Correct Technique And Legality

Note the heel-and-toe contact during the widest spread of the stride, and the straight leg as the heel contacts the ground.
Welcome to this first lecture of 2012, in my maths and sports series, and today, we are going to talk about a collection of topics all related to being successful, by fair means or foul, and mathematical aspects of records and sequences of performance at the Olympics.

The title, if you are not a Latin scholar, “Citius, Altius, Fortius”, “Faster, Higher, Stronger”, is the Olympic motto. It has a curious origin. The founder of the modern games, Baron Coubertin, went to school in Paris, and a local school to the one he went to, not his own school, had this as their motto, and he rather liked it so he just snatched it from the neighbouring school and used it as his Olympic motto.

Well if you want to be successful in the Olympics, let us have a look at who has been the most successful individual in the history of the Olympics. So what I have done here was just to gather some data from reference books on the people who have won the most medals at the Summer Olympic Games, and I have distinguished and separated medals that are won in individual events from medals that are won in team events.

You can see, obviously, there are certain sorts of events, which, if you are desperate to win an Olympic medal, it is good to be in. So, if you are in swimming or you are in athletics or other team sports, there are many more medals on offer than there would be for an individual event alone.

So, what is shown in this table? Here are some names – I will identify some of them in a minute. And here is the pattern of gold, silver and bronze medals won in individual events, and here is the total number of individual medals. Some people think that you should not just add them up, but it is better to win a gold medal than to win silver and bronze, and so you will find, in many medal tables, some type of weighting or scoring is applied. So you might be given three points for each gold medal, two for each silver, and one for each bronze. If you apply that rule and calculate here, you have got nine threes plus one is 28. If you include team medals as well, you get a new total, sixteen for Michael Phelps, and if you weight the whole total, you get a new grand weighted total.

Who are the people on this list? Well, Phelps needs no introduction. He will be in London later this year, probably winning some more gold medals, an American swimmer. He is the most successful Olympic athlete of all time in terms of number of gold medals won – so fourteen, if you include the relays, and nine individual gold medals. Of course, lots of the events are really rather similar. If you win the 100 metres freestyle, you might well be expected to win the 200, and you might win the 50 as well.

Mr Ewry will not be known to you – he won his gold medals back in the 1900s, in the long jump, the high jump, and the triple jump, and those events were standing jumps in those days. So, he was only in individual events. He won the gold medal in every event that he competed in, all eight.

Caslavska is a great Czech gymnast from 1956, ‘60, ‘64 era, a specialist on the beam exercise, and she is the most successful in terms of gold medals.

If you keep going down, you will find somebody else who vies with Phelps for being the most successful Olympian, and that is Latynina, another Russian gymnast, who has a huge total of 31 individual medals and eighteen if you include the team events. So, again, she is in the ‘56, ‘64 period.

But if you introduce these weightings, you see it makes a difference. Although she has fourteen, four more medals than Phelps, if you total up and you introduce the weightings, she falls behind.
Here is the top athlete, Carl Lewis, individually, including the team.

More gymnasts, Andianov and Shakhlin, and there is another runner here, Nurmi, with six gold medals. He has managed to get some team medals because, back in his day, there were team events in the athletics, like cross-country events and so forth.

If you look for the first British entry in this list, if you keep looking down, keep turning the pages over, several times, you can guess who you first find, and it’s Steve Redgrave, but he is in 57th place on this list. He has got five golds and one bronze, and here are his totals.

So, this is what you are aiming at, up at the top.

Well, of course, winning certain sorts of medals, particularly team medals, it really helps to be in a big country, with a large population, and it helps even more if that big country has a big budget, a big GDP, and it spends a good deal of it on sport. By a long way, the most successful country ever in sport, at Summer and Winter Olympics, per head, was the old East Germany, so this was a state that really focused its activity, its international affairs, on being successful in sport, with a state-organised, almost compulsory system for people who turned out to be talented in sport. We know, retrospectively, and suspected at the time, that it was massively and systematically drug-aided.

So here are a group of pictures that show what happens if you were to plot the number of medals – this is the logarithm of the number of medals, so this is ten to the power four, ten to the power three, against some other measure of a country’s investment in sport. So this is a measure of what happens if you take into account the population, and you also take into account the financial investment. You can guess what ought to happen: there ought to be more medals if you have a bigger country – you have got more people to choose from; and you would expect to have more medals if you invest more money. But, if you combine the two, you find the best fit for recent Olympics – I think this is not the last one, it is the previous two or three – is to a sort of weighted combination of population and GDP. This straight line here is the best fit to the data for the major countries, and you can pick from it a number of things.

So, you can see the USA leads the all-time medal list, and then there is Russia, and then there is China and that is the UK here, but there is really a very big variance here. Most of the countries lie some way away from the best-fit line.

If we convert this best fit line to something a bit more recognisable, what the rule is, if you measure the population in millions, and the GDP in billions of US dollars, there’s this approximate power law. So, the number of medals is proportional to the population to about the 0.7 power, times the GDP to about almost 0.3.

If you pick one Games, let us pick Barcelona, and you play this Game with the weighted medals, so you give a score of three times the gold, plus twice the silver, plus the bronze, then you have a simple and rather sort of impressive formula for the score: the one-third power, to quite high precision, of the population, times the two-thirds power of the GDP.

Now, if you just mess around with this a little bit, you can see, if we take the cube root outside of PG squared, and if you just defined another S, let’s call it S’, to be a half of S cubed, so that we can cube both sides, then S primed is a half PG squared, and that is a formula that, if you have done any mechanics, is rather reminiscent. It is rather like the kinetic energy – half the mass times the speed squared is the kinetic energy of motion of a body of mass M at speed V. So this is one of these econo-physics analogues that people talk about in mathematical economics, that half PG squared is like the energy of a nation, the economic and financial energy. You can go on and play other Games with these formula, but you can see, roughly, the number of medals won is linked to this energy of the nation.

Well, you might think it is a rather bad thing that medal success is so strongly a function of GDP, of money, and if you were a journalist producing medal tables, you might want to therefore regard the success of a country in
the Olympic Games medal-winning business as being the extent to which they are above this standard curve. So this standard curve is what you expect simply from your investment and your population, and your degree of success is the extent to which you are above this line, and your degree of failure will be the extent to which you are below.

Well, this issue of money being a big distorter of performance is something that we recognise in this country, particularly in football. So, the Premier League is not really very competitive after you get below the first three or maybe four teams. So, a team like Liverpool at the moment is closer to being relegated than it is to being at the top of the table, and one of the reasons for this is it is a sort of Matthew effect – you know, “Unto he who has shall more be given”, that if you have lots of money, you buy more players, you get more revenue, you buy even more players, and there is a runaway instability.

It is amusing to look at what happens in another country, where they try to directly counter that effect in sport. In the USA, there is a system known as the draft, and it is used for American football, and it is also used in basketball and other sports. What it amounts to is that, at beginning of each season, when new players become available to be hired by the top teams, then there is a selection of those players. They put themselves forward from college to be offered contracts by the top teams. The US system is that the team that came bottom of the league in the previous season gets first pick, so the worst team gets to pick the best player, and so on down the ladder, and eventually, the top team that won the cup final the previous year gets the last pick.

This selection process is fairly easy to model, with a simple type of delay differential equation as it is called. So, your rate of success this year is proportional not to your state of play at the moment but by the state of things as they were at some time in the past, and what that does is to produce a performance prediction that’s [acceleratory] – it goes up and down, like a sinusoidal wave. When you look at the results of American football league performance, if you give people sort of a point for winning and half a point for drawing and none for losing, you will find that, over a period of eight years, four times the delay time in setting up the draft system, teams do indeed behave in this sinusoidal way. So, they improve for a period of years, this puts them at the bottom of the draft, and then they go down for a bit, and then they come up.

So, in American sporting politics, this is something, presumably that is agreed by all the participants – they wish to do it. In Europe, it would be contrary of course to employment law to do that. Again, it would have to be something that people agreed to take part in. But it is an interesting example of how you can have a rule system that actively attempts to prevent rich teams getting more and more of an advantage.

In the American soccer league, for example, there is another factor rather like this: all teams are required to travel in economy class, there cannot be any private aircraft to convey team members to matches, so if you are a very wealthy club, you are not allowed to be being ferried around by private jet while your opponents are checking into the American equivalent of Sleazy-Jet or whatever.

Well, suppose you do want to invest in some way, you want to spend your fraction of GDP, what should you do if you want Olympic success? If you look back to what happened with China, in particular, in the run-up to Beijing that they were hosting, there was an enormous increase in investment, and, to some extent, there has been in this country before the last Games and the current ones. What should you do if you want to get more medals for your buck, as it were?

Well, the first thing is that you should pick sports that not many countries do. So, if you pick cycling, you have got a much better bet of winning something than football or the 100 metres in track athletics.

You should maybe also pick events with lots of relays, lots of team awards, maybe things where there are double-bronzes. So, in boxing, for example, once you lose a bout, you cannot fight another one, so the two losing semi-finalists both get bronze medals. So there are more medals on offer in those sorts of fighting events than in other sports.
If you pick sports where there is a lot of inter-event similarity, you are going to get more for your money. You can see this effect particularly in sports that Britain is good at, like sailing, and especially cycling. It is not a fair comparison to compare the performance of the cycling team, given a certain amount of investment, with that of the athletics team. Athletics is like many, many completely different sports, so what you do to help the pole-vaulters will not help the marathon runners or the hurdlers or the middle-distance runners. But in cycling, there is a very big shared element, so if you invest in aerodynamically-superior bikes, better kit, better velodrome training facilities, it will benefit everyone in your squad.

The other lesson perhaps we learn is that there are certain sports where sheer training and organisation and discipline can win you Olympic medals, over pure talent. So, if you looked at the last Olympics, China invested in all sorts of things, like sort of women’s weightlifting, various types of team event, where systematic squad training, improving pure strength, something that can be done very systematically, really pays off. Whereas, if you decide you are going to invest in trying to have the men’s 100m sprint champion, there is such a big talent factor required that, just by searching people out and training them systematically, you are probably going to fail.

A lesson we have not really learnt in the UK – and it is not necessarily a good lesson to learn – is that do not try and do too many sports. So, if you go to Kenya or to Ethiopia, anyone who is fit and has a very good cardiovascular system will be doing distance running, so there is a great specialisation of fit athletes into a small number of sports. Here, all the throwers, the shot-putters, the hammer-throwers, the discus-throwers, that we do not have are playing rugby! They are playing rugby league, they are playing rugby union...they are in a host of other sports. So, you dilute your pool of talent for any particular sport if you do a very, very wide range.

All the American long jumpers, high jumpers, strength event people, are enormously reduced by the enrolment of very athletic, very powerful athletes into American football, and to some extent also into basketball.

The other thing that you could do to try to counter this is that you have got to encourage some people to change events. So, if you have lots of 400 metre runners, all thinking that they are going to get the last spot to make the team at 400 metres, some of them need to be encouraged, well beforehand, to think about competing at 800 metres.

More radical, you might try to encourage people to change sports, or to join a different sort of sport. And, the sort of sport that you might join, and the factor that you might think about, is something like triathlon. See, the triathlon is massively biased, and so a certain sort of sportsperson should consider seriously whether they should take up the triathlon, and certain other sorts of sportsman should perhaps consider whether they should give it up.

Here are the results from the last Olympics. If you are not familiar with the triathlon, it is a sequence of three events: so you first of all swim, and you then cycle, and you then run. You have a transition, so you come out of the water, you put your cycle shoes on, and you get on your bike, with your horrible little withered toes from swimming 800 metres in the sea, and you then ride for an hour or so, and then you jump off your bike and you run 10,000 metres. So, the times are just added up.

The distances that are spent on each of these sections are an historical accident. The event derives from something called the Iron Man event, which was first competed for in Hawaii I think, back in the early 1970s. It was the brain-child of the San Diego Track Club. Although that event involved much, much longer swims and runs and cycle rides, the proportion of time spent on the run, the ride, and the swim have remained the same, and even in shorter, sort of mini-triathlons, the same proportion is adopted.

But if you look at the results at the last Olympic Games, you see that it is not really very cleverly chosen. So, here are the medallists at the last Olympic Games in the men’s, and here are the women. So, you were swimming for eighteen minutes, and then you are cycling for 59 minutes, and then you are running for 30
minutes – your overall time is about an hour and 48. These were the fastest legs on each of those. For women, the times are sort of a minute and a half slower on the swim, and about five minutes slower on the bike ride, and about sort of three minutes and a bit on the run. But you can see that there is about 54% of the time is spent cycling, sixteen percent swimming, and 28% running. There is a massive bias in this event to cyclists. It is really just a cycle ride with a warm-up run and a shower.

So, if you are thinking of changing events to the triathlon and you are a swimmer, don’t bother! But if you are a cyclist, and you can swim a bit, and you can run a bit - everybody can get better at running, not necessarily at swimming because you might have a bad technique – but you can see that a certain type of transformation, change from one event to this event, could be very advantageous.

Well, because of this, I suggest a revision of this event. A logical, equi-tempered triathlon would be one where competitors spent, essentially, an equal amount of time on each event. To do that, let’s take the one hour 48 as some type of standard, then what would you want to do? A one and a half kilometre swim, okay, probably just a 40 kilometre bike ride, and a ten kilometre run – that is what you do at the moment. But to get those time stages roughly equal, around 36 minutes, you’re looking at a three kilometre swim - athletes will not like that, a 24 kilometre bike ride, and a twelve kilometre run. So that is what the triathlon should be like, and if it was, there would be an equal appeal to swimmers, riders and runners, and you would end up with a rather different type of competitor. It would be attractive to a different type of multi-event sportsperson. Well, you heard it here first!

So, this event is interesting because the times are added directly – you just keep going. It is like a relay with one person in it. But there are other multi-event sports where you convert performances into points, and one is the decathlon.

I came across this on the web, this rather amusing poster, for the 100 [litres], the front crawl, asymmetric bars, rowing, weightlifting, the individual pursuit, the shot, and the three-day event.

But the real decathlon, a formidable event... It did not – I think it existed in the first Olympic Games, and then it was not in the second one, but then, after that, it resumed. It involves, as the name suggests, for men, it is ten events, and for women, it is the heptathlon – Jessica Ennis being one of the prime competitors in that event.

So, what do you do? Well, you have got a mixture of running, throwing and jumping. So, you have got to sprint 100 metres, you have got to hurdle 110 metres, sprint 400 metres, and then stagger round 1500 metres at the end. You have got to throw the shot, the discus and the javelin, and you have got to high jump, long jump, and pole vault. The women have 200 and 800 metres – I have put that in twice, for some reason – 100 metre hurdles, high jump, javelin and shot-put.

The interesting thing about this event is that you do not just add the performances, but the performances are converted into points. There is clearly some arbitrariness, there is some mathematics really involved in that...and here is how the scoring works in this event...

You have got two sorts of goals: for the events where the outcome is a time, you want that time to be as short as possible; where the outcome is a distance, you want it to be as long as possible.

So, in the field events, where you want the distance achieved in a jump or a throw to be as long as possible, the mathematical formula that is used to convert performances to points looks like this. So it is some quantity, let us label it A, times the distance, minus another quantity, all to the power some number which we will call C. So, if your distance is less than B, then you are not going to get any points at all, so that is the sort of rock-bottom.

On the track, where you should get more points for having a smaller time, the formula has the same type of
pattern: some number, times another number, minus the time to the power C.

For each event, the International Athletics Federation, from time to time, decides what these quantities should be, A, B and C – there is one for every event. They are arrived at by looking at all performances, historically and currently, in that event, the world record in that event, and also the performances in decathlons, which are sort of biased in a slightly different way, as we will see.

So, for example, the As and Bs are not terribly fascinating, they are just numbers that come from the averages of the performances. What is more interesting is this number up here, C, which sits in the power. If this number was just equal to one, then the points system would be rather neutral, in that, if you were better, you ran faster, you threw further, you would find it just as hard to win the next 100 points as you had found to win the previous 100 points. But if C is bigger than one, then you are going to find it harder to win the next 100 points to be better than you did to win the previous one, so you get more reward, as it were, for the better performances. If C is less than one, then you get comparatively less reward for the better performances.

Well, if you look at the tables, there is a pattern to what goes on. If you look at the running events, the statistics have ended up giving values of C for the running events, here, here, here and here, they are really quite similar value of C – they are about 1.8, quite close to two. If you look at the jumps, which I have put in orange here, long jump, high jump, pole vault, which is a little bit different, again, you can see they are quite close to 1.4, so there is another cluster for the jumps. Then, if you look at the throws, this is quite close to being neutral – they are very close to one. Now, the interesting thing about these events is that, clearly, you can see that, if we tinker with these tables, if we change these numbers, we are changing the reward system for different sorts of points. So if we were to change these throw values of C and increase them up to 1.8, we would be giving massive rewards for the better throwers, and the results of the whole competition would be different, very different.

In fact, when Daley Thompson won the Olympic Games in Los Angeles, he just missed beating the world record by a few points, but a few months’ later, the tables were recalibrated, and he was suddenly told that he was now the world record-holder because his performance in Los Angeles had given him an increased score on the new tables and he had beaten the world record by a few points.

So, this is an event with a serious type of subjectivity. If you look at the women’s heptathlon, the pattern is really very similar. So, for the values of C, for the running events, around 1.8, for the jumps, around 1.35-1.4, and for the throws, they are quite close to one. So, it is the same for male and female.

So let us have a look what sort of things could you now examine? Well, the world record in the decathlon for the ten events is 9026 points, a fabulous score by Roman Sebrle from the Czech Republic, and the number two all-time performance is not many points behind, by the Czech number two.

Well, if you wanted to score 9000 points, which I can guarantee will win you the gold medal in London, what would you have to do? If you wanted to score 900 points in every event, this is what would do it. You can see, some of the performances, I mean, they are not by any means out of this world – 10.83 for 100 metres, 48.19 400 metres, under eighteen, under nineteen scores - top-class competitors run faster than this, much faster than this. There is a fifteen year old who runs as fast as this in England. 1500 metres, four minutes seven seconds...do that bouncing around on your head – I mean, people, a fifteen year old can run much faster than that. But in the throws and the jumps, the performances become rather more impressive. So, for a good class competitor, competing at a high level in national or even county championships, these individual performances are not really out of reach, but it is being able to put them altogether, over two days, five events per day, that is clearly very, very challenging.

Over here is a picture I have created of where did the 100 best performances in the decathlon, ever, how did they get their points, so how were they spread over these events? That should say 400, not 4500.
So, what you can see from this is then what the top decathletes tend to be like. They are not 1500 metre runners, and they do not bother much preparing for 1500 metre running. They are mostly enormously heavy, strong athletes. It is quite a challenge to run around 1500 metres, although Sebrle runs four minutes twenty seconds or something, very, very fast. But you can see, long jump, hurdles, sprints, 400 metres, they tend to be sprint/jumper type athletes.

Also, you might say from this picture, if you have got a limited amount of training time, where should you invest it? Where will you get more points per hour of training? You can see that, in these jumps, sprints, hurdles, you will get much more in return than spending all your time pole-vaulting or cross-country running for the 1500 metres, because sprint training will help you in the long jump, it will help you in the 100, it will help you in the 400, it will help you in the hurdles. So it is a particular sort of athlete who is going to be successful in the decathlon, in the way that the points scoring structure is set up at the moment.

Well, I thought I would try to invent a new points scoring scheme, which we will call the Barrow Score, or the B-Total. It is an attempt to see what happens if you avoid all this points transformation subjectivity. In the decathlon, we have got six events - the jumps and the throws - where you want to have as large a distance as possible, and you have four running events where you want to have the smallest time possible. So, why don't we just multiply the distances achieved in the jumps and the throws together, convert them all to metres, so they are in the same units, and divide by the times, all in seconds? So, you will have answer that has got units of metres to the power six, divided by seconds to the power four. But it is a very simple system. It is very direct. It is not without biases of course, and, at your leisure, you can think about what happens. I will not put all the arithmetic in. So, you know, we are just multiplying the distances together, dividing by the times.

If we look at the world record-holder, Sebrle, he seems to come out not in first place anymore, but second, and the number two in the world, Dvorak, who scored 8994 points under the old system, he is the top-ranked athlete under this new scheme, with 2.40, and Sebrle at 2.29. But it is interesting that one captures sort of exactly the same two people. You end up with a number that is neither sort of ten followed by 50 zeros or vastly smaller than one. It is a perfectly sensible set of numbers to be dealing with, and it is a much more direct subjectivity-free way of converting performances into points.

And you can think about what sort of performance here would boost this score the most, so where is it easiest to shorten a time or increase a distance. They are all equally weighted. Well, obviously, the longer you are running for, the more chance that you have got to decrease that time. The event that has got the longest distance that you are throwing has got the most scope for increasing that distance.

Well, before we leave the decathlon, there are some funny facts and observations about it that are quite interesting. One case that people have made is that you should have C equal to two for all the events, because this is rewarding people in proportion directly to the amount of that kinetic energy - you remember I mentioned previously, half MV squared, so the amount of work that they have to do in each event, energetically. And if you do that, it of course does change the scores, a lot, and again, Mr Dvorak becomes the new world record-holder, with a very large 9468 score. This change really helps the throwers, because, at the moment, their value of C is 1.1, but it will go all the way up to two.

The point I mentioned here, let us get some sort of normalisation here. The present world record is 9026. Suppose you set the current world record, in every event, okay, what would you get? You would get 12500 points. You can see, therefore, the world records must have played some part in standardising things because it is a nice round number. If you took the best performance that had ever been achieved in a decathlon, then you are up at 10485.

As for individual events, if you try to convert Bolt's 100m world record, he gets 1202 points for that. The fastest ever 100 metres in a decathlon, on the other hand, is much lower, 1042. The best world record of all on these tables is the discus world record, from long ago, and it scores 1383.
Well, I want now to have a look at records, in the simplest way really, and, first of all, to show you that a simple statistical analysis can tell you whether records are being set randomly or not, whether there is something else going on.

Mathematicians have a simple set of arguments for analysing how often you expect a record to be set in a sequence of events that are random. So, this might be the rainfall, high tides, or something like that, that you had some suspicion was random.

So, suppose, in the first year, or the first period where activities took place, you are doing the event for the first time, so it must be a record. So, after one year, the number of record years, record seasons, is obviously one.

In the second year, things are independent of the first year, you have got a one in two chance of beating the record that you set in the first year, and a one in two chance of not beating it. So, after the first two years, the expected number of record years is just one plus a half.

If you go to year three, then the performances that you could have from the first two years could have six possible rankings. So, the performance from year one could be better than year two, could be better than year three, or year one could be best and year three second and year two third, so these are all the permutations of one, two and three. And you can see that there are just two of the rankings here which produce a record in year three, and that is this one and this one. So, this one produces a record in year two, this is a record in year one. So, there are two of those, six possibilities, two over six is a third, so the chance of a record in year three is one in three, a third, plus a half, plus one.

If you keep on going, same sort of argument, after four years, the extra factor is a quarter, and so after N years, the expected number of record years is just going to be the series one plus a half plus a third plus a quarter, and so on, plus one over N. So this is the famous harmonic series that we have seen in these lectures before, and as we know, if we keep on going forever, this series sum can be larger than any number that you specify, but it grows very, very slowly. So, if we call the sum after N terms, H(n), then after one term, it is one, after two, it is one and a half and so on, and after four, it's 2.083, but even if you add in 100 terms to this series, the sum has only grown to 5.19. If you go to 256, it's 6.124, a thousand, 7.49, and even a million terms, the sum is only 14.39. So, roughly, a good approximation, the sum is quite close to 0.6, plus the logarithm, natural logarithm, of the number of terms.

So, from this, you can begin to see how many records you might expect to have. So, if you were looking over 100 years, since 1900 or so, and you took a check, a rain check, each season in athletics, you would find, you would expect there have to have been five records in each event, five world records in the last 100 years.

If you look at the number of Olympiads, there is about 26 of them, Summer Games, since they began, you would expect only to have four records set in any event, if record-setting was just a random process, rather like the rainfall.

Of course, it is not because here is just one example, here is the men's 100 metres, and here are all the records, the times there has been a world record, since about 1910 right up to the present. I have not counted them all, but there is vastly more than five. So the setting of sporting records is not a random process. It is driven by improvements in training, improvements in facilities, in track surface, in time keeping even, starting technique and things like that.

However, there is something rather odd about world records in athletics. And, first approximation, there are not any women's records anymore, certainly not in events that have been competed for the last 30 or 40 years. So, this is a great problem for sponsors and meeting organisers, people who want publicity and excitement - you are never going to get a world record in the 100 metres, the 200 metres, the 400 metres, the throws, and so on.
Here is all the standard events on the Olympic programme. Here is the date of the last world record, in the men's event, and here is the date of the last women's record. I have highlighted some events that seem to have had recent world records, but which do not really count because they are events that were only invented a few years ago for women, so these events were not taking place back in the 1980s and '90s.

So, women's world records stand for a very, very long time, usually, completely unapproachable. So nobody gets anywhere near, say, the women's 400 metre record, which is down around 47.4 seconds, whereas if you ran 49.4, you would be the favourite for the gold medal at the Olympic Games. So what is going on here?

Well, if we just take the numbers from the previous picture, there are eleven women's world records that are more than twenty years old, compared with only two men's records. The problem is that you are looking back to an era when drug taking was systematic, particularly by East German athletes. We know from the records, which have become public after the fall of the East German state, that there was systematic doping and performance enhancement, and those records really are a result of that regime, which can no longer continue because drug-testing became much more stringent and these very old records all tend to pre-date the time when drug-testing for male growth hormone and other drugs that favour improvements in women's performance became sort of essentially impossible really to use without detection.

Well, drug testing is an interesting little mathematical problem. It tests one's sort of logic and understanding of rather simple statistics, and the principles at work there are really the same ones that need to be at work in the court system, if you are trying to determine whether a piece of evidence, say DNA evidence or fingerprint evidence, from an individual is really a decisive piece of evidence for their conviction.

So let us just do a simple example. Let us suppose that one percent of athletes take performance-enhancing drugs - so I am just picking these numbers so that they make the arithmetic come out reasonably clearly – and 80% of drug tests which are performed detect the presence of drugs, but 9.6% of those tests give you a positive result when no drugs are present in the athlete, so these are false-positives. So, what is the situation? We could make a little table here, that one percent of athletes take drugs, 99% do not, in your sample, and your test will find 80% of those to test positive, therefore twenty percent will test negative, 9.6% of those tests give a false-positive, so 9.5% of this sample end up giving a false-positive, so very large. Down here, if you test negative, well, there is one percent of people who took drugs, and the test fails to identify because twenty percent of the tests are bad, so 0.2 of a percent, 0.002, give a false negative, so these are people who take drugs but the test does not find them.

Then, over here, we have got true negatives, so these are people who are not taking drugs and the test correctly identifies that. There are 89% of those.

So, what is the chance that a drug-tester tests positive? Well, it is the 0.008, the true positive, divided by all the...
positives, so 0.008 plus the false-positives, so the chance that a drug user tests positive is 0.0776. That is not very high...7.76%, even though the test is 80% correct... So the reason this number is so small is because of this very large number of false-positives, 9.5%, so bigger than the number of positively tested drug users. So, this is the type of statistical analysis that you have to make in order to determine how effective a particular test is at finding true drug takers.

And you can see that you can play around with these numbers. You perhaps do not know what these numbers are at all, because if people are not being detected, they are successfully evading drug detection, then you just do not know about them, so you are just biased to the people who you do detect. And then you can try and work harder and do better at making these numbers better and this number bigger, but the basic logic of what you have to calculate is essentially the same.

It is an example, if you know any statistics, of course, of the Reverend Thomas Bayes’ famous theorem from the 1700s. So, Bayes did not publish this during his lifetime – it was found by one of his colleagues and published after his death. So, he considered the general problem of, if you want to know the probability of some outcome, given some other condition, how does it depend on the probability of that condition given the outcome times the probabilities of the two possibilities.

So, here, what we have just worked out is a special case of this. We want to know what is the probability that you have been taking drugs, A, given that you test positive, B. Well, it is the probability that you test positive, given that you have been taking drugs, times the probability of taking drugs over the probability of testing positive.

And in the court system, where this would be applied to DNA evidence or something else, the prosecutor will try to fool you, if you are a juror, into thinking that PA given B is the same as P given A. It is even known as the Prosecutors’ Fallacy.

So, if we just went through the arithmetic and we put all our numbers in, P(A) as one percent. This is the effectiveness of the test, 80%, down here. This is the false-positive probability, PB not A. And this is the probability you are not taking drugs. Put all the numbers in...we get the answer that we just got by our slower application of pure thought, if you like. But this is the general formula that Bayes found, a very, very important formula of statistics, which can deal with any problems of this variety.

Well, the last thing I want to talk about is a situation where an unfair advantage can be gained, sometimes not deliberately, as in the case of taking drugs, but where there is an interesting little piece of mechanics involved, and this is race-walking.

So, these are the sorts of people walking along in a rather unusual style. Here is a picture of the 1908 Olympics - this is White City. This is the start of the men’s 3500 metre walk. Almost everybody seems to be lifting at the start, but I think they did not worry about that. Well, race-walking is not like walking down the street. Race-walking looks different, you see people waddling along, swinging their hips. How does walking differ from running? In everyday life, the point about running is that your centre of gravity bobs up and down, so it does not move in a straight line, it moves in a sort of sinusoidal line. When you walk, your centre of gravity moves pretty much in a horizontal line – it does not bob up and down.

If you are a race-walker, there is a more stringent definition of walking that you have to adhere to. Here is an official statement of it. It looks as though it was amended by the bracketed condition at some stage. What has to happen is that you have to maintain contact with the ground with one foot at all times. So, your next step, your foot must not come off the ground there at the back until the front foot has touched the ground, so you must have continuous contact with the ground. Whereas, running of course, you are jumping up in the air, you are not in continuous contact. The other requirement is that your back leg must be straight, so your joint at your heel must lock out. So you cannot walk with a bent leg in the way that you do when you run -
so this runner has always got a bent leg. So this is
the rule, a progression of steps so the walker makes contact with the ground, and the second rule is that the
back leg must be rigid. The point about this rule is that it is visible to the human eye. Nowadays, where people
are scrutinised by television cameras and even private videos, this is a very difficult criteria to adhere to. The
human eye is like looking at something with a camera that just takes a shot 25 frames per second – it is a
really rather slow camera. Whereas, the video cameras and TV cameras that are looking at walkers are taking
many, many more frames per second than that, so they see things that you do not see, you cannot pick up with
the naked eye.

So here is an international race-walker. He is clearly lifting. The rules are, by the way, on these events, twenty
kilometres and 50 kilometres, on the road, you are allowed two warnings, so you get two warning cards shown
to you by judges round the course, for lifting or not straightening your leg, and the third one you get, you are
out, you are disqualified. There are cases at major championships of athletes being disqualified just a few metres
from the finish in the stadium. You have to get the warnings from different judges, so it cannot be the
same person that is always giving you a warning of the same thing. So, this fellow here, you see he is clearly
breaking contact. This film is rather slowed down. His leg is straight, but he is breaking contact.

The next video, this is the star man. He is no longer active. So, he was the world record-holder, Olympic
champion, and you can see his style is absolutely perfect, and, certainly to the naked eye, he is not lifting. The
way you get these fantastic
speeds is to have extremely high cadence rate. So, his step rate is very similar to that that you would have in a
400 m sprinter, and this is what requires tremendous leg strength. Now, let us just look at the mechanics of
the forces involved to see if we can work out a criterion for when you lift. I have got two pictures to do this.

So, if we imagine that this is the centre of our walker, here is his weight going downwards, and this is where the
back foot is. This is the stride length. This is the leg length. That back foot then has a reaction, a force going
upwards from the ground, as you change this angle. If that reaction force were to go to zero, then the
athlete must be lifting. If it is positive, then he must be lifting. So, by working out the equation of motion in the
vertical direction, so R upwards minus Mg is equal to the acceleration in the vertical direction, and also looking at
the rotation around this point, so the acceleration of this angle times the moment of inertia of the body is equal
to the moment of the weight that’s causing the force. What we find is that there is actually a very simple
criterion for whether this force can ever go to zero, and what it tells you is what is the maximum speed, what is
the maximum rate at which this angle can be turned, and therefore the maximum linear speed that a walker can
achieve. It is given by this formula here. It depends on the leg length, depends on the stride length, leg
length, and the acceleration due to gravity.

So, if we sort of write it rather bigger here, and we put some numbers in - leg length of one metre, stride length
of 1.3 - then the maximum speed that you can achieve without breaking contact is only about 1.7 metres per
second. But the top walkers, we have seen, like Mr Perez, they are going faster than four metres per second,
over distances of up to twenty kilometres. So this is not mechanically possible without breaking contact, and this
simple calculation just shows you that.

Well, you could say what could you do, finally, to try and make things worse? Well, you could reduce your stride
length, so if you make $S$ smaller, you take a smaller amount there, and you get a bigger $V$, but you cannot make
it much smaller than the leg length and there is a limit to how small you can make it. Instead, what you do is
to effectively increase your stride length by swinging your hips, so that is why
walkers swing their hips and walk in this unusual style. It is a way of artificially lengthening your stride by
increasing this…adding a distance in here, so your stride length becomes much bigger than it would have
become if you did not swing your hips.

So, if you do this, what you find is that you might increase your leg length a little bit, but you still cannot get
anywhere near the four metres per second speeds that are observed. You would have to have a leg that was 2.3
metres long to do that, or your hip sway would have to be about three-quarters of a metre. So you would have
to have a very unusual physiology indeed.

This rather simple analysis tells you that there is lifting going on here, but it might not be terribly dramatic and it
might only be perceptible by really quite high-speed photography.
I hope I have given you a few different insights into some aspects involving sort of records and what you can do to gain an advantage, both fairly and unfairly. Thank you.

© Professor John Barrow 2012