Formalisation of Product Cannibalisation**
as a mathematical concept that can be estimated, monitored, and predicted.

[Deutsch and Ross, 2022]
arxiv.org/abs/2201.05009

isabella.deutsch@ed.ac.uk
isabelladeutsch.com
@BayesianBella

# References I


- Copulsky, W. (1976).
  **Cannibalism in the Marketplace.**
  *Journal of Marketing*, 40(4):103–105.

- Daley, D. J. and Vere-Jones, D. (2003).
  ***An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods.***
  Probability and Its Applications, An Introduction to the Theory of Point Processes. Springer-Verlag, New York, NY, 2 edition.

- Deutsch, I. and Ross, G. J. (2022).
  **Bayesian Estimation of Multivariate Hawkes Processes with Inhibition and Sparsity.**
  *arXiv:2201.05009 [stat].*

- Hawkes, A. G. (1971).
  **Spectra of some self-exciting and mutually exciting point processes.**
  *Biometrika*, 58(1):83–90.

# References II


- Rizoiu, M.-A., Lee, Y., Mishra, S., and Xie, L. (2017).
  **A Tutorial on Hawkes Processes for Events in Social Media.**
  *arXiv:1708.06401.*