Racing and Wagering Western Australia

Evaluation of Box Draws

Berwin A. Turlach*

27 July 2012

*Centre for Applied Statistics (M019), The University of Western Australia, 35 Stirling Highway, Crawley WA 6009, Australia.

1 Introduction

The Centre for Applied Statistics (CAS) of the University of Western Australia (UWA) was contacted by Mark Bottcher (Manager Greyhound Racing) from *Racing and Wagering Western Australia* and asked to perform a statistical analysis of results from their box draws algorithm for greyhounds races. Specific questions of interest are whether (a) each greyhound is equally likely to be assigned to each of the eight boxes and (b) whether the draw from any one race is not related to the draw of other races.

The box draw randomiser within the software developed by *Racing and Wagering Western Australia* relies on the CryptGenRandom function (`http://en.wikipedia.org/wiki/CryptGenRandom`) provided by Microsoft's Cryptographic Application Programming Interface. For the analysis, CAS was provided with a file (draws20thJune2012.csv) that contained 100,000 simulated draws; the amount of races that are held approximately over 3 years.

2 Analysis

Various graphical and numerical summaries were used to investigate whether there is any evidence that the box draw data supplied in the file draws20thJune2012.csv were produced in a manner that does not have the required desired properties, namely that (a) each greyhound is equally likely to be assigned to each of the eight boxes and (b) the draw from any one race is not related to the draw of other races. Specifically, the analyses performed included:

- For each runner/box combination a plot was produced that showed how the relative frequency with which the runner was assigned to the box evolved over the simulated 100,000 draws. If draws are produced in the desired manner then, in the long run, in each plot one would expect these relative frequencies to approach 0.125. The speed with which the relative frequencies should approach 0.125 is also quantifiable. The plots were inspected to see whether they showed the expected behaviour.
- For each runner, the sequence of boxes to which it was allocated was treated as a time series and the auto-correlation function of this series calculated to check whether there is a noticeable correlation between box positions over several races. Likewise, for each pair of runners the cross-correlation function of their box allocations was calculated to check whether the position of any pair of runners is correlated over several races.
- χ^2 goodness of fit tests were used to investigate
  - for each runner, whether it was equally likely to be assigned to any of the boxes,
  - for each pair of runners, whether they were equally likely to be assigned to any possible pair of boxes,
  - for each triplet of runners, whether they were equally likely to be assigned to any possible triplet of boxes; and
  - for each group of four runners, whether they were equally likely to be assigned to any possible quadruplet of boxes.

  The χ^2 goodness of fit test statistics were evaluated for the complete set of 100,000 draws. Additionally, the set of draws was also split into several non-overlapping blocks (using a range of sizes for these sub-blocks) and the χ^2 goodness of fit test statistics were evaluated within each block.

3 Conclusion

Based on the analysis performed, we found no evidence that the box draw data supplied in the file draws20thJune2012.csv were not produced in a manner that has the desired properties. Overall, the graphical and numerical results were what one would expect to see if the draws are produced such that (a) each greyhound is equally likely to be assigned to each of the eight boxes and (b) the draw from any one race is not related to the draw of other races.