Available projects with Ben Stevenson

November 15, 2018

Please e-mail me at ben.stevenson@auckland.ac.nz if you are interested in any of the projects listed below.

A comparison of spatial capture-recapture and random encounter models for camera trap data

Camera-trap surveys are commonly used to estimate density of wildlife populations. Over the last decade, spatial capture-recapture (SCR) and random encounter models (REMs) have gained traction in their application to the resulting data. They each require slightly different information—for example, SCR usually needs individuals to be recognised when they are detected, while REMs usually require *a priori* knowledge of average animal speeds. The two methods also make different assumptions about the way animals move and behave. This project will assess the performance of SCR and REM estimators in terms of properties such as bias, precision, and robustness to assumption violations.

Requirements: An interest in programming, with good grades in statistical computing papers.

Parameter estimation for void point processes

Spatial point processes model the observed locations of objects or events. They can explain patterns in the locations of stars, animals, earthquakes, criminal acts, and terrorism attacks, to name just a few applications. Jones-Todd et al. (in press) described a new type of spatial point process, the void process, which was used to model spatial patterns formed by cancer cell nuclei on tissue slides from colon cancer patients. A void process comprises two types of points, Type A and Type B, which are uniformly scattered with intensities λ_A and λ_B, respectively. Type B points are not observed at all. Only Type A points that are greater than distance τ from all Type B point are observed. In other words, each hidden Type B point 'deletes' any Type A point within its 'void' of radius τ. Jones-Todd et al. (in press) proposed estimation of λ_A, λ_B, and τ via maximum Palm likelihood. Although this is a computationally efficient approach, it does not appear to be particularly statistically efficient. This project will develop a Bayesian (or possibly a maximum likelihood) estimator for void processes.

Requirements: A good grade in STATS 331 or STATS 731. A good grade in STATS 310 would also be an advantage.

Reference: Jones-Todd, C. M., Caie, P., Illian, J. B., Stevenson, B. C., Savage, A., Harrison, D. J., and Bown, J. L. (in press) Identifying prognostic structural features in tissue sections of colon cancer patients using point pattern analysis. *Statistics in Medicine.*

Who's skating on thin ice? Evaluating the performance of hockey goaltenders in the NHL

North American sports are well known for their meticulous recordkeeping, and ice hockey is no different. Vast quantities of data are collected from the performance of every player who steps onto the ice during an NHL match. The goaltender (or 'goalie') is an important player who attempts to prevent the opposition from scoring by blocking shots directed at their goal. The save percentage (SP) is a key summary statistic that is commonly used to measure goaltender performance. However, comparison of the SP either between players (to assess who is better) or within a single player (to assess how performance changes over time) is typically achieved simply by seeing which is bigger, without any regard for sampling variation. This project aims to model goaltending performance in order to answer various questions, which may include some of the following:

1. How much better is a 'good' goaltender than a 'bad' one? How much does a 'good' goaltender improve the performance of their team?
2. How does goaltender skill vary across a career? Is there an optimal age?
3. Do goaltenders with higher salaries actually perform any better? Is it better to hire an older, experienced, expensive goaltender, or a younger, inexperienced, cheap goaltender?
4. What role does form play? For example, how does a good goaltender on a bad day compare to a bad goaltender on a good day?
5. Is it possible to predict how well a goaltender will perform based on recent performances? In other words, do goaltenders go through cycles of 'good' and 'bad' form? If so, should a coach pick their best goaltender in 'poor' form or a back-up goaltender in 'good' form?
6. If a coach deems that the starting goaltender is performing poorly in a match, they may be 'relieved' by the back-up goaltender. When this occurs, how often is the goaltender truly performing poorly, and how often have they just gotten unlucky? Is there any evidence to suggest back-up goaltenders perform any better, or does being introduced mid-game negatively affect their performance?

Requirements: At least two of the following would be beneficial, but a particularly good student could get away with fewer: (1) Experience with hierarchical models (or 'latent variable models' or 'mixed effects models'). (2) A good grade in statistical inference papers, such as STATS 310 and/or STATS 730. (3) Some experience using Template Model Builder (TMB). (4) A good understanding of ice hockey.