FUSSBALL

Ist Fußball ein Glücksspiel?

ADIDAS

Nicht ganz. Aber der Zufall wird auch bei der anstehenden Weltmeisterschaft eine so überragende Rolle spielen, dass man mit einer statistischen Analyse ziemlich weit kommt.

Von Holger Dambeck

Der Frust bei den Bayern saß tief. Mit einem Sieg gegen den englischen Verein Bolton Wanderers hätten die Münchner im November 2007 den Einzug in die nächste UEFA-Pokal-Runde klarmachen können. Doch das Spiel in München endete 2:2. Hinterher gab es Kritik an der Personalpolitik von Bayern-Trainer Ottmar Hitzfeld, der wichtige Spieler wie Philipp Lahm oder Franck Ribéry nicht aufgestellt oder früh ausgewechselt hatte. »Fußball ist keine Mathematik«, stichelte Bayern-Vorstand Karl-Heinz Rummenigge nach dem Spiel in Richtung des einstigen Mathelehrers – und lag damit falsch.

Schon seit Jahren analysieren Wissenschaftler das Spiel auf dem Rasen, vor allem mit mathematischen Methoden. Statistiken über verwandelte Elfmeter oder die Auswirkungen Roter Karten sind dabei noch vergleichsweise simpel. Anspruchsvoller wird es, wenn man versucht, den Einfluss von Glück und Pech abzuschätzen oder gar eine ganze Liga zu modellieren.

Forscher wie Fans sind sich einig: Die Unvorhersehbarkeit macht den Reiz von Fußball aus. Alles ist möglich, auch eine Drittligamannschaft kann Spitzenteams wie Bayern, Schalke oder Leverkusen schlagen. Eine Mannschaft kommt in einem Spiel nur auf etwa 10 bis 20 Torschüsse, von denen obendrein nur ein geringer Teil erfolgreich ist. Bei so geringen Zahlen spielt unweigerlich der Zufall eine wichtige Rolle.

Dagegen fallen in einem Handballspiel oft 50 und mehr Tore, und die Erfolgsquote bei Angriffen ist viel höher. Es ist wie beim Würfeln: Je öfter geworfen wird, umso näher kommt man dem Ergebnis, das zu erwarten ist, also einer Gleichverteilung der Augenzahlen oder beim Ballspiel dem Sieg des Favoriten. Wer nur wenige Male würfeln oder schießen darf, muss vor allem Glück haben.

Theorie der Torschüsse: die Poisson-Verteilung

Statistiken bestätigen dieses Phänomen. Der Physiker Eli Ben-Naim vom Los Alamos National Laboratory hat 2007 untersucht, wie oft Spiele im Fußball, Eishockey, Basketball, Baseball und American Football überraschend ausgegangen sind – also mit einer Niederlage des Favoriten. Fußball erwies sich dabei mit einer Quote von 45 Prozent als Spitzenreiter vor Basketball mit 36 Prozent.

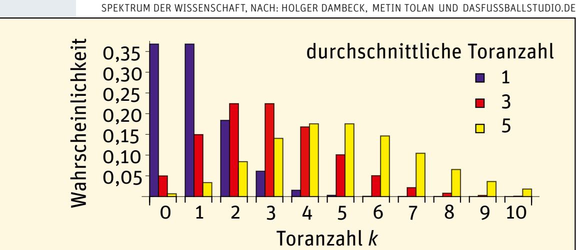
Noch besser verstehen lässt sich der Einfluss des Zufalls, wenn man das Fußballspiel mathematisch modelliert. Das erscheint zunächst als verwegenes Vorhaben. Doch es ist verblüffend einfach, solange man nur hinreichend ungenau hinschaut. Einzelne Spielzüge oder gar Torschüsse werden nicht analysiert; vielmehr fasst man sie als Ergebnis eines Zufallsprozesses auf. Man tut also zum Beispiel so, als würde während des Spiels immer wieder eine Art Würfel mit sehr vielen Seiten geworfen. Nur sehr wenige dieser Seiten tragen die Aufschrift »Tor für die Heimmannschaft« oder »Tor für die Gäste«, und ein Tor fällt genau dann, wenn bei einem Wurf eine entsprechende Seite nach oben zu liegen kommt. Die Aufgabe besteht dann darin, aus den Endergebnissen vieler Spiele zu erschließen, wie der Würfel gebaut ist.

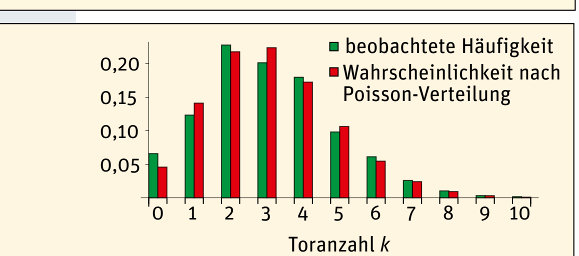
Pro Bundesligaspiel fallen im Schnitt ungefähr drei Tore. Der Dortmunder Physiker Metin Tolan hat für sein Buch »So werden wir Weltmeister« die Verteilung der Tore bei mehr als 14000 Bundesligabegegnungen ab der Saison 1963/64 untersucht. Am häufigsten, und zwar fast 3300-mal, fielen zwei Tore. Das Ergebnis lautete dann also 2:0, 1:1 oder 0:2. Bei mehr als 2800 Spielen gab es drei Treffer (3:0, 2:1, 1:2, 0:3). Die Wahrscheinlichkeit für noch mehr Tore sinkt danach rasant. So fielen nur bei 47 Spielen neun Tore.

Trägt man die Anzahl der Spiele, bei denen genau k Tore gefallen sind, gegen diese Zahl k auf, so ergibt sich ein Bild, das dem Statistiker vertraut ist: die Poisson-Verteilung. Sie gilt nicht nur für Fußballtore, sondern für alle Ereignisse, die in einem gewissen Sinn als selten gelten können: Blitzeinschläge, Autounfälle, radioaktive Zerfälle … In der Theorie tritt die Poisson-Verteilung dann auf, wenn drei Voraussetzungen erfüllt sind:

- Man kann die Zeit in so kleine Intervalle einteilen, dass in jedem Intervall höchstens ein Ereignis eintritt;
- die Wahrscheinlichkeit, ein Ereignis in einem Zeitintervall irgendeiner Länge zu finden, ist proportional dieser Länge;
- und das Eintreten des Ereignisses wird nicht durch vorherige Ereignisse beeinflusst.

SPEKTRUM DER WISSENSCHAFT, NACH: HOLGER DAMBECK, METIN TOLAN UND DASFUSSBALLSTUDIO.DE





Unter der Voraussetzung, dass die durchschnittliche Anzahl der Tore, die während eines Fußballspiels fallen, bekannt ist, berechnen sich die Wahrscheinlichkeiten, dass 0, 1, 2, … oder 10 Tore fallen, nach einer Poisson-Verteilung (oben). Die Häufigkeit der Toranzahlen in allen Bundesligaspielen von 1963/64 bis 2009/10 (30. Spieltag) folgt dieser Verteilung sehr genau – mit charakteristischen Abweichungen (unten).

All dies trifft für Tore beim Fußball mit großer Genauigkeit zu.

Wenn t die mittlere Anzahl der Ereignisse pro Zeiteinheit ist, dann kann man die Wahrscheinlichkeit, dass in einer Zeiteinheit genau n Ereignisse eintreten, mit folgender Formel berechnen:

$$P(n) = t^n/n! \cdot e^{-t}$$

Dabei ist e = 2,71828… die eulersche Zahl. Je größer t ist, desto breiter und flacher ist die Kurve (Grafik S. 68, oben).

Zeichnet man nun die Poisson-Verteilung für t = 3,0836 (das ist die genaue Durchschnittszahl der Tore pro Spiel) in die Torstatistik ein (Grafik S. 68, unten), fällt zunächst die gute Übereinstimmung ins Auge. Auf den zweiten Blick entdeckt man einige Abweichungen zwischen Realität und Theorie. Ein 0:0 tritt häufiger auf als prognostiziert, Ergebnisse mit einem Tor (1:0, 0:1) sind dafür seltener. Auch drei Tore fallen weniger oft, als man laut Poisson-Verteilung erwarten würde. Dafür treten Ergebnisse wie 1:1 und 2:2 etwas häufiger auf.

Trotzdem liefert die Formel verblüffend genaue Prognosen für Torzahlen, wie folgendes Beispiel zeigt: Wie groß ist die Wahrscheinlichkeit, dass in einem Bundesligaspiel mehr als ein Tor fällt? Dazu brauchen wir nur die Wahrscheinlichkeiten für null und für ein Tor zu berechnen und diese Werte von eins abzuziehen. Für $t = 3$ erhalten wir $P(0) = e^{-3}$ und $P(1) = 3 \cdot e^{-3}$. Die gesuchte Wahrscheinlichkeit für mehr als ein Tor ist dann $P(>1) = 1 - P(0) - P(1) = 1 - 4 \cdot e^{-3} = 0{,}8$. In 80 Prozent aller Spiele müssten demnach zwei oder mehr Tore fallen. Tatsächlich waren es 81 Prozent.

Das Modell lässt sich noch verfeinern, wenn man berücksichtigt, dass es torstarke und weniger treffsichere Mannschaften gibt. Nehmen wir an, Team A schießt im Durchschnitt zwei Tore je Spiel, Team B nur eins. Wie wahrscheinlich ist dann ein 1:0? Es ist das Produkt $P_A(1) \cdot P_B(0)$, also $P(1:0) = 2 \cdot e^{-2} \cdot e^{-1} = 2 \cdot e^{-3} = 0{,}1$ = 10 Prozent. Auf dieselbe Weise errechnet man, dass die Wahrscheinlichkeit für ein 0:1 – eine Niederlage des Favoriten – genau halb so groß ist wie die Wahrscheinlichkeit für ein 1:0. Das ist plausibel, schließlich schießt Mannschaft A im Durchschnitt doppelt so viele Tore wie Team B.

Bislang ist unser Modell unvollständig, denn wir haben nur die Torgefährlichkeit

eines Teams berücksichtigt, nicht jedoch seine Abwehrstärke. Eine Mannschaft, die im Schnitt pro Spiel zwei Tore schießt, aber zugleich fast immer drei kassiert, wird auf Dauer kaum besser spielen als eine, die durchschnittlich 1,5 Tore erzielt und 1,2 Tore einstecken muss.

Tordifferenz als Fitnessmaß

Interessanterweise muss man Angriffs- und Abwehrstärke nicht als zwei getrennte Größen modellieren. Andreas Heuer vom Institut für Physikalische Chemie der Universität Münster hat gezeigt, dass es genügt, die Differenz beider Größen zu berücksichtigen. Man berechnet für jede Mannschaft individuell deren durchschnittliche Tordifferenz ΔG; sie dient als Maß für die Leistungsstärke (Fitness) der Mannschaft. Verblüffenderweise beschreibt ΔG die Stärke eines Teams besser als die erreichte Punktzahl.

Ein fiktives Beispiel: Mannschaft A hat nach drei Spielen mit den Ergebnissen 1:0, 1:0, 1:0 zwar einen besseren Tabellenplatz als Mannschaft B, die 3:0, 4:0 und 1:2 gespielt hat. Aber B hat ein ΔG von 2 und ist damit im Prinzip stärker als A mit $\Delta G = 1$. Das dürfte sich auch im Endergebnis zeigen.

Allgemein ist die in der ersten Saisonhälfte erreichte mittlere Tordifferenz der beste Indikator für das Endergebnis: Mit ihr korrelieren sowohl die zu Saisonende erreichte Punktzahl als auch die endgültige Tordifferenz stärker als mit der Punktzahl aus der Hinrunde (Grafik S. 70). So konnte man bereits zur Halbzeit der Saison 2007/08 den rasanten Aufstieg des VfL Wolfsburg anhand seiner für den Tabellenplatz ungewöhnlich guten Tordifferenz von 0 vorausahnen.

Wie ist nun das individuelle ΔG für jede Mannschaft zu berechnen? Im Prinzip, indem man den Durchschnitt über die Tordifferenzen vergangener Spiele bildet. So geht Heuer auch vor; allerdings gehen die Ergebnisse mit einem umso geringeren Gewicht in die Berechnung ein, je älter sie sind. Schließlich ändern sich Mannschaften allmählich mit der Zeit, so dass die Fitness von vor drei Jahren nicht mehr allzu viel über die gegenwärtige Stärke aussagt.

Bei den umfangreichen statistischen Untersuchungen von Spielergebnissen hat Heuers Team einige Überraschungen erlebt, die auch in die Modellierung mit einfließen. So beschreibt er den Heimvorteil als eine Konstante, die zu der Fitness der Heimmannschaft hinzuaddiert wird – und hat bei der Berechnung herausgefunden, dass der Heimvorteil für alle Bundesligateams innerhalb der statistischen Ungenauigkeit gleich ist, sich aber von Jahr zu Jahr ändern kann. Dass eine einzelne Mannschaft besonders heimstark ist, gehört demnach zu den typischen Fußballmythen.

Auch Positiv- und Negativserien haben die Forscher analysiert, also die Frage, ob Siege oder Niederlagen in vorherigen Spielen Auswirkungen auf das nächste Spiel haben. Erstaunlicherweise trifft das nur in einem Fall zu: Nach drei Niederlagen in Folge pflegt eine Mannschaft auch im nächsten Spiel »unter Wert« abzuschneiden.

Erreichter Platz	Team	Punktzahl	Tordifferenz	Wahrscheinlichkeit für			Platzwechsel
				Platz 1	Platz 1 bis 3	Abstiegs-platz	
1	Bayern München	36	23	0,59	0,96		
2	Werder Bremen	36	18		0,93		
3	Hamburger SV	32	11		0,44		
4	Bayer Leverkusen	30	16		0,45		
5	Schalke 04	29	9				
6	Karlsruher SC	28	–2				
7	Hannover 96	27	–1				
8	VfB Stuttgart	25	–1				
9	Eintracht Frankfurt	23	–4				
10	Borussia Dortmund	21	–4				
11	VfL Wolfsburg	20	0				
12	Hertha BSC	20	–5				
13	VfL Bochum	19	–2				
14	Arminia Bielefeld	18	–19				
15	Hansa Rostock	17	–10			0,54	
16	1. FC Nürnberg	15	–7			0,27	
17	Energie Cottbus	15	–10			0,60	
18	MSV Duisburg	13	–12			0,84	

SPEKTRUM DER WISSENSCHAFT, NACH: HOLGER DAMBECK

Die vier linken Spalten zeigen das Ergebnis der Hinrunde in der Saison 2007/08. Aus diesen Daten sowie den Ergebnissen der Spielzeiten 2004/05 bis 2006/07 hat Andreas Heuer eine Prognose für das Ergebnis am Saisonende errechnet (Balkendiagramme). Lesebeispiel: Nach der Prognose ist es fast sicher (93 Prozent), dass Werder Bremen unter die ersten drei kommt, und einigermaßen wahrscheinlich (54 Prozent), dass Hansa Rostock absteigt. Die Pfeile (letzte Spalte) kennzeichnen die Verschiebungen, die in der Tabelle bis zum Saisonende stattgefunden haben.

Wie mit dieser Methode Prognosen für die Bundesliga berechnet werden, sei am Beispiel des Spiels Bayer Leverkusen gegen Bayern München am 10. April 2010 (30. Spieltag) demonstriert. Das Match war wichtig für beide Teams: Es ging um die Meisterschaft und um die Teilnahme an der Champions League.

Aus den Tordifferenzen der Vergangenheit ergibt sich zunächst ein ΔG von 0,55 für Leverkusen und 0,80 für Bayern. Der Heimvorteil, der in der Bundesliga normalerweise etwa 0,50 beträgt, lag in der aktuellen Saison bei lediglich 0,15. Demnach lag die erwartete Tordifferenz bei 0,55 + 0,15 – 0,80 = –0,10, ein sehr geringer Vorsprung für die Bayern, der ein spannendes Spiel versprach. Nur hatte Leverkusen unmittelbar zuvor drei Niederlagen in Folge einstecken müssen, ein Demoralisierungseffekt, der mit –0,15 einzuschätzen ist. Damit fällt die erwartete Tordifferenz auf –0,25 zu Ungunsten von Leverkusen.

Die Gesamtzahl der in dem Spiel zu erwartenden Tore wurde, wie hier nicht näher erläutert wird, auf 2,85 geschätzt. Das teilt sich nach obigen Berechnungen auf in 1,55 Tore für Bayern und 1,30 für Leverkusen.

Mit diesen Angaben kann man Voraussagen treffen: In 41 Prozent der Fälle siegen die Münchner. Die Wahrscheinlichkeiten für ein Unentschieden ist 28 Prozent, die eines Bayer-Siegs liegt bei 31 Prozent. Diese Werte liegen sehr nah an den Wahrscheinlichkeiten, die sich aus den Oddset-Wettquoten ergeben haben (Sieg Bayern: 40 Prozent, Sieg Leverkusen: 31 Prozent, Remis: 29 Prozent).

Berechnet ist das alles in Sekundenbruchteilen. Die Spieler auf dem Rasen in Leverkusen kämpften 92 Minuten um den Sieg – am Ende stand es 1 : 1.

Wenn man diese Methode auf alle noch zu spielenden Begegnungen anwendet, lassen sich zu jedem Zeitpunkt einer Saison Prognosen für die Abschlusstabelle berechnen. Statt Ribéry, Pizarro oder Kießling versenkt jedoch dann ein Zufallsgenerator, der einen Poisson-Prozess nachbildet, die Bälle im Netz. Das verstärkte Auftreten von Ergebnissen wie 0 : 0 oder 1 : 1 wird durch eine kleine Korrektur berücksichtigt.

So wird die ganze verbleibende Saison 10 000-mal simuliert. Die zu erwartende Endplatzierung einer Mannschaft ist nichts anderes als der Mittelwert über alle 10 000 Simulationen. Und meist steht dabei Bayern München am Ende ganz oben. Diese Art von Mathematik könnte womöglich selbst Skeptiker wie Karl-Heinz Rummenigge überzeugen.

Wie tauglich die Methode für Weltmeisterschaften oder die englische Premier League ist, muss Heuer noch untersuchen. Zumindest bei einer WM dürfte es größere Probleme geben, weil diese nur alle vier Jahre stattfindet und sich die Zusammensetzung der Mannschaften in einem solchen Zeitraum stark ändert. Eine Berechnung der Spielstärke anhand der mittleren Tordifferenz ΔG wird so erschwert.

Sicher ist auf jeden Fall eins: die Spannung bei jedem Fußballspiel. Und selbst wenn die mathematische Beschreibung immer besser gelingt, Tore lassen sich auch künftig nicht vorhersagen. Zum Glück für alle Beteiligten. ◁

Holger Dambeck ist Redakteur bei »Spiegel online« und schreibt regelmäßig eine Mathematik-Kolumne, die jetzt auch als Buch erschienen ist (»Numerator – Mathematik für jeden«).

Ben-Naim, E. et al.: What is the Most Competitive Sport? In: Journal of the Korean Physical Society 50, S. 124 – 126, 2007; online unter http://arxiv.org/abs/physics/0512143.

Tolan, M.: So werden wir Weltmeister. Piper, München 2010.

Heuer, A., Rubner, O.: Fitness, Chance and Myths: an Objective View on Soccer Results. In: The European Physical Journal B67, S. 445 – 458, 2009; online unter http://arxiv.org/abs/0803.0614.

Heuer, A. et al.: Soccer: Is Scoring Goals a Predictable Poissonian Process? In: Europhysics Letters 89(3), 38007, 2010; online unter http://arxiv.org/abs/1002.0797.