Relative Age Effect in Japanese Professional Soccer

Demorest, Lemire and Wilson

12/5/2021

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. Much research has been done to look at RAE in North American and European athletes. However very little research has attempted to extend findings surrounding RAE to other parts of the world, specifically Asia. This is why we are interested in Hideaki Ishigami's work in the paper *Relative age and birthplace effect in Japanese professional sports: a quantitative evaluation using a Bayesian hierarchical Poisson model* (published in the Journal of Sports Sciences in 2016). In our paper we will replicate Ishigami's modelling process using simulated data. In order to answer the question: what is the RAE on soccer players in Japan between the ages of 23 and 25 in the 2012 season?

Our simulation and modeling process will mirror the author's study design in which he uses data sourced from the Japan Professional Football League (J. League) from the 2012 soccer season. The J. League data consist of 40 teams, representing a total of 1013 players registered in the 2012 season and since we will focus on players between the ages of 23 and 25, we have 227 observations. The author characterized "becoming a professional sports player" as an event and thus we are dealing with discrete count data. Using a poisson regression model it is possible to estimate the magnitude of the RAE on the likelihood of becoming a professional athlete. More details on the poisson model can be found below. The authors then ran an MCMC algorithm [1] using JAGS (Plummer, 2012) and confirming with Stan (Stan Development Team, 2013).

The work replicated in our paper is an important step in extending the RAE's scope of inference. Additionally this paper and our analyses allow the magnitude of relative age to be quantified. In contrast many other analyses have simply stated whether an association is statistically significant using, for example, a χ^2 test. Although it is beyond the scope of this paper, these results impact parents' decisions surrounding red shirting and informs public understanding of factors influencing success in athletics (particularly professional athletics).

Gelman and Shalizi (2012)

Our research question of interest is: what is the RAE on soccer players in Japan between the ages of 23 and 25?

[1] The MCMC ran for 25,000 iterations with five chains where the first 5,000 samples were discarded as burn-in. Only every 100 iteration was saved for a total of 1,000 MCMC samples.

Model

References

Data

Results

Posterior Predictive Checks

Assessing Convergence with $\hat{R}$ The authors used the Gelman-Rubin statistic (Gelman & Rubin, 1992), $\hat{R}$, which assesses convergence by comparing the estimated between-chains and within-chain variances for each model parameter.

Assessing Convergence with Geweke's Diagnostic Additionally, the authors used the Geweke's convergence diagnostic (Geweke, 1992) to check the convergence of the MCMC algorithms. The Geweke convergence diagnostic is a test for equality of the means of the first and last part of a Markov chain, typically the first 10% and the last 50%. If the samples are drawn from a stationary distribution of the chain, then the two means are equal and Geweke's statistic has a standard normal distribution. The test statistic is a standard Z-score: the difference between the two sample means divided by its estimated standard error. The standard error is estimated from the spectral density at zero, and so takes into account any autocorrelation and the Z-score is calculated under the assumption that the two parts of the chain are asymptotically independent. Both of these tests are available in the `coda` package in R.

Conclusion

References

Gelman, Andrew, and Cosma Rohilla Shalizi. 2012. "Philosophy and the Practice of Bayesian Statistics." *British Journal of Mathematical and Statistical Psychology* 66 (1): 8–38. https://doi.org/10.1111/j.2044-8317.2011.02037.x.

Ishigami, Hideaki. 2016. "Relative Age and Birthplace Effect in Japanese Professional Sports: A Quantitative Evaluation Using a Bayesian Hierarchical Poisson Model." *Journal of Sports Sciences* 34 (2): 143–54. https://doi.org/10.1080/02640414.2015.1039462.