

THE JAMES WILLOUGHBY COLUMN



This month James looks at Jack Mitchell's remarkable record for Charlie Johnston. Comparing his strike rate for the yard with other jockeys who have had more rides motivates the use of a simple but powerful technique called Split Half Reliability.

IT MAY be a bit of a mouthful, but Split Half Reliability allows statisticians to turn data achieved in the past into predictions about the future. The relevant concept is the prior sample, a record of dummy wins and losses to add to a jockey's real statistics which puts all riders on an equal footing.

And we can apply this same framework to other racing statistics, such as a jockey's record at a course and a trainer's recent form.

For many years, I have highlighted the underrated prowess of jockey Jack Mitchell. The method I use to rate jockeys is known as analysis of variance and isolates a rider's true contribution to winning by controlling for the quality of their mounts. Every year, Mitchell comes out well, and only Oisin Murphy ranked above him in the most recent figures I published here.

Mitchell's talent was first recognised as a 16-year-old amateur in 2005, but it is fair to say that his career since has been something of a grind. Many times, he could be found at one of the sport's outposts while big meetings were taking place elsewhere. Finally, in the last few years, he has been used more regularly by the top yards, particularly the Newmarket stable of Roger Varian.

Notwithstanding Mitchell's 16 winners from 85 rides for Varian in 2023, including a Royal Ascot success on Royal Champion, his most remarkable numbers have been posted on Charlie Johnston-trained horses. For the year to September 28, Mitchell has won 12 races from just 33 mounts, a 36 percent strike rate. While this rate of success is not sustainable in the long term, it does suggest there may be what statisticians call an 'interaction' between Mitchell's approach and the needs of Johnston-trained horses. The question is: what might the Mitchell-Johnston combination achieve in the longer term?

Jockey	Trainer	Year	Wins1	Rides1	SR1	Wins2	Rides2	SR2
David Allan	Tim Easterby GB	2023	25	226	11	15	226	7
Oisin Orr	Richard Fahey GB	2023	23	163	14	25	163	15
Joe Fanning	Charlie Johnston GB	2023	21	140	15	24	141	17
Connor Beasley	Michael Dods GB	2023	9	133	7	13	135	10
Andrew Mullen	Iain Jardine GB	2023	7	124	6	12	124	10
Tom Marquand	William Haggas GB	2023	29	122	24	30	123	24
Paul Mulrennan	Jim Goldie GB	2023	19	114	17	17	115	15
Oisin Murphy	Andrew Balding GB	2023	22	114	19	21	115	18
Clifford Lee	K R Burke GB	2023	13	112	12	22	112	20
Duran Fentiman	Tim Easterby GB	2023	2	109	2	2	110	2

Table 1: The most common jockey-trainer records in Britain during 2023 from random split of the data

When this kind of question is presented to statisticians, a commonly used technique is Split Half Reliability. The idea is this: take every season of jockey-trainer records, such as Mitchell-Johnston 2023 or Frankie Dettori-John Gosden 2022, for instance, and randomly divide the data into halves. An example of some of these split halves for 2023 can be found in Table 1.

This is how to understand Table 1. Joe Fanning had 45 winners from 281 rides for Charlie Johnston for the year to September 28. A random split of this data resulted in one bucket containing 21 wins from 140 rides at a 15 percent strike rate and another 24 wins from 141 rides at a 17 percent strike rate. The point of this exercise is to examine the stability of the strike rate statistic – look down the two shaded columns for a visual guide. While Fanning's strike rate was similar in both buckets, Clifford Lee had two very different strike rates when his data for Karl Burke was randomly split.

A computer can perform billions of these random splits and compare the volatility of the two strike rates for a jockey-trainer combination within a season. Better still, the machine can learn what is known as a prior sample, a dummy record of wins and losses, to add to each half in order to make the best prediction about the other. The idea is that a large prior sample is required if the statistic in question is volatile, but only a small prior sample is needed if the statistic tends not to fluctuate.

The optimal prior sample for the average trainer turns out to be to add 5 wins and 43 rides to each jockey's record. This is a surprisingly small dummy. It suggests that a jockey-trainer strike rate tends to stabilise quickly.

The prior must be adjusted to match the trainer's overall strike

Jockey	Prior wins	Prior rides	Actual wins	Actual rides	Actual SR	Predicted SR
Joe Fanning	6	43	45	281	16.0	15.7
Franny Norton	6	43	23	185	12.4	12.7
Jason Hart	6	43	13	88	14.8	14.5
Jack Mitchell	6	43	12	33	36.4	23.7
Oliver Stammers	6	43	7	46	15.2	14.6
Andrea Atzeni	6	43	5	39	12.8	13.4
Richard Kingscote	6	43	4	42	9.5	11.8
Oisin Murphy	6	43	4	27	14.8	14.3
Rossa Ryan	6	43	3	8	37.5	17.6
Archie Young	6	43	3	22	13.6	13.8
Ben Curtis	6	43	3	29	10.3	12.5
William Buick	6	43	3	26	11.5	13.0
James Doyle	6	43	3	18	16.7	14.8
Andrew Mullen	6	43	3	36	8.3	11.4

Table 2: Predicted future strike rates for jockeys with three or more winners for Charlie Johnston

rate, however, as an average rider will have more wins for a high strike-rate yard than a low one.

Charlie Johnston's strike rate so far in 2023 is 14 percent, compared with the strike rate described by the prior sample of 11.6 percent. So, instead of adding 5 wins and 43 rides to each of his jockeys, we need to add $5 \times 14.0/11.6 = 6$ wins.

Table 2 contains the predicted future strike rates for jockeys who have ridden for Charlie Johnston in 2023, based on their in-season data.

IN Table 2, the jockey's actual strike rate, shown in the column second from the right, is regressed – moved towards – the yard's overall strike rate of 14 percent relative to the size of the sample. So, the predicted strike rate for Joe Fanning is not much different to his actual strike rate, since his sample size of 281 is large compared to the prior sample of 43 rides. By contrast, Rossa Ryan's actual strike rate of 37.5 percent is heavily regressed towards 14 percent because he has taken only 8 rides and there is less evidence to go on.

So it is that a profoundly important aspect of statistics is revealed: regression to the mean. Any statistic which is remarkably high or low tends to become less remarkably high or low as time goes on. And the smaller the sample size, the more regression tends to take place. In other words, it is much easier for a trainer or jockey to have a high strike rate while involved with few runners.

In between the 281 rides of Joe Fanning and the 8 of Rossa Ryan is the sample size of 33 rides which Jack Mitchell has taken. Bearing in

mind that the prior sample of 43 rides is relatively small, this is a decent amount of evidence to go on that Mitchell's future strike rate is likely to be higher than that of the yard in general.

Of course, all of this takes no account of the SP of future mounts. It assumes that, for each jockey, their opportunities will remain the same in future as in the actual sample. But this uncertainty is already accounted for in the prior.

In short, Mitchell's record is unusually good and point to him having either a natural approach which suits the needs of Charlie's horses or being able to understand this and adapt his style accordingly.

Other interesting prior samples

Remembering that the larger the prior sample, the more a statistic tends to reflect random factors than real ones, it

may be interesting to see how other racing statistics regress to the mean. Table 3 provides prior samples for commonly used metrics. Again, to find the predicted strike rate in future, add the competitor's actual record to the prior sample.

The higher up the table, the more reliable the observed SR tends to be. Here are some interesting takeaways regarding these strike rates:

A trainer's record with a jockey and at a course is more stable than a trainer's record in general. This means that trainers tend to have course-related, and jockey-related efficiency which is different from their overall performance. By contrast, a trainer's record in the last 14 days is a less reliable guide to their future performance – in most cases, so-called 'trainer form' is an illusion.

Similarly, jockeys tend to have course- and trainer-specific strike rates, and there is evidence for a 'hot-hand' effect – a jockey's record over the last 14 days is a better guide to their

next performance level than their strike rate overall. This is a long-held belief of the Johnstons. Not surprisingly, jockeys tend not to have going-specific capabilities.

In contrast to trainers and jockeys, a large amount of evidence is required in order to believe that a sire will have a better or worse strike rate than average. Interestingly, sires do have going-related capabilities which is difficult to tease out of their observed strike rates but is clear from Split Half Reliability. Sires also perform at different levels for different trainers, which is again not surprising considering that more powerful trainers have access to the best-bred, best-looking and most expensive of the sire's progeny. ■

Statistic	Prior wins	Prior runs
Jockey-Trainer SR	5	43
Trainer-Course SR	6	50
Trainer SR	7	59
Jockey-Course SR	7	61
Jockey last 14 days SR	7	63
Trainer-Going SR	9	79
Trainer-Sire SR	10	83
Trainer last 14 days SR	10	85
Jockey SR	10	90
Jockey-Going SR	26	222
Sire-Going SR	31	266
Sire SR	71	611

Table 3: Prior samples for some interesting racing statistics