
Singo, Gai and a Little Bit of Baye

Adjusting probabilities as the world unfolds in front of you is an important skill in life as much as investing. Fortunately, there's a little bit of probability theory to help us on our way.

There were all sorts of accusations being hurled at yesterday's inquiry into the performance of the race horse More Joyous at Randwick on 27 April, a spectacular flop as favourite in the Group 1 All Aged Stakes.

Owner John Singleton reckons he was told prior to the race – by someone who knew something – that there was something wrong with the horse. Worse than that, this someone turned out to be someone who had been told by someone who had been told by trainer Gai Waterhouse's son, bookmaker Tom Waterhouse.

It's a loose string of events. And there are hundreds of similar rumours at the race track every weekend. It turns out that one of middlemen – rugby league legend Andrew Johns – had had a few too many drinks on the Friday night in question. Singo himself had had a few beers before making his race-day accusations on national television. So everyone was drunk, it was all rumour an innuendo and it's time to move on, right?

Not so fast, Your Honour. There is an important piece of information here. The horse DID run second last in the race in question – the worst performance of its career. And that changes everything.

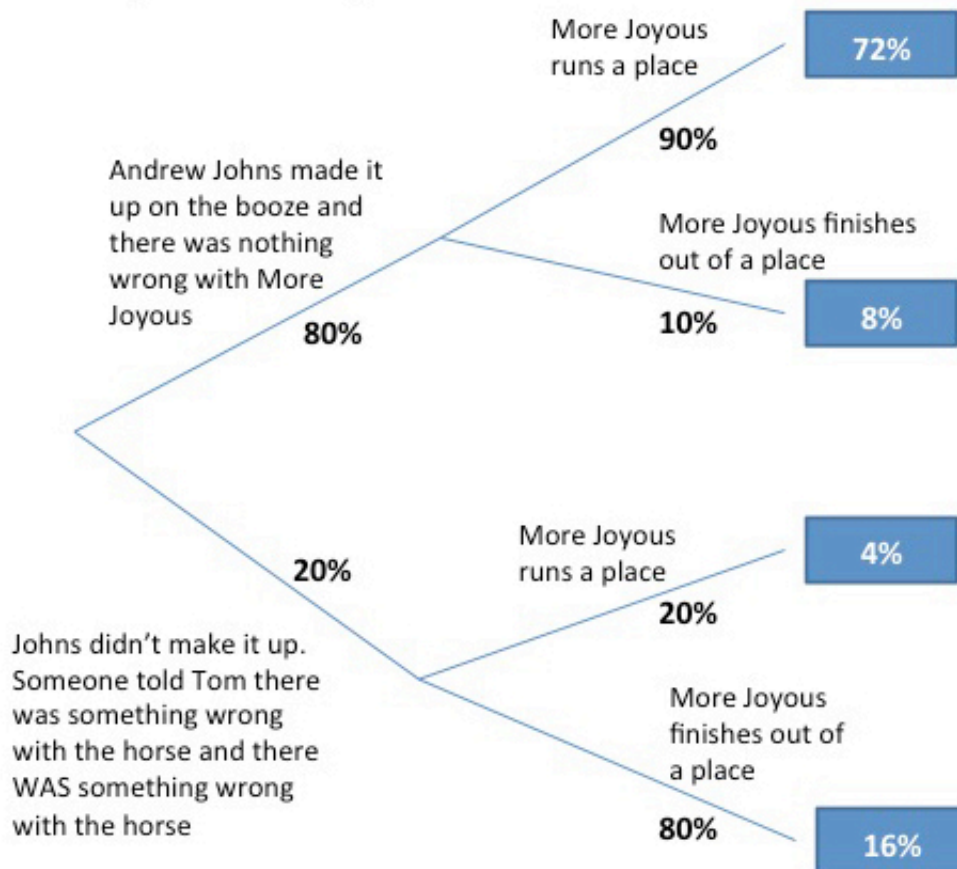
Nate Silver's book [The Signal and the Noise](#) is all about statistics and probability (it's a great read). His favourite topic is Bayes' Theorem, an area of probability that helps us update prior probabilities as the real world unfolds. And the Gai/Singo saga is a perfect example on how you might put it into practice.

First, let's lay down some assumptions.

Given the characters involved, the alcohol consumed and racing's penchant for rumours and innuendo, absent any other information you would probably say there's a pretty good probability that the whole saga is a beat up. Let's say the prior probability that there was nothing wrong with the horse was 80%, and that there was a 20% chance everything Singo is saying is on the money.

Next we need to need to assign some probabilities to More Joyous performing well under both scenarios. Prior to the controversial All Aged Stakes, More Joyous had raced 32 times and run a place on 24 of those occasions, a strike rate of 75%. The record is much better in Sydney and over the 1400m of the All Aged Stakes, but let's estimate that she runs a place 90% of the time when fit and well, and 20% of the time when she's not. You can see the prior probabilities mapped out in the diagram below. Adding them up, we have assumed there is a 76% chance she runs a place (72% fit and runs a place plus 4% not fit and runs a place). And there's a 24% chance she doesn't (16% unfit and runs poorly and 8% fit and runs poorly).

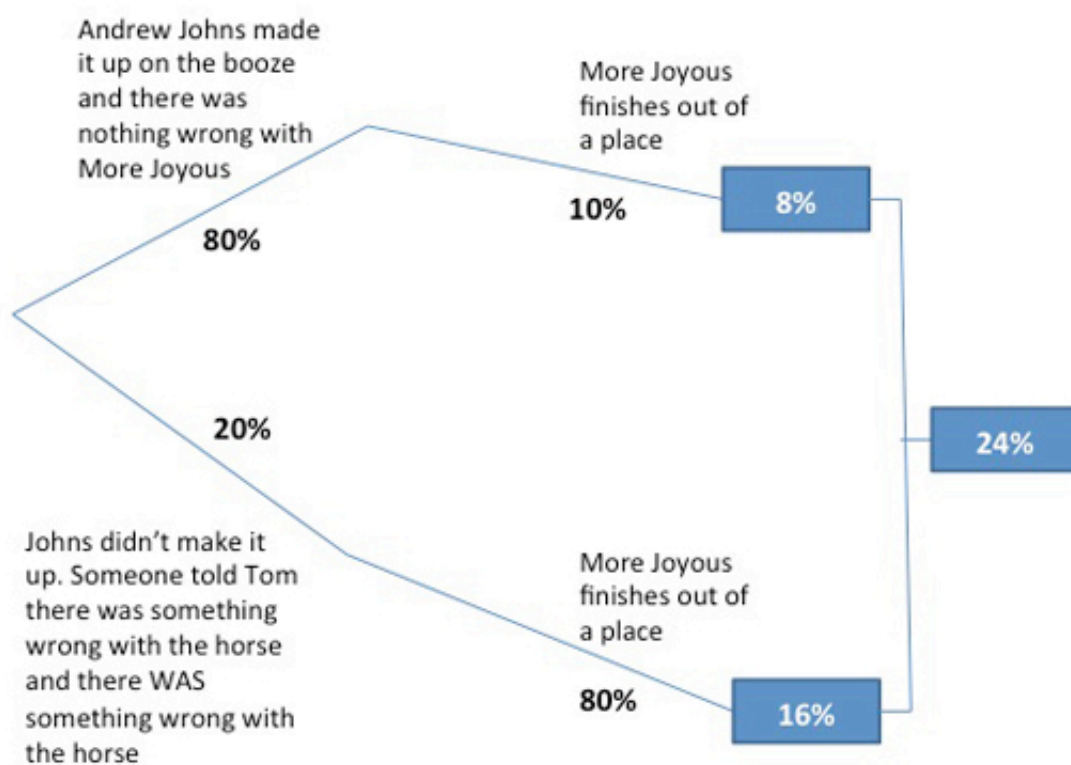
Prior probability tree



Now this is where Bayes' Theorem comes in. We've assigned a pretty low prior probability (20%) that Johns was privy to important information that Gai Waterhouse should have made available to the rest of the punters. But, **given that the horse ran poorly**, how should we update our probabilities?

There is a [complicated formula](#) for calculating this which I have never bothered to remember. I prefer diagrams, and all we need to do is update our probability tree for the fact that More Joyous ran second last. Like this.

Post probability tree



We're only left with the two scenarios where she runs out of a place, and all we need to do is compare them to each other to update our probabilities.

There is the 16% branch where there was something wrong and she ran poorly, and the 8% branch where there was nothing wrong and she still ran poorly. As you can see, it is twice as likely that we are in the bottom half of the tree – and that Singo's rumours were true – than the top half. Our updated probability, therefore, is that there is a 66% chance ($16\%/24\%$) that there was actually something wrong with her, given what we know about her performance.

That's a fairly dramatic change in probabilities, from 20% prior to 66% given the new information. In this example the probabilities have changed so much because More Joyous is a consistent racehorse and doesn't often run poorly. Other examples of Bayes' Theorem in practice show that the probabilities don't change as much as you might think (perhaps a future blog post if this hasn't confused you too much).

There are plenty of opportunities to apply Bayes Theorem in the real world. Especially relevant for us are its applications to updating your stock valuations for new information. And Singo certainly deserves a bit more leeway.