

Hawkes process as models for some genomic data
Dr. Patricia Reynaud-Bouret
(Université Nice Sophia-Antipolis, France)

May 4, 2011

It seems that some of the genomic data, such as positions of words on the DNA or positions of transcription regulatory elements may hint for synergies between them. One of the statistical possible model to catch those interactions is the Hawkes process, which has been first introduced to model earthquakes. Gusto and Schbath have introduced this model for genome analysis. However if maximum likelihood methods exist and if AIC criterion is usually used to select a correct number of parameters, this combination has been proved to be not accurate when the complexity of the family of parametric models is high. After discussing the Hawkes models (multivariate or not) and explaining what has been done from a parametric point of view (eventually combined with AIC), I will explain what adaptive model selection can and cannot do and also what thresholding in certain cases and Lasso methods may improve.