Time-rank duality in a simple model for Formula 1 racing

John Fry[*], Tom Brighton[†] and Silvio Fanzon[‡]

December 2023

Abstract

Two natural ways of modelling Formula 1 race outcomes are a probabilistic approach, based on the exponential distribution, and statistical regression modelling of the ranks. Both approaches lead to exactly soluble race-winning probabilities. Equating race-winning probabilities leads to a set of equivalent parametrisations. This time-rank duality is attractive theoretically and leads to new ways of dis-entangling driver and car level effects as well and a simplified Monte Carlo simulation algorithm. Results are illustrated by applications to the 2022 and 2023 Formula 1 seasons.

Keywords: Exponential Distribution; Formula 1; Regression; Time-rank duality.
2020 Mathematics Subject Classification: 62P25, 91-10, 62M99, 65C05.

1 Introduction

The modelling of Formula 1 race outcomes is both an interesting statistical problem in its own right (Bell et al., 2016; van Kesteren and Bergkamp, 2023) and of significant wider interest (Maurya, 2021). Building on recent analytical modelling of sports (see eg. Baker et al., 2022; Singh et al., 2023), including classical Poisson models for soccer (Maher, 1982), intra-season race outcomes are most easily modelled assuming car finishing times correspond to a sequence of independent exponential random variables. Under this assumption race-winning probabilities can be written down in closed from. This tractability also enables relatively easy model calibration using readily available bookmakers' betting odds. See Section 2.

However, this probabilistic approach is at odds with much of the publicly-available race data. Race finishing times are typically not available in the conventional sense. For example, lapped cars do not typically finish the full race. In contrast, the most convenient statistical approach for

[*]Centre for Mathematical Sciences, School of Natural Sciences, University of Hull, Hull, HU6 7RX, UK. Email: J.M.Fry@hull.ac.uk
[†]Centre for Mathematical Sciences, School of Natural Sciences, University of Hull, Hull, HU6 7RX, UK. Email: thomasbrighton02@gmail.com
[‡]Centre for Mathematical Sciences, School of Natural Sciences, University of Hull, Hull, HU6 7RX, UK. Email: S.Fanzon@hull.ac.uk

publicly-available race data seems to be a regression model for the final race ranking obtained (Eichenberger and Stadelmann, 2009).

Thus, in this paper we provide a hybrid approach that combines the elegance and simplicity of the probabilistic approach with the practical utility of classical regression modelling. The duality of both approaches is established in Section 3. Firstly, we show that a regression model for ranks can be used to estimate race-winning probabilities and then to an equivalent exponential-distribution parameterisation using the method in Section 2. Secondly, we show that under the simplifying assumption of homoscedasticity appropriate regression parameters can be reverse-engineered from a set of race-winning probabilities e.g. those corresponding to a given exponential-distribution parameterisation or a set of bookmakers' odds. Combined use of both probabilistic and statistical approaches then leads to new insights regarding driver-level versus car-level effects (Section 5) as well as a simplified Monte Carlo simulation algorithm (Section 6).

The layout of this paper is as follows. Section 2 outlines a simple probabilistic approach to modelling race finishing times and model calibration via bookmakers' odds. Section 3 establishes theoretical duality between this probabilistic approach and statistical regression modelling of the final rank. Section 4 discusses the empirical regression modelling of historical results. Sections 5-6 make further use of the duality between the probabilistic and statistical approaches. Section 5 discusses combined use of both approaches to disentangle driver-level and car-level effects. Previously, such an analysis has only been possible over longer time periods (Bell et al., 2016; Eichenberger and Stadelmann, 2009; van Kesteren and Bergkamp, 2023). A Monte Carlo simulation algorithm based around converting statistical regression output to an equivalent exponential-distribution parameterisation is outlined in Section 6. Section 7 concludes and discusses the opportunities for future research.

2 Probabilistic approach to modelling finishing times

Classical queuing theory suggests models based around the exponential distribution form the most natural probabilistic approach to modelling Formula 1 race outcomes. Suppose, for the sake of simplicity, that a race consists of n cars whose finishing times T_1, T_2, …, T_n are independent exponential distributions with parameters λ_1, λ_2, …, λ_n. A standard result in probability theory (Grimmett and Stirzaker, 2020) states that the winning race time $\min\{T_1, T_2, \ldots, T_n\}$ has an exponential distribution, i.e.,

$$\min\{T_1, T_2, \ldots, T_n\} \sim \exp\left(\sum_{i=1}^{n} \lambda_i\right). \qquad (1)$$

Equation (1), as well as a method to calculate race-winning probabilities, are discussed in the proposition below. Further analytical results for this exponential model are discussed in Powell (2023).

Proposition 1

i. *If* T_1, T_2, ..., T_n *are independent and exponentially distributed with parameters* λ_1, λ_2, ..., λ_n *then*

$$\min\{T_1, T_2, \ldots, T_n\} \sim \exp\left(\sum_{i=1}^{n} \lambda_i\right).$$

ii. *If* X *and* Y *are independent exponential distributions with parameters* λ_X *and* λ_Y *then*

$$Pr(X \leq Y) = \frac{\lambda_X}{\lambda_X + \lambda_Y}. \tag{2}$$

iii. *Consider the Formula 1 race with independent and exponentially distributed finishing times as outlined above. Then the probability of the* j*-th car winning is*

$$Pr(Car\ j\ wins) = \frac{\lambda_j}{\sum_{i=1}^{n} \lambda_i}. \tag{3}$$

Proof of Proposition 1

i. $Pr(T_i \geq x) = e^{-\lambda_i x}$. Since all the T_i are independent

$$Pr(T_1 \geq x, \ldots, T_n \geq x) = e^{-\lambda_1 x} \ldots e^{-\lambda_n x}.$$

This gives

$$Pr(\min\{T_1, \ldots, T_n\} \leq x) = 1 - e^{-(\sum_{i=1}^{n} \lambda_i)x}.$$

ii.

$$\begin{aligned}
Pr(X \leq Y) &= \int_0^\infty \int_0^y \lambda_X \lambda_Y e^{-(\lambda_X s + \lambda_Y y)} dx dy \\
&= \int_0^\infty \lambda_Y e^{-\lambda_Y y} \left[-e^{-\lambda_X x}\right]_0^y \\
&= \int_0^\infty \lambda_Y e^{-\lambda_Y y} \left[1 - e^{-\lambda_X y}\right] dy \\
&= \int_0^\infty \lambda_Y e^{-\lambda_Y y} dy - \int_0^\infty \lambda_Y e^{-(\lambda_X + \lambda_Y) y} dy \\
&= 1 - \lambda_Y \left(\frac{1}{\lambda_X + \lambda_Y}\right) = \frac{\lambda_X}{\lambda_X + \lambda_Y}.
\end{aligned}$$

iii. For the sake of argument suppose $j = 1$. Then

$$Pr(\text{Car 1 wins}) = Pr(T_1 \leq \min\{T_2, T_3, \ldots, T_n\}).$$

Now T_1 and $\min\{T_2, T_3, \ldots, T_n\}$ are independently and exponentially distributed with parameters λ_1 and $\sum_{i\geq 2}\lambda_i$ respectively. Hence the result follows from (2).

□

The implication of Proposition 1 is that supposing we are given a sequence of win probabilities p_1, p_2, …, p_n, calculated e.g. from bookmakers' odds, we can estimate the parameters λ_i. To do this we can minimise the Residual Sum of Squares (RSS) or squared difference between the observed and model-predicted probabilities:

$$\text{RSS} := \sum_{i=1}^{n}\left(\frac{\lambda_i}{\lambda_1 + \ldots + \lambda_n} - p_i\right)^2. \tag{4}$$

The minimisation of the function in equation (4) can be done numerically – here using the function `optim` in R. The results of the procedure applied to bookmakers' data for a Formula 1 race for the 2023 season are shown in Table 1.

A few points about the construction of Table 1 are in order. Odds can be converted to probabilities as follows. The probability corresponding to odds of 25/1 for a Lewis Hamilton victory can be calculated via

$$\frac{1-p}{p} = 25;\ p = \frac{1}{26}.$$

Win probabilities for the remaining drivers are calculated in the same way, and then renormalised (Štrumbelj, 2014) so that they sum to 1. These renormalised win probabilities are then recorded in the fourth column of Table 1. The estimated $\hat{\lambda}$ values obtained from the minimisation of the function in equation (4) are recorded in the fifth column.

3 Statistical regression approach to modelling final ranking

Empirical Formula 1 data are most commonly listed in terms of the rank rather than the strict finishing times. The analysis of historical race data is therefore most easily accomplished by regression modelling of the final rank obtained. This follows a similar approach in Eichenberger and Stadelmann (2009). This implicitly assumes a Gaussian model – used for approximate predictions of sporting outcomes elsewhere (see e.g. Scarf et al., 2019).

This section therefore considers two related problems. Section 3.1 considers the calculation of approximate win probabilities given a regression model for the expected rank. Section 3.2 considers the related problem of reverse-engineering regression parameters given a sequence of win probabilities $p_1, p_2, \ldots, p_n$.

Team	**Car**	**Bookmakers odds**	**Implied win probability**	$\hat{\lambda}$
Mercedes	Lewis Hamilton	25/1	0.031655049	0.0081037902
Mercedes	George Russel	25/1	0.031655049	0.0081037902
Red Bull	Max Verstappen	2/9	0.673389233	0.1723897481
Red Bull	Sergio Perez	12/1	0.063310099	0.0162075831
Ferrari	Charles Leclerc	25/1	0.031655049	0.0081037902
Ferrari	Carlos Sainz	28/1	0.028380389	0.0072654675
Mclaren	Lando Norris	12/1	0.063310099	0.0162075831
Mclaren	Oscar Piastri	16/1	0.048413605	0.0123940343
Alpine	Estaban Ocon	500/1	0.001642777	0.0004205564
Alpine	Pierre Gasly	500/1	0.001642777	0.0004205564
Aston Martin	Fernando Alonso	80/1	0.01016088	0.0026012171
Aston Martin	Lance Stroll	500/1	0.001642777	0.0004205564
Haas	Kevin Magnussen	500/1	0.001642777	0.0004205564
Haas	Nico Hulkenburg	500/1	0.001642777	0.0004205564
Alfa Tauri	Yuki Tsunoda	500/1	0.001642777	0.0004205564
Alfa Tauri	Daniel Riccardo	500/1	0.001642777	0.0004205564
Alfa Romeo	Valterri Bottas	500/1	0.001642777	0.0004205564
Alfa Romeo	Zhou Guanyu	500/1	0.001642777	0.0004205564
Williams	Alex Albon	500/1	0.001642777	0.0004205564
Williams	Logan Sergant	500/1	0.001642777	0.0004205564

Table 1: Results of the model applied to betting data for the 2023 Qatar Grand Prix.

3.1 Calculation of win probabilities from a regression model

Suppose that there are n cars in the race and the final ranking r_i of car i can be approximated by a normal distribution: $r_i \sim N(\mu_i, \sigma_i^2)$. As an illustration if a classical normal linear regression model was fitted to ranking data this would correspond to $\sigma_i^2 = \sigma^2$ in the above with the values of μ_i and σ being determined by the fitted regression model (Fry and Burke, 2022). The approximate probability that car i wins the race is given by

$$p_i = Pr(r_i \leq 1.5) = \Phi\left(\frac{1.5 - \mu_i}{\sigma_i}\right), \tag{5}$$

where $\Phi(\cdot)$ denotes the standard normal CDF.

3.2 Estimating regression parameters from a sequence of win probabilities

Suppose we are given a sequence of win probabilities $p_1, p_2, \ldots, p_n$ for Cars $1, 2, \ldots, n$. Under the simplifying assumption of $\sigma_i^2 = \sigma^2$, equivalent to the classical normal linear regression model

(Fry and Burke, 2022), from equation (5) set

$$\Phi\left(\frac{1.5-\mu_i}{\sigma}\right)=p_i;\ \mu_i=1.5-\sigma\Phi^{-1}(p_i). \tag{6}$$

Since the sum of the ranks is equal to $\frac{n(n+1)}{2}$ summing equation (6) over i gives

$$\frac{n(n+1)}{2}=1.5n-\sigma\sum_{i=1}^{n}\Phi^{-1}(p_i);\ \sigma=\frac{n-\frac{n^2}{2}}{\sum_{i=1}^{n}\Phi^{-1}(p_i)}. \tag{7}$$

Combining equations (6-7) therefore gives the estimated μ_i values corresponding to the given win probabilities p_i. Table 2 applies this approach to estimate a set of $\hat{\mu}_i$ and $\hat{\sigma}^2$ regression parameters corresponding to the bookmakers' data shown in Table 1.

Team	**Car**	**Bookmakers odds**	**Implied win probability**	$\hat{\mu}_i$
Mercedes	Lewis Hamilton	25/1	0.031655049	8.704026
Mercedes	George Russel	25/1	0.031655049	8.704026
Red Bull	Max Verstappen	2/9	0.673389233	-0.242969
Red Bull	Sergio Perez	12/1	0.063310099	7.426002
Ferrari	Charles Leclerc	25/1	0.031655049	8.704026
Ferrari	Carlos Sainz	28/1	0.028380389	8.890783
Mclaren	Lando Norris	12/1	0.063310099	7.426002
Mclaren	Oscar Piastri	16/1	0.048413605	7.941444
Alpine	Estaban Ocon	500/1	0.001642777	12.904103
Alpine	Pierre Gasly	500/1	0.001642777	12.904103
Aston Martin	Fernando Alonso	80/1	0.01016088	10.501519
Aston Martin	Lance Stroll	500/1	0.001642777	12.904103
Haas	Kevin Magnussen	500/1	0.001642777	12.904103
Haas	Nico Hulkenburg	500/1	0.001642777	12.904103
Alfa Tauri	Yuki Tsunoda	500/1	0.001642777	12.904103
Alfa Tauri	Daniel Riccardo	500/1	0.001642777	12.904103
Alfa Romeo	Valterri Bottas	500/1	0.001642777	12.904103
Alfa Romeo	Zhou Guanyu	500/1	0.001642777	12.904103
Williams	Alex Albon	500/1	0.001642777	12.904103
Williams	Logan Sergant	500/1	0.001642777	12.904103

Table 2: Implied regression parameters corresponding to betting data for the 2023 Qatar Grand Prix ($\hat{\sigma}=3.879374$).

4 Regression modelling of historical results

In this section we calibrate the model to historical results (observed race rankings) from the last fully completed 2022 season. This follows a similar approach to modelling historical results in Fry et al. (2021). Following a similar approach in Eichengreen and Stadelmann (2009) we regress the finishing position against the dummy variables corresponding to each of the constructors. We then use stepwise regression (Fry and Burke, 2022) to automatically choose the best model. We constrain all models fitted to including a dummy variable indicating the teams' second (less-favoured) driver. Forwards and stepwise regression choose the same model indicated below in Table 3. In contrast, backward selection suggests a more complex model. However, an F-test, not reported, proved non-significant indicating that the simpler model in Table 3 should suffice. Negative and significant parameters in Table 3 indicate that the constructors have lower than expected final finishing performance. Results therefore lead to the following categorisation of teams loosely based on standard bond-rating terminology shown in Table 4.

Coefficient	**Estimate**	**Std. Error**	***t*-value**	***p*-value**
(Intercept)	13.8420	0.3794	36.484	0.000
Second driver	0.2160	0.4056	0.533	0.5946
Red Bull	-9.6500	0.7170	-13.459	0.000
Mercedes	-8.2700	0.7170	-11.534	0.000
Ferrari	-7.6900	0.7170	-10.725	0.000
Mclaren	-3.5500	0.7170	-4.951	0.000
Alpine	-3.5500	0.7170	-4.951	0.000
Aston Martin	-1.7900	0.7170	-2.496	0.0129

Table 3: Stepwise regression results obtained (constrained to include driver order term). R^2 value=0.3914.

Rating	**Teams**
AAA	1. Red Bull 2. Mercedes 3. Ferrari 4 eq. McLaren, Alpine 6. Aston Martin
AA+	Alfa Romeo, Alfa Tauri, Haas, Williams

Table 4: Suggested categorisation of teams based on stepwise regression results.

5 A regression approach to disentangling driver-car effects

Based on the regression output in Table 3 a 95% confidence interval for the second driver term is

$$\text{Second driver confidence interval} = (-0.581, 1.013). \quad (8)$$

This means that if we compare equation (8) with implied regression parameters in Table 2 a difference between two drivers of the same team bigger than 1.013 implies an extraordinary level of performance over-and-above the quality of the car. Comparing drivers in this way the suggestion is that two drivers Max Verstappen (Red Bull) and Fernando Alonso (Aston Martin) exhibit extraordinary performance levels over-and-above the quality of their respective cars. Past academic research has previously highlighted Verstappen's level of performance as historically significant (van Kesteren and Bergkamp, 2023).

6 Monte Carlo simulation of Formula 1 races and seasons

In this section we illustrate the duality of both approaches by showing how a regression model applied to historical data can be used to develop a simplified Monte Carlo simulation algorithm for historical competitions. A similar approach to modelling Rugby Union matches is reported in Fry et al. (2021).

Monte Carlo simulation proceeds as follows. Firstly, expected values are extracted from the regression output in Table 3. Using equation (5) a set of race-winning probabilities is estimated and then normalised so that the probabilities sum to 1 (Štrumbelj, 2014). Using the method in Section 2 a set of λ parameters are then estimated. The results obtained are summarised in Table 5.

Race outcomes are simulated by ranking a set of randomly sampled exponential finishing times according to the λ values listed in Table 5. Points are awarded according to the following: 1st place=25 points, 2nd place=18 points, 3rd=15 points, 4th=12 points, 5th=10 points, 6th=8 points, 7th=6 points, 8th=4 points, 9th=2 points, 10th=1 point. A bonus point is also available for drivers who secure the fastest lap and also finish in the top 10. For the purposes of the simulations it is assumed that the fastest-lap bonus point is awarded at random to one of the drivers finishing in the top 10. The whole procedure is then repeated 23 times to simulate a whole F1 season.

Monte Carlo simulation results for position and points are reported below in Tables 6-7. Results reflect the rigid segmentation into elite and non-elite teams as suggested in Table 4.

Team	Car	Regression estimate $\hat{\mu}_i$	Estimated win probability	Normalised win probability	$\hat{\lambda}$
Mercedes	Lewis Hamilton	5.572	0.184618051	0.138761898	0.113775773
Mercedes	George Russel	5.788	0.172192974	0.129423010	0.106118486
Red Bull	Max Verstappen	4.192	0.276388070	0.207737721	0.170331478
Red Bull	Sergio Perez	4.408	0.260685058	0.195935085	0.160654084
Ferrari	Charles Leclerc	6.152	0.152493089	0.114616260	0.093977910
Ferrari	Carlos Sainz	6.368	0.141539365	0.106383264	0.087227385
Mclaren	Lando Norris	10.292	0.026269000	0.019744203	0.016188970
Mclaren	Oscar Piastri	10.508	0.023498427	0.017661796	0.014481530
Alpine	Estaban Ocon	10.292	0.026269000	0.019744203	0.016188971
Alpine	Pierre Gasly	10.508	0.023498427	0.017661796	0.014481526
Aston Martin	Fernando Alonso	12.052	0.009988179	0.007507276	0.006155478
Aston Martin	Lance Stroll	12.268	0.008788244	0.006605386	0.005415992
Haas	Kevin Magnussen	13.8420	0.003249325	0.002442245	0.002002483
Haas	Nico Hulkenburg	14.058	0.002810319	0.002112281	0.001731932
Alfa Tauri	Yuki Tsunoda	13.8420	0.003249325	0.002442245	0.002002483
Alfa Tauri	Daniel Riccardo	14.058	0.002810319	0.002112281	0.001731932
Alfa Romeo	Valterri Bottas	13.8420	0.003249325	0.002442245	0.002002483
Alfa Romeo	Zhou Guanyu	14.058	0.002810319	0.002112281	0.001731932
Williams	Alex Albon	13.8420	0.003249325	0.002442245	0.002002483
Williams	Logan Sergant	14.058	0.002810319	0.002112281	0.001731932

Table 5: Regression estimates, implied win probabilities and estimated $\hat{\lambda}$ values for the 2022 F1 season.

7 Conclusions

Two natural ways of modelling Formula 1 races are a probabilistic approach based on the exponential distribution and statistical regression modelling of the ranks (Eichenberger and Stadelmann, 2009). Both approaches enable the race-winning probabilities to be exactly solved analytically. This tractability facilitates model calibration to either bookmakers' betting odds or historically-observed race rankings.

Equating race-winning probabilities means that both approaches can be seen as equivalent to each other. This time-rank duality is attractive theoretically and allows for a separation of driver-level and car-level effects. Results suggest that of the current crop of drivers Max Verstappen and Fernando Alonso out-perform the level of the car that they drive. Results match previous suggestions that Verstappen's performance level is historically significant (van Kesteren and Bergkamp, 2023). Time-rank duality also leads to a simplified Monte Carlo simulation algorithm for individual Formula 1 races and, by extension, entire Formula 1 seasons.

Future work will adjust the above models to account for cars that fail to finish races. There remains substantial interest in the analytical modelling of sports (Baker et al., 2022; Singh et

Team	Car	Expected position	95 % Confidence Interval
Mercedes	Lewis Hamilton	3.661	(1-6)
Mercedes	George Russel	4.014	(1-6)
Red Bull	Max Verstappen	1.776	(1-4)
Red Bull	Sergio Perez	2.005	(1-5)
Ferrari	Charles Leclerc	4.606	(2-6)
Ferrari	Carlos Sainz	4.939	(2-6)
Mclaren	Lando Norris	8.430	(7-11)
Mclaren	Oscar Piastri	8.832	(7-11)
Alpine	Estaban Ocon	8.426	(7-11)
Alpine	Pierre Gasly	8.831	(7-11)
Aston Martin	Fernando Alonso	11.956	(9-17)
Aston Martin	Lance Stroll	12.435	(9-17.5)
Haas	Kevin Magnussen	16.043	(11-20)
Haas	Nico Hulkenburg	16.482	(11.5-20)
Alfa Tauri	Yuki Tsunoda	16.041	(11-20)
Alfa Tauri	Daniel Riccardo	16.479	(11.5-20)
Alfa Romeo	Valterri Bottas	16.039	(11-20)
Alfa Romeo	Zhou Guanyu	16.484	(11.5-20)
Williams	Alex Albon	16.040	(11-20)
Williams	Logan Sergant	16.481	(11.5-20)

Table 6: Monte Carlo simulation of championship position (based on 1,000,000 simulations).

al., 2023). Financial aspects of professional sport (Plumley et al., 2021), including its ultimate financial sustainability (Richau et al., 2021), are also worthy of further investigation.

References

[1] Baker, R., Chadwick, S., Parma, R. and Scarf, P. "The binomial-match, outcome uncertainty, and the case of netball." *Journal of the Operational Research Society 73*, (2022): 1856-1872.

[2] Bell, A., Smith, J., Sabel, C. E. and Jones, K. "Formula for success: multivariate modelling of Formula One driver and constructor performance, 1950-2014." *Journal of Quantitative Analysis in Sports 12*, (2016): 99-112.

[3] Eichenberger, R. and Stadelmann, D. "Who is the best Formula 1 driver? An economic approach to evaluating talent." *Economic Analysis and Policy 39*, (2009): 398-409.

[4] Fry, J. M. and Burke, M. *Quantitative methods in finance using R.* (Open University Press, 2022).

Team	**Car**	**Expected points**	**95 % Confidence Interval**
Mercedes	Lewis Hamilton	299.142	(239-362)
Mercedes	George Russel	290.431	(230-354)
Red Bull	Max Verstappen	349.986	(291-410)
Red Bull	Sergio Perez	342.483	(283-403)
Ferrari	Charles Leclerc	275.384	(215-339)
Ferrari	Carlos Sainz	266.173	(206-330)
Mclaren	Lando Norris	91.606	(46-147)
Mclaren	Oscar Piastri	83.580	(40-137)
Alpine	Estaban Ocon	91.597	(47-146)
Alpine	Pierre Gasly	83.352	(40-137)
Aston Martin	Fernando Alonso	38.939	(10-81)
Aston Martin	Lance Stroll	34.550	(7-75)
Haas	Kevin Magnussen	13.210	(0-41)
Haas	Nico Hulkenburg	11.459	(0-38)
Alfa Tauri	Yuki Tsunoda	13.186	(0-41)
Alfa Tauri	Daniel Riccardo	11.436	(0-38)
Alfa Romeo	Valterri Bottas	13.193	(0-41)
Alfa Romeo	Zhou Guanyu	11.446	(0-38)
Williams	Alex Albon	13.207	(0-41)
Williams	Logan Sergant	11.461	(0-38)

Table 7: Monte Carlo simulation of championship points (based on 1,000,000 simulations).

[5] Fry, J., Smart, O., Serbera, J-P. and Klar, B. “A Variance Gamma model for Rugby Union matches.” *Journal of Quantitative Analysis in Sports 17*, (2021): 67-75.

[6] Grimmett, G. and Stirzaker, D. *Probability and random processes*, fourth edition. (Oxford University Press, 2020).

[7] Maher, M. J. “Modelling association football scores.” *Statistica Neerlandica 36*, (1982): 109-118.

[8] Maurya, A. “Formula one (f1) car: A scientometric study.” *International Journal of Advance Research and Innovative Ideas in Education 7*, (2021): 1463-1479.

[9] Plumley, D., Serbera, J-P. and Wilson, R. “Too big to fail? Accounting for predictions of financial distress in English professional clubs.” *Journal of Applied Accounting Research 22*, (2021): 93-113.

[10] Powell, B. “Generalizing the Elo rating system for multiplayer games and races: why endurance is better than speed.” *Journal of Quantitative Analysis in Sports*, (2023): (forthcoming).

[11] Richau, L., Follert, F., Frenger, M. and Emrich, E. “The sky is the limit?! Evaluating the existence of a speculative bubble in European football.” *Journal of Business Economics 91*, (2021): 765-796.

[12] Scarf, P., Parama, R. and McHale, I. “On outcome uncertainty and scoring rates in sport: The case of international Rugby Union.” *European Journal of Operational Research 273*, (2019): 721-730.

[13] Singh, A., Scarf, P., Baker, R. “A unified theory for bivariate scores in possessive ball-sports. The case of handball.” *European Journal of Operational Research 304*, (2023): 1099-1112.

[14] Štrumbelj, E. “On determining probability forecasts from betting odds.” *International Journal of Forecasting 30*, (2014): 934-943.

[15] van Kesteren, E-J. and Bergkamp, T. “Bayesian analysis of Formula One race results: disentangling driver skill and constructor advantage.” *Journal of Quantitative Analysis in Sports 19*, (2023): 273-293.