Racing and Wagering Western Australia
An Audit of Box Draw Data
SA Data

Berwin A Turlach[1]

09/12/2024

Disclaimer

This report is an official document of the University of Western Australia, compiled with the expertise available to the Centre for Applied Statistics. Whilst all reasonable care is used in all work produced by the Centre for Applied Statistics, neither the University nor the Centre represents that the work is suitable in respect of any particular use or function. Accordingly the University and the Centre as well as the authors of the report accept no liability in respect to any use or reliance upon such work unless specifically acknowledged in writing.

[1] Centre of Applied Statistics (M019), The University of Western Australia, 35 Stirling Highway, Perth WA 6009, Australia

1 Introduction

The Centre for Applied Statistics (CAS) of the University of Western Australia (UWA) was contacted by Colin Stutley (IS Business Analyst) from *Racing and Wagering Western Australia* (RWWA) and asked to perform a statistical analysis of results from actual greyhounds races in various Australian states and territories (NSW, NT, QLD, SA, TAS and WA) in which RWWA's box draw algorithm was used. This box draw algorithm is supposed to randomly allocate, for a given race, the greyhounds that run in the race to the boxes that are used in the race[2], such that each greyhound is equally likely to be allocated to any of the boxes used.

Specific questions of interest are whether there is any evidence in the data that the box draw algorithm does not perform as designed when looking at:

1. the allocation of individual greyhounds to boxes over all the races that they ran in,
2. the allocation of the greyhound(s) of any owner to boxes over all races in which that owner had greyhound(s) running,
3. the allocation of the greyhound(s) of any trainer to boxes over all races in which that trainer had greyhound(s) running, and
4. the allocation of greyhounds trained by the same trainer to neighbouring boxes, if a trainer had multiple greyhounds in the same race.

The analysis was proposed to be done for each state and territory separately. The following section will describe the methodology used to analyse the data followed by the results of the analyses for each state and territory in Section 3. A summary conclusion from all these analyses is given in Section 4.

2 Methodology

All analyses were performed using the statistical computing environment R (R Core Team, 2024). To address the questions of interest, the data were analysed using a Monte Carlo permutation test approach. This approach is described next in some details for the analyses of the data on individual greyhounds.

First, a contingency table was calculated by cross tabulating the data on greyhound IDs with the data on box allocations. This yields an $M \times 8$ table[3] with entries $o_{ij,\text{obs}}$, where

[2] Typically, 8 greyhounds run in a race and 8 boxes are used. But there are races with less then 8 greyhounds, in which specific boxes should not be used, e.g. in races with 7 greyhounds box 5 should not be used.

[3] Different states and territories have, of course, different values for M. Details are given in the appropriate subsection of Section 3.

$o_{ij,\text{obs}}$ is the number of times greyhound i ($i = 1, \dots, M$) was allocated to box j ($j = 1, \dots, 8$). Next, based on the boxes used in the races in which each greyhound started, we calculated e_{ij}, the expected number of times that greyhound i should be allocated to box j (under the assumption that in each race it is allocated equally likely to any of the boxes used in that race).

Based on these two quantities, we can calculate for each greyhound a Pearson's χ^2 style statistic for how well the observed frequencies match the fitted frequencies. These individual statistics are:

$$\chi^2_{i,\text{obs}} = \sum_{j=1}^{8} \frac{\left(o_{ij,\text{obs}} - e_{ij}\right)^2}{e_{ij}}, \qquad i = 1, \dots, M$$

A test statistic for judging whether the overall observed contingency table is consistent with the expected frequencies is then given by $\chi^2_{\text{all,obs}} = \sum_{i=1}^{M} \chi^2_{i,\text{obs}}$.

As long as a greyhound started in sufficiently many races, such that for all boxes that the greyhound was allocated to the expected number of times that the greyhound is allocated to each box is at least 5, a χ^2 distribution would be a suitable reference distribution for its individual test statistic $\chi^2_{i,\text{obs}}$. However, for greyhounds who participated in a relative small number of races it is not immediately clear what the correct reference distribution for their individual test statistics $\chi^2_{i,\text{obs}}$ would be. For the overall test statistic $\chi^2_{\text{all,obs}}$ this issue is compounded with the additional issue that its individual components are stochastically dependent in a rather subtle way.

A Monte Carlo permutation test approach is one way of addressing these issues. In this approach the box allocations for each race are repeatedly randomly permuted. For the results reported in Section 3, the number of times B that the box allocations are randomly permuted was chosen to be 10,000. For each of these random permutations, the contingency table obtained by cross tabulating data on greyhound ID with data on (permuted) box allocations is recalculated to obtain $o_{ij,k}$, the number of times greyhound i ($i = 1, \dots, M$) was allocated to box j ($j = 1, \dots, k$) in permutation k ($k = 1, \dots, B$)[4]. All statistics are then re-evaluated on each contingency table based on permuted data yielding values $\chi^2_{i,k}$ ($i = 1, \dots, M$) and $\chi^2_{\text{all},k}$ for $k = 1, \dots, B$. These values form (an empirical approximation of) the reference distribution of the test statistics $\chi^2_{i,\text{obs}}$ and $\chi^2_{\text{all,obs}}$ under the assumption that the box draw algorithm works as designed.

This allows us to calculate a p-value associated with the (null) hypothesis that the observed pattern of box assignments for greyhound i is consistent with the expected pattern as follows:

$$p_i = \frac{\#\{k: \chi^2_{i,k} \geq \chi^2_{i,\text{obs}}\} + 1}{B + 1}$$

[4] Note that the expected allocations e_{ij} do not change.

where $\#\{k: \chi^2_{i,k} \geq \chi^2_{i,\text{obs}}\}$ is the number of $\chi^2_{i,k}$ values $(k = 1, \dots, B)$ that are larger or equal to $\chi^2_{i,\text{obs}}$.

Likewise, we can calculate a p-value associate with the (null) hypothesis that the overall observed pattern of box assignments is consistent with the expected pattern as follows:

$$p_{\text{all}} = \frac{\#\{k: \chi^2_{\text{all},k} \geq \chi^2_{\text{all,obs}}\} + 1}{B + 1}$$

where $\#\{k: \chi^2_{\text{all},k} \geq \chi^2_{\text{all,obs}}\}$ is the number of $\chi^2_{\text{all},k}$ values $(k = 1, \dots, B)$ that are larger or equal to $\chi^2_{\text{all,obs}}$.

A small value for p_{all} would indicate that the observed overall pattern of box allocations is not consistent with the assumption that greyhounds are randomly allocated to boxes with each greyhound being equally likely to be allocated to any of the boxes used. In Section 3 we use a 1% significance level, i.e. any p-value p_{all} less than 1% will be declared (statistically) significant indicating that the data provides evidence that the box draw algorithm does not allocate greyhounds to boxes such that in every race every greyhound is equally likely to be allocated to any of the boxes.

When assessing the M p-values p_i, $i = 1, \dots, M$, for the box allocations for each individual greyhound, care has to be taken of the fact that M (simultaneous) tests are performed and, already for moderately large M, one would expect some of the p-values to be, just by chance, below conventionally used significance levels of 1% or 5%. The discussion in Section 3 addresses this issue in two ways. First a 1% significance level with a Bonferroni correction is used for performing the multiple tests, i.e. the p-values are compared to $1/M$. This Bonferroni correction is designed to control the *probability* of incorrectly declaring one or more p-values as being sufficiently small, i.e. to declare for one or more greyhounds that their observed allocation to boxes is not consistent with the assumption that they are allocated in each race randomly to their boxes with each box being equally likely, when performing M simultaneous tests.

The second approach employed to address the problem of performing multiple tests, is the usage of *false discovery rate* (FDR) methods. FDR methods attempt to control the *expected proportion* of rejected null hypotheses that were incorrect rejections. Here, as the p_i are not independent of each other, we used the FDR procedures proposed by Benjamini and Yekutieli (2001) and Blanchard and Roquain (2008) (see also Sarkar 2008) as implemented in the `R` package `mutoss` (Team *et al.*, 2023).

Both approaches, the Bonferroni correction method and the FDR methods, are used on all p-values p_i, $i = 1, \dots, M$, but also on the subset that corresponds to greyhounds that started at least 40 times. The rationale for looking at this subset only is based on the recommendation for the standard χ^2 test that each expected frequency should be at least 5. As there are 8 boxes this recommendation would be fulfilled if a greyhound starts in 40 races with 8 runners.

The analyses for the allocation of greyhounds owned by the same owner and the allocation of greyhounds trained by the same trainer were performed in an analogous manner.

Several issues arose for the analysis of the frequencies with which greyhounds trained by the same trainer were allocated to neighbouring boxes:

- If the race has only 7 runners and box 5 is not used, are boxes 4 and 6 regarded as neighbouring boxes? Likewise, how should races with even less runners be treated?
- If a trainer has 3 runners in a race and they are allocated to 3 consecutive boxes, does this count as two allocations to neighbouring boxes? Or should the analysis look at all possible cases:
    - None of the 3 greyhounds in neighbouring boxes.
    - Two of the 3 greyhounds in neighbouring boxes and the third in a box separated by at least one other box from the first two.
    - All 3 greyhounds in consecutive boxes.

  How to handle all the various possibilities if a trainer has more than 3 runners in a race?

To address these issues, a decision was taken to consider only races with 8 runners and only those cases in which a trainer has exactly 2 runners in the race. Under this scenario, if greyhounds are randomly assigned to boxes with every greyhound being equally likely to be assigned to any of the boxes used, the probability that the two greyhounds of a trainer are assigned to neighbouring boxes in a race with 8 runners is 0.25.

Thus, every trainer who had exactly two greyhounds running in at least one race was used in the cross tabulation of how often the two greyhounds of the trainer were allocated to neighbouring boxes or not. In terms of the description of the analysis above, an $M \times 2$ contingency table was created with M being the number of trainers that had exactly two greyhounds running in at least one race. The frequencies o_{i1} and o_{i2} counted how often the two greyhounds of trainer i $(i = 1, \ldots, M)$ were allocated to non-neighbouring and neighbouring boxes, respectively. The expected frequencies were calculated as $e_{i1} = \frac{3}{4} n_i$ and $e_{i2} = \frac{1}{4} n_i$, where $n_i = o_{i1} + o_{i2}$ is the number of races in which trainer i had exactly two greyhound running. The remainder of the analysis followed the Monte Carlo permutation test approach outlined above. The only deviation from the procedure described above is that when analysing a subset of p-values for individual trainers the chosen subset included the p-values of all those trainers who had at least 20 times exactly 2 greyhounds running in a race[5].

3 Results for Races in South Australia

The provided spreadsheet contained information on 19,209 races held at 1,694 meetings between 1 January 2020 and 31 December 2023. There were 6,012 unique greyhound IDs, 1,984 unique owner IDs and 514 unique trainer IDs.

[5] With $n_i \geq 20$, we have $e_{i1} \geq 15$ and $e_{i2} \geq 5$.

3.1 Allocation of individual greyhounds

The analysis of the actual overall allocation of greyhounds to boxes provided no evidence against the assumption that in every race every greyhound is equally likely to be assigned to any one of the boxes used ($p_{\text{all}} = 0.609$). Analysing each individual greyhound's allocation to boxes, taking into account that this involves the analysis of 6,012 p-values by either using a 1% significance level with a Bonferroni correction or controlling the false discovery rate at 1%, provided no evidence that the box allocation of any individual greyhound was suspicious. Analysing the 1,033 p-values for the subset of greyhounds that started in at least 40 races yielded the same result, i.e. no evidence was found that the assumption, that each greyhound is equally likely to be assigned to any one of the boxes used in races in which the greyhound ran, is not tenable.

3.2 Allocation of greyhounds by owner

The analysis of the actual overall allocation of greyhounds owned by the same owner to boxes provided no evidence against the assumption that in every race every greyhound is equally likely to be assigned to any one of the boxes used ($p_{\text{all}} = 0.751$). Analysing for each individual owner the allocation of his or her greyhound(s) to boxes, taking into account that this involves the analysis of 1,984 p-values by either using a 1% significance level with a Bonferroni correction or controlling the false discovery rate at 1%, provided no evidence that the box allocation of greyhound(s) belonging to the same owner was suspicious for any of the owners. Analysing the 659 p-values for the subset of owners who had greyhounds starting at least 40 times yielded the same result, i.e. no evidence was found that the assumption, that the greyhound(s) of any owner is/are equally likely to be assigned to any of the boxes used in races in which the greyhound(s) ran, is not tenable.

3.3 Allocation of greyhounds by trainer

The analysis of the actual overall allocation of greyhounds trained by the same trainer to boxes provided no evidence against the assumption that in every race every greyhound is equally likely to be assigned to any one of the boxes used ($p_{\text{all}} = 0.862$). Analysing for each individual trainer the allocation of his or her greyhound(s) to boxes, taking into account that this involves the analysis of 514 p-values by either using a 1% significance level with a Bonferroni correction or controlling the false discovery rate at 1%, provided no evidence that the box allocation of greyhound(s) trained by the same trainer was suspicious for any of the trainers. Analysing the 269 p-values for the subset of trainers who had greyhounds starting at least 40 times yielded the same result, i.e. no evidence was found that the assumption, that the greyhound(s) of any trainer is/are equally likely to be assigned to any of the boxes used in races in which the greyhound(s) ran, is not tenable.

3.4 Allocation of greyhounds to neighbouring boxes

Of the 19,209 races in the spreadsheet, 5,761 were races with 8 runners and in 2,830 of those at least one trainer had exactly two greyhounds running. Altogether there were 226 trainer IDs involved and 3,687 cases in which trainers had exactly two greyhounds running in a race.

The analysis of the overall contingency table provided no evidence against the hypothesis that allocation mechanism works as it is designed to do ($p_{\text{all}} = 0.512$). Analysing the results for individual trainers, taking into account that this involves the analysis of 226 p-values by either using a 1% significance level with a Bonferroni correction or controlling the false discovery rate at 1%, provided no evidence that any trainer has his or her two greyhounds allocated more often to neighbouring boxes than what would be expected under the allocation mechanism. Analysing the 38 p-values for the subset of trainers who have exactly two greyhounds starting in at least 20 races yielded the same result, i.e. no evidence was found that the assumption, that the two greyhounds of any trainer are equally likely to be assigned to any of the 8 boxes, is not tenable.

4 Conclusions

RWWA's box draw algorithm is supposed to allocate, in any given race, the participating greyhounds to boxes in such a manner that each greyhound is equally likely to be assigned to any of the employed boxes. The summary conclusion from all the analyses reported in Section 3 is that there is no evidence in the data to suggest that RWWA's box algorithm does not work as designed when analysing the data with respect to

- the allocation of individual greyhounds to boxes over all the races that they ran in,
- the allocation of the greyhound(s) of any owner to boxes over all races in which that owner had greyhound(s) running,
- the allocation of the greyhound(s) of any trainer to boxes over all races in which that trainer had greyhound(s) running, and
- the allocation of greyhounds trained by the same trainer to neighbouring boxes, if a trainer had two greyhounds in the same race.

5 References

Benjamini, Y. and Yekutieli, D. (2001) The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, **29**, 1165–1188.

Blanchard, G. and Roquain, E. (2008) Two simple sufficient conditions for FDR control. *Electronic Journal of Statistics*, **2**, 963–992. Available at: https://doi.org/10.1214/08-EJS180.

R Core Team (2024) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.R-project.org/.

Sarkar, S. K. (2008) On methods controlling the false discovery rate. *Sankhyā: The Indian Journal of Statistics, Series A*, **70**, 135–168. Available at: http://www.jstor.org/stable/41234409.

Team, M. C., Blanchard, G., Dickhaus, T., et al. (2023) *Mutoss: Unified Multiple Testing Procedures*. Available at: https://CRAN.R-project.org/package=mutoss.