# MT4537 - Project 1

A. Johnston

01/03/2022

This project accounts for 10% of your overall grade. It is due at 11.59 on Wednesday the Wed 16th March. Submit your project through MMS, which you can access through MySaint.

## Project outline

There are two main phases to this project - a series of simulations and then fitting models to data.

### Simulations

- Create a simulation for a Thomas cluster process, using the theory of a Cox process. Devise a random driving intensity $\Lambda(x)$ which comprises Gaussian densities at parent locations. (See Baddeley et al section 12.3.3 "Thomas cluster process" for a visual of this).
- Create the simulation using thinning ($P(x)$-thinning, where $p(x)$-thinning is a single realisation of this).
- Although your simulation will be general, present two example simulations.
- For each of the simulations, fit a Thomas and Matérn process model using the minimum contrast method (use `spatstat` for this). Describe the minimum contrast method (including which summary characteristic you've used).
- Clip the simulations to some nominal window prior to the model fitting and specify an edge-correction.
- Can you distinguish between the two models based on their fits? How might you determine which is more appropriate?

### Fitting to data

We'll use the `spatstat` dataset `bei` and `bei.extra` for the following part. Refer to the helpfiles for details about the dataset, but in short provides the locations for over 3000 trees (*Beilschmiedia*) in a tropical rainforest. There are additional covariates for elevation/altitude and slope of elevation (gradient).

- Use a kernel smooth to give an initial estimate of the intensity function (e.g. `density` function). Determine which kernel is being used, what the bandwidth is and how this was selected. Try other bandwidths, or bandwidth selection methods. Present two markedly different intensity functions.
- Fit models relating the location of the trees to covariates, using the `spatstat` function `ppm`. Fit models with the following covariates:
  - The $x$ and $y$ coordinates, polynomial variants of these and interaction terms. You can choose a variety of complexities as you see fit, but select between a few candidates on the basis of AIC (don't do a comprehensive search over complexity - only consider a few). Describe the models considered and indicate which was favoured and why. Present the intensity function for the best of these.
  - Combinations of the altitude, gradient and an interaction of these. Select between these models on the basis of AIC. Interpret the best of these models (e.g. the parameter estimates). Present

the fitted intensity function and examine inhomogeneous summary characteristic functions *based on your fitted model*. Interpret these.

- Using the model structure from your preferred model in the previous question, refit the model as a Cox model where the scattering process is Gaussian (i.e. a Thomas process). Interpret the outputs of the fitted model (in particular you will now have more estimated parameters). Plot the fitted intensity function, inhomogeneous summary characteristic functions, and interpret.
  - Provide 2 simulations from this fitted model.
  - Which of the models is best in your view, and why?

## Submission

The submission will be a PDF report, potentially with an associated code file.

#### Simulations

- A description of your simulation process.
- Commented code as an appendix or link to a repository.
- Two realisations from your simulation, showing the point pattern and driving intensity.
- The fitted parameters for the two cluster models indicated. Include the pair-correlation functions for each, with interpretation.

#### Models fitted to data

- Model outputs and comparisons as requested i.e. AICs where required, brief descriptions of the models, fitted intensity surfaces for those indicated.
- Interpretations of fitted parameters for the simple inhomogeneous Poisson point model and the subsequent Cox model.
- Simulations for the final model.
- Commentary as relevant throughout.

## Assessment

You'll be assessed by:

- Whether your simulation process does what it is supposed to.
- Your understanding of the process as assessed by the code, commentary and description.
- Your understanding of the models being fitted and the interpretation of outputs.
- The readability/quality of your report - make it clear, concise and presentable.