Final Project, Task 1: Description of paper and model

Demorest, Lemire and Wilson

1. A brief description of the background motivation of your chosen article and a *single* research question you will choose to investigate that was presented in the paper. Note that your article may address more than 1 question but I want you to focus on just 1 for this project.

2. A brief description of the methods used by the authors of your paper to address the question of interest you are focusing on.

Data

The participants in the original study were professional baseball and soccer players, but we will focus only on professional soccer. The Japan Professional Football League (J. League) consists of 40 teams, representing a total of 1013 players registered in the 2012 season and we will focus on players between the ages of 23 and 25, for a total of 227.

The school year in Japan begins April and ends on March of the following calendar year, which corresponds with the competitive season of most professional sports. Birth date was treated as monthly data and because school year and competitive season both begin in April in Japan, relative age was coded as 0 (April) to 11 (March).

An athlete's birthplace is defined as the prefecture the player was born in. A prefecture is a first-order administrative district in Japan, and there are 47 prefectures in total.

The author's model includes the total number of male children born in each month as an offset term, which is a variable whose coefficient is fixed at one. The variable in the estimated equation is the total number of male births over the years when the sampled players were born.

Methods

Becoming a soccer player can be seen as an "event", and a number of them can be regarded as a "count of events". Thus, the author applied a Bayesian hierarchical Poisson regression model.

3. The full probability model used in the paper written out in mathematical notation. That is, all likelihood and prior components.

$$y_{ij} \sim \text{Poisson}(\lambda_{ij}) \text{ for } i = 1, \ldots, 12; j = 1, \ldots, 47$$

$$\lambda_{ij} = \theta_{ij} exp\left\{ a_j + \beta RA_i \right\}$$

y_{ij} is the number of professional sports players where subscripts i, j indicate birth month and prefecture, respectively. θ_{ij} is the total number of male children born in month i in prefecture j. The a_j intercept term is meant to capture the differences in the likelihoods between birthplaces.

RA_i is the relative age of those born in month i, while the coefficient β measures the relative age effect (RAE). For example, April would be month 1, with $RA_1 = 0$. The exponential term gives the probability of becoming a professional soccer player.

The authors found the posterior estimate for β to be $\beta \sim \text{Normal}(-0.0934, 0.214^2)$ which we will use as our prior.

We did not see an updated posterior for a_j so we assume the prior in the paper where $a_j \sim \text{Normal}(\mu_a, \sigma_a^2)$ where $\mu_a \sim \text{Normal}(0, 100^2)$ and $\sigma_a^2 \sim \text{Uniform}(0, 100)$.

Introduction

The Relative Age Effect (RAE) is a term used to describe how those born early in the academic year tend to have an advantage both athletically and academically. An earlier birth is typically associated with increased physical ability and this advantage may occur because those who are older are typically more physically, emotionally or cognitively developed than those who are younger. Additionally, individuals born earlier in the year may be more likely to persevere in sports since they are able to actively dominate in physical activities and key performance measures. This suggests that as time goes on, those who are less mature drop out of sports as they are not as successful, motivated, or fulfilled by the experience in their younger years of competition, contributing further to their underperformance.

In the paper *Relative age and birthplace effect in Japanese professional sports: a quantitative evaluation using a Bayesian hierarchical Poisson model*, Hideaki Ishigami explores the effect sizes of the relative age and birthplace of becoming a professional athlete for two of the most popular sports in Japan, soccer and baseball. Previous studies suggest that where an athlete was born and developed is another important factor for determining the likelihood of becoming an elite athlete. In Japan, for example, "northern areas may receive, on average, over 100 snow days per year, whereas in the southern areas there are few to no snow days" [@RAE]. As soccer and baseball are not typically played in snow, an athlete's birthplace may have an effect on their athletic development.

Code Appendix

References