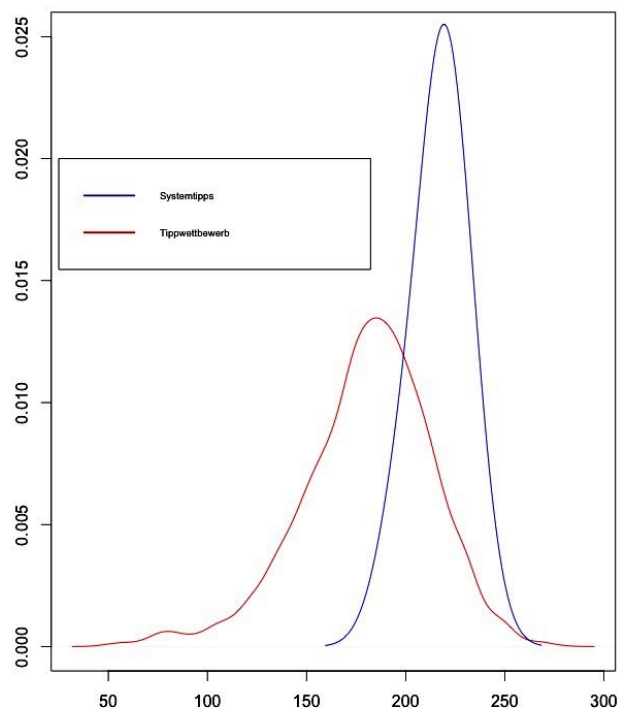


Wer hat den besten WM-Tipp - die Experten, die große Masse, die Mathematik oder doch das Wettbüro?



Antonia Lücke
Malte Braband
Christoph Friedrich

Betreuer: Herr Eberhardt, Herr Maul

Cäcilien Schule Wilhelmshaven
Neue Oberschule Braunschweig
Projekt-Nr.: 130423
Januar 2015

Kurzfassung

Die Cäcilienchule Wilhelmshaven führte zur WM 2014 einen Tippwettbewerb mit über 2400 Teilnehmern durch. Da stellt sich sofort die Frage, wer die besten WM-Tipps abgibt - und natürlich warum der (selbsternannte) „Fussballexperte“ mit seinem Fachwissen, eine Tippgemeinschaft, die die „Weisheit der Vielen“ ausnutzt, ein mathematisches Modell, das die Ergebnisse z.B. aus Ranglisten berechnet (Systemtipp) oder das Wettbüro mit seinen Quoten den größten Erfolg hat. Aus Selbsteinschätzungen der Teilnehmer haben wir Experten identifiziert und weitere Daten zu den Teilnehmern ausgewertet.

Zur wissenschaftlichen Auswertung mussten wir teilweise neue mathematische Beziehungen zwischen den verschiedenen Prognosemodellen herleiten. Außerdem haben wir eine Methodik entwickelt, mit der wir den Verlauf der WM aus den Modellen numerisch berechnen konnten, anstatt ihn wie bisher in der Literatur nur zu simulieren.

Bei der Auswertung ergab sich insbesondere, dass zwar die Experten statistisch signifikant besser abgeschnitten haben als die Durchschnittsteilnehmer und diese wiederum besser als die Anfänger (weiterhin haben die Männer besser als die Frauen und die Alten besser als die Jungen getippt). Allerdings waren die Unterschiede eher gering. Die Systemtipps haben im Mittel aber deutlich besser abgeschnitten als die menschlichen Tipper. Allerdings konnten wir nachweisen, dass auch der beste Systemtipp realistischere einen Tippwettbewerb mit sehr vielen Teilnehmern nicht gewinnen kann. Unter den Systemtipps gab es quantitativ und qualitativ relativ wenige Unterschiede, wenn man von zufälligen Schwankungen absieht. Allerdings hat sich auch herausgestellt, dass die von den Teilnehmern erhobenen Daten nicht dazu geeignet waren, das Abschneiden der Teilnehmer verlässlich vorherzusagen.

Insgesamt weisen die Ergebnisse unseres WM-Tippspiels also ähnlich wie Bundesliga-Tipprunden einen hohen Zufallsanteil auf. Verlässt man sich auf Systemtipps statt auf die eigene Intuition, dann wird man mit hoher Wahrscheinlichkeit zu den besten 20% der Teilnehmer gehören. Bei großen Tipp- runden hat man aber keine realistische Gewinnwahrscheinlichkeit, da es fast immer viele Tipps gibt, die zufällig besser sind.

Inhaltsverzeichnis

Einleitung.....	3
Forschungsstand.....	3
Vorgehensweise.....	7
Ergebnisse.....	8
Diskussion.....	14
Zusammenfassung.....	16
Abkürzungen und Symbole.....	17
Quellen- und Literaturverzeichnis.....	17

Einleitung

Die Cäcilienchule Wilhelmshaven führt seit 2006 zu jeder Welt- und Europameisterschaft einen Tippwettbewerb durch, an dem sich zur WM 2014 über 2400 Teilnehmer beteiligt haben. Wir wollten wissenschaftlich bewerten, wer die besten WM-Tipps abgab - und natürlich warum:

- Der (selbsternannte) „Fussballexperte“ mit seinem Fachwissen oder
- eine Tippgemeinschaft, die die „Weisheit der Vielen“ ausnutzt oder
- ein Computer, der die Ergebnisse z.B. aus Ranglisten berechnet (Systemtipp) oder
- das Wettbüro mit seinen Quoten.

Neben den normalen Teilnehmern werden daher im aktuellen Tippspiel zur WM 2014 auch Computertipps mitspielen, die mithilfe verschiedener Ranglisten und Wettquoten gegen die Menschen konkurrieren werden. Anhand von Selbsteinschätzungen haben wir zudem Experten identifiziert und weitere Daten zu den Teilnehmern erfragt, sodass sich noch viele weitere Untersuchungen durchführen lassen -- z.B. ob die "alten Hasen" mit ihrer Erfahrung besser abschneiden als Jugendliche oder Männer ein besseres Händchen haben als Frauen. Am Ende der WM werden wir es dann sehen!

Forschungsstand

Im Schulbuch [4] sind wichtige Wahrscheinlichkeitsverteilungen, z.B. Binomialverteilung und Normalverteilung sowie Verfahren der beurteilenden Statistik, wie z.B. Konfidenzintervalle oder Regressionsanalyse, erläutert. In der Fachliteratur, insbesondere [2] oder [19], wird vor allem die Regressionsanalyse verwendet, d.h. es wird statistisch versucht, einen möglichst guten Zusammenhang zwischen verschiedenen Einflussgrößen auf das Spielergebnis zu finden. Um Zusammenhänge statistisch zu beurteilen, werden Testverfahren benutzt, die eng verwandt sind mit Konfidenzintervallen. Bei Tests ermittelt man alle Werte, die mit einem Parameter oder einer Verteilung bei Vorliegen einer bestimmten Stichprobe vereinbar sind [6], und zwar auch unter Berücksichtigung einer Irrtumswahrscheinlichkeit α . Das heißt, das Konfidenzintervall besteht dann genau aus den Werten, deren zugehörige Wahrscheinlichkeit größer als α ist.

Prognose von Fussballspielen mittels Poissonverteilungen

Der einfachste Ansatz zur Prognose von Fußballergebnissen wird im Buch von Tolan [1] beschrieben und beruht im Wesentlichen auf der Annahme, dass die Tore jeder Mannschaft in einem Spiel Poisson-verteilt sind und zwar unabhängig von der gegnerischen Mannschaft.

Die Poisson-Verteilung ist eine diskrete Wahrscheinlichkeitsverteilung. Sie wird durch einen einzigen Parameter $\lambda > 0$, der gleichzeitig der Erwartungswert für die Anzahl von Toren pro Spiel für eine Mannschaft ist, bestimmt:

$$P(T = t) = \frac{\lambda^t}{t!} e^{-\lambda} \quad (1)$$

Gibt man diese Anzahl λ vor, kann man mit Formel (1) die Wahrscheinlichkeit ausrechnen, dass diese Mannschaft t Tore in einem Spiel schießt. Abbildung 1 (links) zeigt die Poisson-Verteilung für Werte, wie sie typischerweise in Fußballspielen vorkommen können.

Die Poisson-Verteilung hat folgende wichtige Eigenschaften [7]: Die Summe von Poisson-verteilten Zufallsgrößen ist wieder Poisson-verteilt. Die Parameter addieren sich dabei. Aber die Differenz von Poisson-verteilten Zufallsgrößen ist nicht wieder Poisson-verteilt, sondern es ergibt sich eine andere Verteilung, die sogenannte Skellam-Verteilung [14] (Abbildung 1 rechts zeigt typische Beispiele).

Das konkrete Spielergebnis ergibt sich als Realisierung der Torzahlen T_A und T_B zweier Poisson-Zufallsexperimente mit Parameter $\lambda_A = E(T_A)$ und $\lambda_B = E(T_B)$. Die erwarteten Torzahlen beider Mannschaften müssen geschätzt werden. D.h. das Fußballspiel kann man statistisch als ein Würfelspiel mit einem Poisson-Würfel interpretieren [5].

Die Wahrscheinlichkeit für ein spezielles Spielergebnis für Mannschaft A gegen Mannschaft B, z. B. 3:0, ergibt sich dann nach Formel (1) zu $P(T_A = 3)P(T_B = 0)$. So kann man sich für jedes Spieler-

gebnis die Wahrscheinlichkeiten ausrechnen. Außerdem kann man sofort die Wahrscheinlichkeiten für eine bestimmte Tordifferenz k ausrechnen (Summe über alle Spielausgänge mit der Tordifferenz). Genauso kann man die Wahrscheinlichkeiten für ein Unentschieden, einen Sieg oder eine Niederlage berechnen (Summe über die entsprechenden Tordifferenzen).

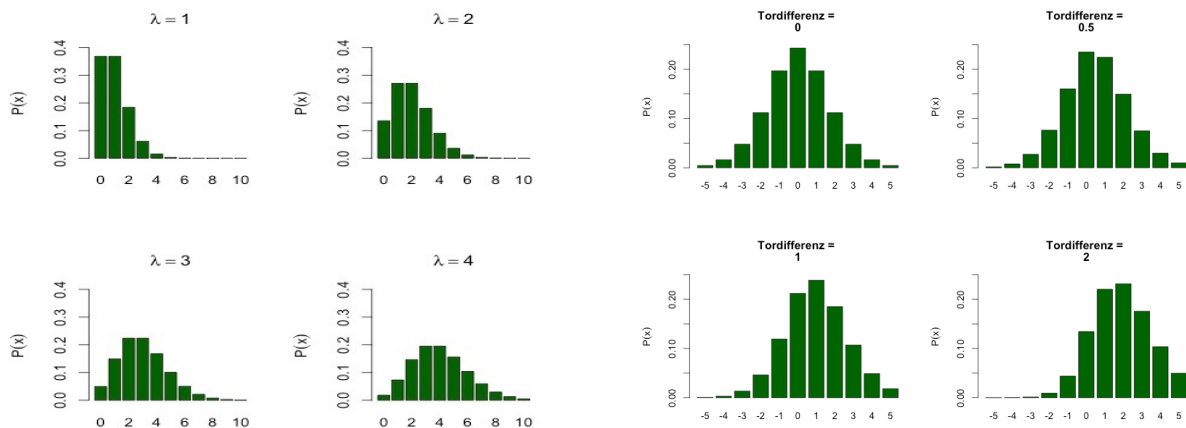


Abbildung 1: Typische Poisson- und Skellam-Verteilungen

Das Modell nach Tolan berücksichtigt aber nicht den Einfluss des Gegners. Heuer [2] wählt einen Ansatz mittels Tordifferenzen zur Prognose des Spiels von A gegen B:

$$\lambda_A - \lambda_B = LS_A - LS_B \text{ (erwartete Tordifferenz)} \quad (2)$$

$$\lambda_A + \lambda_B = TS_A + TS_B - TS \text{ (erwartete Torsumme)} \quad (3)$$

Die Leistungsstärke LS_A ist bei Heuer definiert als die mittlere Tordifferenz beim Spiel gegen einen Gegner mittlerer Stärke auf neutralem Platz. Dieser Ansatz passt besonders gut für Meisterschaften von Vereinsmannschaften, wo quasi dieselben Mannschaften jedes Jahr gegeneinander spielen und es Statistiken von Vorsaisons gibt. Wenn man alle Parameter auf der rechten Seite schätzen kann, kann man die Gleichungen (2) und (3) zur konkreten Spielprognose nach λ_A und λ_B auflösen. Zusätzlich kann der Heimvorteil $HV_A = LS_A^H - LS_A^A$ von A berücksichtigt werden, indem man HV_A zur in (2) berechneten Leistungsstärke der Heimmannschaft addiert. Man könnte aber auch wie in [5] gleich eine Heim- und eine Auswärtsleistungsstärke für jede Mannschaft ansetzen. Schließlich kann man nach Heuer noch den Marktwert der Mannschaft zu Saisonbeginn berücksichtigen (Koeffizienten s. [2] S.144):

$$LS_A = f_1 TD_{A,V} + f_2 TD_{A,S} + f_3 (\ln MW_A - \ln MW) \quad (4)$$

Soccer Power Index

Der Soccer Power Index (SPI) wurde von dem für Wahlprognosen bekannten amerikanischen Statistiker Nate Silver entwickelt [21] und bereits relativ erfolgreich für die Prognose der WM 2010 angewendet. Die genaue Herleitung ist kompliziert, aber letztendlich wird für jede Mannschaft eine Offensiv- und Defensivstärke ermittelt, die der mittleren erzielten bzw. erhaltenen Torzahl gegen eine Mannschaft mittlerer Spielstärke auf neutralem Platz entspricht. Die Differenz entspricht der Leistungsstärke nach Heuer und man kann Spielprognosen nach (2) und (3) erstellen.

Verteilung des Toto-Tipps

Toto-Tipps können mit der Skellam-Verteilung berechnet werden. Die Berechnung ist sehr kompliziert, nach [8] handelt es sich um sogenannte modifizierte Besselfunktionen, aber zum Glück hat sich jemand die Mühe gemacht, die Skellam-Verteilung für R zu programmieren. Daraus ist dann relativ einfach die Wahrscheinlichkeit für den klassischen Toto-Tipp (Heimsieg, Unentschieden, Auswärtssieg) abzuleiten. Die Funktion `dske1lam(0, x, y)` der Sprache R liefert auf einfache Weise die Wahrscheinlichkeit für ein Unentschieden (Tordifferenz 0) für die mittleren Torzahlen x und y .

$\text{pskellam}(0, x, y)$ entspricht der Verteilungsfunktion für Werte <0 und damit der Wahrscheinlichkeit für einen Auswärtssieg.

Abbildung 3 auf Seite 10 zeigt die so berechneten Wahrscheinlichkeiten für Toto-Tipps, die man in R mit dem Befehl `curve` zeichnen kann. Dabei wurde von einer bei der WM im Mittel vorkommenden mittleren Torsumme von 2,5 Toren ausgegangen, die zwischen Heim- und Auswärtsmannschaft aufgeteilt wurde, d.h. für die Heimmannschaft läuft die mittlere Torzahl x von 0 bis 2,5 und für die Auswärtsmannschaft wird als mittlere Toranzahl $2,5-x$ angesetzt. Die grüne Kurve zeigt die Wahrscheinlichkeit für Heimsieg, die rote die für Auswärtssieg und die gelbe die für Unentschieden.

Besonderheiten bei Weltmeisterschaften

Weltmeisterschaften kommen im Vergleich zu Vereinsmeisterschaften selten vor und es werden auch nur relativ wenige Spiele gespielt. Da die Qualifikation über die verschiedenen Kontinentalverbände erfolgt, spielen Mannschaften unterschiedlicher Kontinentalverbände außerhalb der Weltmeisterschaft eigentlich nur Freundschaftsspiele. Dies bedeutet, dass nur wenige Statistiken vorliegen und dass diese teilweise auch noch recht alt und damit wenig aussagekräftig sind. Deswegen ist der Ansatz aus [2] und [5] nicht direkt übertragbar.

Ranglisten

In vielen Sportarten, auch für Fussball-Nationalmannschaften, werden Ranglisten geführt. Dabei wird angenommen, dass eine Mannschaft über einen gewissen Zeitraum eine etwa gleich große Leistungsstärke LS besitzt, die aus den Spielergebnissen geschätzt wird, und mit Hilfe derer die Mannschaften nach Leistungsstärke sortiert werden können, um z. B. Setzlisten für Turniere zu bilden. Für alle Ranglisten-Systeme sollte gelten:

- Eine Mannschaft mit einer besseren Platzierung sollte gegen eine schlechter platzierte Mannschaft bei einer großen Anzahl von Spielen häufiger gewinnen
- Die Gewinnwahrscheinlichkeit sollte umso größer sein, desto größer der Abstand in der Rangliste oder in der Leistungsstärke ist

Die FIFA-Rangliste [11] beruht auf einem Ansatz, bei dem für jedes Länderspiel eine Wertzahl ermittelt wird, die von den folgenden Parametern abhängt: dem Spielergebnis, der Wichtigkeit des Spiels, der Stärke des Gegners sowie dem Mittelwert der Stärken der Kontinentalverbände beider Mannschaften. Der Gewinner erhält diese Punktzahl. Die Ranglistenpunkte ergeben sich aus einem gewichteten Mittel aller Punkte aus den letzten 4 Jahren. Der Haupt-Kritikpunkt am FIFA-System besteht darin, dass sich auch die Punktzahl einer Mannschaft, die alle Spiele gewinnt, verringern kann, wenn sie z. B. als Gastgeber keine Pflichtspiele machen muss oder nur gegen schwächere Gegner spielt. Andererseits kann sich aber auch die Punktzahl ohne Spiele erhöhen, nur weil sich das Zeitfenster für die Berechnung geändert hat.

Daneben gibt es noch inoffizielle Elo-Ranglisten für Nationalmannschaften [12], die auf den im Schach bekannten System des ungarischen Physikers Elo beruhen. In diesem System bekommt das Team, das gewinnt, direkt Punkte vom unterlegenen Team. Der Haupt-Kritikpunkt ist hier, dass ein Team, das nicht spielt, oder sich nur schwächere Gegner aussucht, keine Punkte verliert. Interessanterweise nutzt die FIFA dieses System für Frauen-, aber nicht für Männer-Nationalmannschaften.

Prognose von Fussballspielen aufgrund von Ranglisten

Für die FIFA-Rangliste scheint es derzeit kein Verfahren zu geben, um aufgrund der Ranglisten-Position oder den -Punkten konkrete Spielprognosen abzuleiten. Die Elo-Rangliste dagegen beruht auf einem statistischen Modell [13]. Wenn r die Elo-Zahl einer Mannschaft ist, dann gilt für das Modell die Annahme, dass die wahre Spielstärke R normalverteilt mit dem Mittelwert r und einer Streuung von 200 (für jede Mannschaft gleich) ist. Spielt A gegen B, so gewinnt A, wenn deren tatsächlich realisierte Spielstärke größer ist als die von B. Die Differenz der Spielstärken ist dann wieder normalverteilt mit dem Mittelwert $\Delta r = r_A - r_B$ und der Streuung $200\sqrt{2}$. Damit kann man die Wahrscheinlichkeit berechnen, dass A gewinnt:

$$W(\Delta r) = P(A \text{ gewinnt}) = P(R_A > R_B) = P(R_A - R_B > 0) = 1 - \Phi_{\Delta r, 200\sqrt{2}}(0) \quad (5)$$

wobei Φ die Verteilungsfunktion der Normalverteilung ist.

Interessanterweise ist im Elo-Modell die Wahrscheinlichkeit für ein Unentschieden gleich Null, da die Normalverteilung eine stetige Verteilung ist (siehe S.8). Da die Berechnung der Normalverteilung ohne Computer mühsam ist, hat Elo noch eine einfache Näherungsformel angegeben:

$$W(\Delta r) = \frac{Q_B}{Q_A + Q_B} = \frac{1}{1 + 10^{-\Delta r/400}} \quad \text{mit } Q_A = 10^{r_A/400}, Q_B = 10^{r_B/400} \quad (6)$$

Außerdem kann der Heimvorteil berücksichtigt werden. Hier wird pauschal vorgeschlagen, die Punktzahl der Heimmannschaft um 100 zu erhöhen. Heuer hat diesen Zusammenhang allerdings nicht benutzt, sondern er hat für Elo-Zahlen für Fußball-Bundesliga-Vereine mit einer Regressionsanalyse gezeigt [2], dass die Tordifferenz linear von deren Elo-Zahlen abhängt. Leider hat er aber nicht das Bestimmtheitsmaß angegeben. Wenn $\bar{R}$ die mittlere Elo-Zahl aller WM-Teilnehmer ist, gilt

$$LS = 0,0045(R - \bar{R}) \quad (7)$$

Prognose von Fussballspielen aufgrund von Wettquoten

Im Internet veröffentlichen die Wettanbieter die Quoten für Fussballwetten, manchmal gibt es bis zu 300 verschiedene Wetten für ein einziges Spiel. Interessant sind dabei vor allem die Quoten für eine gewisse Anzahl von Toren bzw. von Toto-Tipps. Leider ist die exakte Grundlage der Wettquoten ein Geschäftsgeheimnis der Wettanbieter. Im Prinzip sollte zwischen der Quote q für einen Tipp (Auszahlung bei Gewinn der Wette für 1 € Einsatz) und der Wahrscheinlichkeit p für diesen Tipp ein einfacher Zusammenhang bestehen, nämlich

$$p = 1/q. \quad (8)$$

In diesem Fall würde sich eine faire Wette ergeben. Allerdings würde der Wettanbieter hier keinen Gewinn machen. Daher verringert der Wettanbieter die Quote um einen bestimmten Prozentsatz. In [5] wurde der Gewinnanteil vom Wettanbieter b_{win} mit ca. 8% abgeschätzt.

In [5] wurde außerdem anhand von Beispielen gezeigt, dass es plausibel ist, dass auch die Wettanbieter Poisson-Verteilungen zur Prognose benutzen. Wenn die Quote für die Wette „Die Mannschaft schießt null Tore“ bekannt ist, so weiß man aus Formel (1) ($t = 0$), dass dann bei einer Poisson-Verteilung $p = e^{-\lambda} = 1/q$ gelten muss und der Mittelwert aus der Quote berechnet werden kann:

$$\lambda = \ln q \quad (9)$$

Man kann die Kurven in Abbildung 3 (Seite 10) auch dazu nutzen, um aus den Wettquoten bzw. den Wahrscheinlichkeiten für den Toto-Tipp die mittleren Torzahlen der Mannschaften zurückzurechnen. Dabei muss man natürlich auch noch die Torsumme schätzen.

Prognose von Fussballspielen aufgrund der „Weisheit der Vielen“

Surowiecki beschreibt das Prinzip in seinem Buch [14]. Das Prinzip ist vielen auch z. B. aus der Quiz-Show „Wer wird Millionär“ als Publikums-Joker bekannt. Wenn der Kandidat die Antwort nicht sicher weißt, lässt er das Publikum abstimmen. In der Regel wählt er dann die Antwort, die die meisten getippt haben und liegt damit in über 90% der Fälle richtig. Insbesondere ist die Quote deutlich besser als beim Experten- oder Telefon-Joker, bei dem nur eine Person befragt wird.

Dieses Prinzip kann man auch bei der Prognose von Fussballspielen anwenden. Hier gibt es ebenfalls nur eine richtige Antwort, die allerdings erst später feststeht. Nach Surowiecki muss aber eine ausreichende Meinungsvielfalt und Unabhängigkeit in der Gruppe gegeben sein, damit man gute Ergebnisse bekommt.

WM-Prognosen

Vor jeder WM werden zahlreiche Prognosen veröffentlicht, in der Regel Tipps auf das Abschneiden bestimmter Mannschaften oder „Wer wird Weltmeister?“. In Jahr 2014 hat sich sogar der weltberühmte Physiker Stephen Hawking daran beteiligt [22]. Dabei konnte man beobachten, dass die Weltmeister- oder Halbfinal-Tipps alle relativ ähnlich waren, nur Tolan wich mit seinem eindeutigen Favoriten Deutschland stark ab [23].

Allen Prognosen scheint gemein, dass sie entweder Formeln aufgrund von Regressionsmodellen aufstellen oder den WM-Verlauf simulieren. So wurden z. B. kurz vor der WM an der FU Berlin ein WM-Simulator im Rahmen einer Bachelorarbeit [19] und an der Universität Frankfurt am Main neue Simulationsergebnisse[24] veröffentlicht.

Vorgehensweise

Im Tippwettbewerb der Cäcilienchule Wilhelmshaven müssen die Tipps komplett vor dem ersten Spiel abgegeben werden. Dadurch werden einige Berechnungen vereinfacht, da sonst die geschätzten Leistungsstärken der Mannschaften während des Turniers in Abhängigkeit der erzielten Ergebnisse angepasst werden müssten. Für korrekt vorhergesagte Platzierungen gibt es Zusatzpunkte und nur, wer die KO-Partien aufgrund seiner Vorrundentipps korrekt vorhergesagt hat, kann in der Hauptrunde noch weitere Punkte machen. Dies weicht z. B. von der Kicktipp-Plattform [17] ab, bei der man jedes Spiel bis zum Anpfiff tippen kann.

Zusätzlich zu den Tipps wurden die folgenden Daten über die Teilnehmer erhoben: Vor- und Nachname, Wohnort und Postleitzahl, E-Mail Adresse, Alter und Geschlecht sowie Angaben bezüglich der sportlichen Aktivität, dem Interesse an Fußball und dem Lieblingsverein. Aufgrund dieser Daten wurden die Teilnehmer in „Experten“ und „Nicht-Experten“ eingeteilt. Für das Interesse am Fußball, wurden Punkte von „1: gar nicht“ bis „5: sehr stark“ vergeben. Außerdem wurden für die Fussball-Aktivität folgende Punkte vergeben:

- 1: gar nicht → 1 Punkt
- 2: in der Freizeit → 3 Punkte
- 3: im Verein → 5 Punkte

Falls Angaben bezüglich des Lieblingsvereins gemacht wurden, konnten noch 2 Extrapunkte erreicht werden. Wenn Aktivität bzw. Interesse nicht angegeben waren, wurden auch keine Punkte vergeben. Mit unserem Bewertungssystem waren insgesamt Punktzahlen zwischen 0 und 12 erreichbar. Die Teilnehmern wurden danach wie folgt eingeteilt:

- 0 – 3 Punkte → Kategorie Anfänger
- 4 – 8 Punkte → Kategorie Normalos
- 9 – 12 Punkte → Kategorie Experten

Außerdem wurden aufgrund der abgegebenen Einzel-Tipps nach dem Ansatz der „Weisheit der Vielen“ Tipps für „Tippgemeinschaften“ gebildet. Zusätzlich haben wir noch zum Vergleich auf Tipps größerer Tippgemeinschaften wie Kicktipp [17] zurückgegriffen. Hier ist allerdings nur die Vorrunde vergleichbar.

Für die Systemtipps wurden erst einmal Ranglisten untersucht und versucht, die Informationen aus den Ranglisten auf das Poisson-Modell abzubilden. Außerdem wurde untersucht, wie aus den Quoten der Wettbüros Daten für das Poisson-Modell ermittelt werden können. Zusätzlich wurde noch angestrebt, aus den Daten der Marktwerte der Nationalteams ein Poisson-Modell zu ermitteln. Schließlich wurden auch noch Tipps aus kombinierten Modellen, z. B. Rangliste und Marktwert, gebildet. Die Daten wurden aus den folgenden Quellen zu den jeweils angegebenen Zeitpunkten gewonnen:

Quelle	Datum	Bemerkung
Tippwettbewerb [16]	12. Juni 2014	wird komplett vor WM-Beginn getippt
Wettquote bwin [28]	1. Juni 2014	
FIFA-Rangliste [11]	8. Mai 2014	wird nur monatlich aktualisiert, nachher fanden nur noch Testspiele statt
ELO-Rangliste [12]	19. Mai 2014	Zur Vergleichbarkeit dieselbe Datenbasis wie [19]
Kicktipp [21]	12. Juli 2014	Daten erst nach WM-Ende erhältlich
Transfermarkt [15]	1. Juni 2014	
Soccer Power Index [17]	1. Juni 2014	Zur Vergleichbarkeit dieselbe Datenbasis wie [19]

Nach der WM wurden die Tipps ausgewertet und es wurde versucht, Erklärungen für die Ergebnisse zu finden, z. B. warum gewisse Gruppen oder Tipps besser abgeschnitten haben als andere. Für statistische Berechnungen bzw. Simulationen wurden Excel und die Programmiersprache R [9] benutzt, die frei erhältlich ist und mit der solche Berechnungen einfach programmiert und Grafiken erzeugt werden können. Alle für diese Arbeit benötigten Verteilungen und Verfahren sind in R vorhanden. Für jede Verteilung kann man einfach alle Berechnungen durchführen. Konkret gearbeitet wurde mit der R-Umgebung RStudio [10] sowie einer Anleitung [3], die die wichtigsten Konzepte erklärt.

Ergebnisse

Abbildung 2 stellt in einer Übersicht die Zusammenhänge und Abhängigkeiten der verschiedenen WM-Spielprognosen dar. Die in schwarz dargestellten Zusammenhänge entsprechen dem oben dargestellten Forschungsstand. Die farbig dargestellten Zusammenhänge wurden im Rahmen dieser Arbeit entwickelt und werden in den folgenden Unterabschnitten erläutert.

WM-Tipps aus Ranglisten

Beim Elo-System besteht der größte Nachteil bei Einzelspiel-Prognosen darin, dass es nach Formel (5) und (6) keine Unentschieden gibt. Im Schach spielt das keine so große Rolle, da meistens ein komplettes Turnier für einen Spieler ausgewertet wird und man sich nur für die erwartete Punktzahl interessiert. Dabei werden Remis mit 0,5 bewertet.

In Abbildung 3 haben wir einen Zusammenhang zwischen den Toto-Tipps und den erwarteten Torzahlen dargestellt. Um einen Zusammenhang mit dem Elo-Modell herzustellen, könnten wir in (5) die Unentschieden als „halbe Siege“ mitzählen. Dann könnte man diesen Zusammenhang zwischen der Elo-Gewinnwahrscheinlichkeit und der Toto-Tipp-Wahrscheinlichkeit im Poisson-Modell bekommen

$$W(\Delta r) = P(A \text{ gewinnt}) + \frac{1}{2} P(\text{Unentschieden}) = F(\lambda_A, TS) \quad (10)$$

Allerdings gilt $\Delta r = 0$ im Elo-Modell bei gleich starken Mannschaften, während dies im Poisson-Modell für $\Delta \lambda = \lambda_A - \lambda_B$ gilt. Um die beiden Modelle für eine vorgegebene Torsumme TS zu vergleichen, müssen wir im Elo-Modell den Ursprung noch um TS/2 verschieben. Dann können wir die Kurven in dasselbe Diagramm zeichnen, siehe Abbildung 3 für eine Torsumme TS=2,5. Neu hinzugekommen ist die Gewinnwahrscheinlichkeit im Poisson-Modell (blaue Kurve, die linke Seite von (10)) sowie die Gewinnwahrscheinlichkeit im Elo-Modell (die orange Kurve zeigt die Näherung nach (6), die schwarze Kurve die exakte Formel (5)).

Man erkennt vor allem im zentralen Bereich die sehr gute Näherung von (6) für (5) als auch die für praktische Auswertungen gute Übereinstimmung der beiden Seiten von (10). Die Elo-Formel hat im Vergleich zum Poisson-Modell eine geringere Streuung, d. h. der Vorteil der stärkeren Mannschaft ist etwas höher als im Poisson-Modell.

Im Vergleich zu den anderen Ungenauigkeiten erscheint uns dieser Unterschied aber gering und wir können jetzt die Modelle ineinander umrechnen. Formal ergibt sich dann für ein vorher bestimmtes TS

$$\lambda_A = F^{-1}(W(\Delta r + TS/2), TS) \quad (11)$$

Mathematisch ist die Umkehrfunktion kompliziert. Wir lesen hier entweder die Werte aus Abbildung 3 ab oder berechnen sie direkt mit R.

Bei der FIFA-Rangliste sollte ein ähnlicher Zusammenhang gelten, allerdings gibt es keinen Zusammenhang zwischen den Punkten und der Gewinnwahrscheinlichkeit. Wir versuchen, direkt einen Zusammenhang mit der erwarteten Torzahl herzustellen. Der Spitzenreiter Spanien hat derzeit etwa 1500 Punkte, die schlechtesten WM-Teilnehmer etwas über 500 Punkte. Bei gleicher Punktzahl ordnen wir den Mannschaften dieselben erwarteten Torzahlen zu und beim maximalen Abstand von 1000 Punkten nehmen wir an, dass die schlechtere Mannschaft quasi chancenlos ist. Der Nachteil ist, dass man etwa 200 Punkte Abstand braucht, damit die bessere Mannschaft gewinnt (Umskalierung).

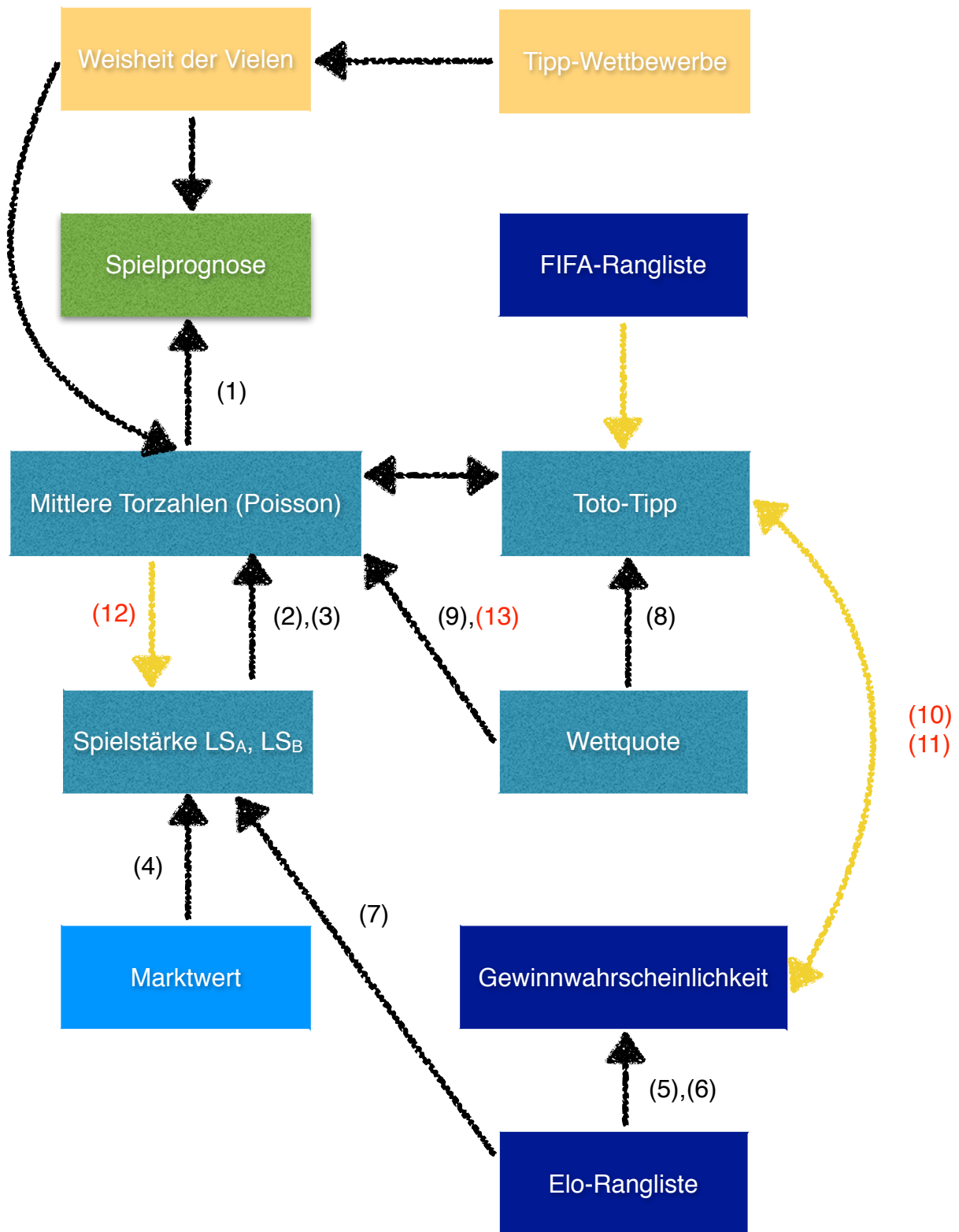


Abbildung 2: Übersicht über Prognose-Modelle und Ergebnisse der Arbeit

WM-Tipps aus Wettquoten

Bei der Spielprognose haben wir die Formeln (2) und (3) benutzt, um aus den geschätzten Leistungsstärken (Statistiken) die erwarteten Torzahlen der beiden Mannschaften zu berechnen. Da wir keine verlässlichen Statistiken für Nationalmannschaften ähnlich der Bundesliga haben, versuchen wir nun die Leistungsstärken aus den erwarteten Torzahlen für ein Spiel zu schätzen. Die erwarteten Torzah-

len werden mit Hilfe von (9) berechnet. Allerdings stellen die Wettanbieter vor WM-Beginn nur die kompletten Wetten für die ersten Spielpaarungen bereit.

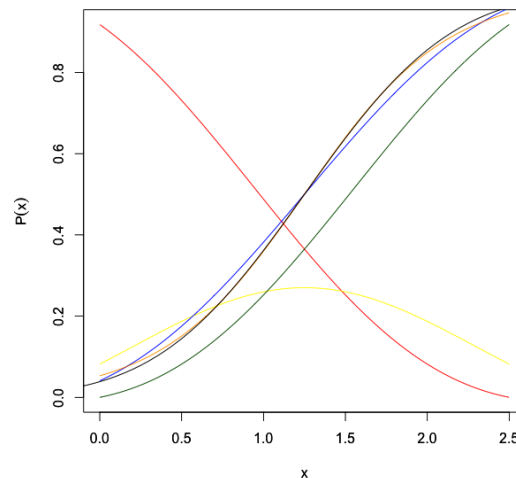


Abbildung 3: Wahrscheinlichkeiten für Toto-Tipps mit Skellam-Verteilung und Elo-Modell

Beispiel: Für das erste Spiel Deutschland gegen Portugal betragen die Quoten für Null Tore 4,50 bzw. 2,50. Um den Gewinn der Wettanbieter von etwa 8% auszugleichen, müssen wir diese Quoten noch entsprechend erhöhen. Dann ergeben sich mit (2) erwartete Torzahlen von 1,58 und 0,99, d.h. $LS_D - LS_P = 0,59$.

Dies bedeutet auch, dass wir für ein Spiel nur eine Gleichung mit zwei Unbekannten erhalten, denn in (3) kommen die Leistungsstärken nicht vor. Mit dem zweiten Gruppenspiel erhalten wir also zwei Gleichungen für vier Unbekannte. Für die anderen Spiele sind vor WM-Beginn nur die Quoten für Toto-Tipps bekannt. Aus den Wettquoten für die Toto-Tipps kann man die Wahrscheinlichkeiten für die Spielausgänge berechnen und mithilfe von Abbildung 3 die erwarteten Torzahlen schätzen. Dabei ist allerdings noch die Torsumme unbekannt, die bei Weltmeisterschaften geringer ist als in der Bundesliga. Wir nehmen hier wieder die Torsumme 2,5 an.

Beispiel (fortgesetzt): Für das zweite Gruppenspiel von Deutschland gegen Ghana betragen die Toto-Tipp-Quoten 1,33 bzw. 8,25. Nach Berücksichtigung des Gewinns ergeben sich erwartete Torzahlen von etwa 1,9 und 0,6, d.h. $LS_D - LS_G = 1,3$. Daraus folgt, dass man mit den ersten vier Spielen ein lineares Gleichungssystem mit vier Gleichungen und vier Unbekannten für die Leistungsstärken der Mannschaften A, B, C und D erhält.

$$\begin{aligned} LS_A - LS_B &= \lambda_A - \lambda_B \\ LS_C - LS_D &= \lambda_C - \lambda_D \\ LS_A - LS_C &= \lambda_A - \lambda_C \\ LS_B - LS_D &= \lambda_B - \lambda_D \end{aligned} \quad (12)$$

Leider erweisen sich die Gleichungen als nicht unabhängig, d. h. das Gleichungssystem ist nicht eindeutig lösbar. Zieht man von der ersten Gleichung die zweite und dann noch einmal die dritte Gleichung ab, so erhält man die vierte Gleichung. Außerdem ergibt sich aus der Differenz der ersten und der dritten Gleichung die Gleichung für das letzte Gruppenspiel zwischen B und C und aus der Differenz der ersten und der vierten Gleichung die Gleichung für das noch fehlende Gruppenspiel A gegen D. Dies kann man auch leicht nachvollziehen, da alle Gleichungen nur Differenzen darstellen und irgendwie ein fester Punkt fehlt, auf den sich die Differenzen beziehen.

Wir können uns aber diesen Fixpunkt beschaffen, indem wir für eine Mannschaft die Leistungsstärke aus einer anderen Quelle beschaffen. Z. B. können wir für die jeweils schwächste Mannschaft in der Gruppe die Wettquote im Internet nachschauen, dass diese Mannschaft Weltmeister wird. Daraus können wir die Wahrscheinlichkeit berechnen. Leider fehlte in der Literatur ein Zusammenhang zur Leistungsstärke. Diesen haben wir selbst entwickelt und können aus der WM-Prognose mit Formel (13) daraus die Leistungsstärke einer Mannschaft bestimmen und dann das Gleichungssystem lösen.

WM-Tipps aus Marktwerten

Heuer hat mit Formel (4) den grundsätzlichen Ansatz zur Berücksichtigung von Marktwerten geliefert, den man auch hier verwenden kann. Man kann z. B. die aus Wettquoten resultierende Leistungsstärke mit dem aktuellen Marktwert der Mannschaft [15] mitteln. Hier ist allerdings die Frage, welchen Marktwert man für eine Nationalmannschaft ansetzen sollte: den der Startelf oder den der Startelf zusätzlich Einwechselspieler (z. B. 15 Spieler) oder sogar den des gesamten Kaders (23 Spieler)?

Die Bachelor-Arbeit von Dormagen [18] benutzt den mittleren Marktwert des Kaders. Wir folgen hier dem Ansatz von Heuer, benutzen aber den Marktwert der teuersten positionstreu besetzten Startelf, wie man sie auf der Transfermarkt-Webseite [15] berechnen lassen kann.

WM-Tipps aufgrund der „Weisheit der Vielen“

Aus den Daten des eigenen Tippwettbewerbs sowie des offenen kicktipp-Wettbewerbs wurde der am häufigsten vorkommende Spielausgang ermittelt. Außerdem wurde ein Tipp ermittelt, bei dem zuerst der am meisten getippte Toto-Ausgang und dann das dabei am meisten getippte Ergebnis ermittelt wurde (sog. Tendenz-Tipp).

WM-Tipps aus dem Kicker Sonderheft

Vor der WM haben wir uns das Kicker Sonderheft [27] zur Weltmeisterschaft in Brasilien gekauft. Dort wurden für jede Mannschaft eins bis fünf Sterne für die Mannschaftsstärke verteilt. Daraus haben wir heuristische Tipps folgendermaßen erstellt:

- bei gleicher Anzahl Sterne wurde Unentschieden getippt, also 0:0, bei mindestens drei Sternen 1:1
- ansonsten wurde die Differenz der Sterne genommen, wobei eine Mannschaft mit drei oder mehr Sternen immer ein Tor zugeschrieben bekommen hat
- ab dem KO-System erhielt Brasilien im Endspiel den Gastgeberbonus und die FIFA-Rangliste entschied bei gleicher Anzahl von Sternen

WM-Prognosen

Wenn man für einzelne Spiele z. B. die erwarteten Torzahlen bestimmt hat, dann braucht man das Ergebnis nicht zu simulieren, sondern man kann mit (1) die Wahrscheinlichkeiten für alle Spielergebnisse exakt berechnen. Als Tipp kann man dann das wahrscheinlichste Ergebnis nehmen. Diese Vorgehensweise hat aber den Nachteil, dass zu viele knappe Ergebnisse vorkommen und die mittlere Torzahl nicht erreicht wird. Statt der erwarteten Torzahl von 2,5 kommen nur etwa 1,5 Tore pro Spiel vor. Dies liegt an der Poisson-Verteilung (Abbildung 1). Bei einer Poisson-Verteilung mit einer mittleren Torzahl zwischen 1 und 2 liegt die maximale Wahrscheinlichkeit immer bei einem Tor, d.h. es kommen im Mittel 0 bis 1 Tor zu wenig vor. Um zu korrigieren, dass im Schnitt pro Mannschaft pro Spiel 0,5 Tore fehlen, kann man für jeden Tipp eine weitere Poisson-Zufallszahl mit Erwartungszahl 0,5 erzeugen und addiert die so erhaltene Torzahl bei beiden Mannschaften. Damit steigt die erwartete Torsumme um eins und der Toto-Tipp wird nicht verändert.

Um die Reihenfolge der Mannschaften in den Gruppen vorherzusagen, braucht man nur die Mannschaften nach ihrer Leistungsstärke z. B. Position in der Rangliste zu ordnen. Dies ist die wahrscheinlichste Reihenfolge. Wenn man Aussagen über die Unsicherheit der Prognose machen will, kann man die Wahrscheinlichkeiten bestimmen, mit denen eine Mannschaft einen bestimmten Platz belegt, z. B. Gruppensieger wird. Dies wird normalerweise per Simulation gemacht, wie z. B. in [18-20].

Wir haben bei unseren Recherchen keine Quelle gefunden, die die Prognose berechnet anstatt zu simulieren, obwohl in dem Poisson-Modell eigentlich der komplette Spielausgang statistisch beschrieben wird. Daher haben wir uns folgenden Ansatz für eine exakte Berechnung überlegt:

- I. Bestimme für jedes Gruppenspiel die erwarteten Torzahlen beider Mannschaften
- II. Bestimme für jedes (realistische) Ergebnis dessen Wahrscheinlichkeit
- III. Bestimme für die Gruppe alle Ergebnis-Kombinationen und deren zugehörige Wahrscheinlichkeit
- IV. Ermittle die Abschlusstabelle und Platzierung der Mannschaften
- V. Addiere die entsprechenden Platzierungswahrscheinlichkeiten für jede Mannschaft

Tordifferenz	>5	5	4	3	2	1	0	-1	-2	-3	-4	-5	<-5
A-B	P(>5)	P(5)	P(4)	P(3)	P(2)	P(1)	P(0)	P(-1)	P(-2)	P(-3)	P(<-4)	P(-5)	P(<-5)
C-D	P(>5)	P(5)	P(4)	P(3)	P(2)	P(1)	P(0)	P(-1)	P(-2)	P(-3)	P(<-4)	P(-5)	P(<-5)
A-C	P(>5)	P(5)	P(4)	P(3)	P(2)	P(1)	P(0)	P(-1)	P(-2)	P(-3)	P(<-4)	P(-5)	P(<-5)
B-D	P(>5)	P(5)	P(4)	P(3)	P(2)	P(1)	P(0)	P(-1)	P(-2)	P(-3)	P(<-4)	P(-5)	P(<-5)
A-D	P(>5)	P(5)	P(4)	P(3)	P(2)	P(1)	P(0)	P(-1)	P(-2)	P(-3)	P(<-4)	P(-5)	P(<-5)
B-C	P(>5)	P(5)	P(4)	P(3)	P(2)	P(1)	P(0)	P(-1)	P(-2)	P(-3)	P(<-4)	P(-5)	P(<-5)

Tabelle 1: Exakte WM-Prognose mit mehrstufigem Zufallsexperiment

Dieser Ansatz entspricht direkt den mehrstufigen Zufallsexperimenten, die schon in der 7.Klasse behandelt werden. Allerdings ist die Auswertung komplizierter. Der Ansatz ist für alle Prognose-Modelle möglich, für die erwartete Torzahlen ermittelt werden können (vergleiche Abbildung 3). Es gibt in Schritt III sechs Stufen (die Gruppenspiele), aber in jeder Stufe (Schritt II) gibt es unendlich viele Spielausgänge, deswegen muss man sich auf die typischerweise bei einer WM auftretenden Ergebnisse beschränken. Wir haben statt der exakten Ergebnisse mit den Tordifferenzen gearbeitet, und zwar mit Tordifferenzen zwischen -6 und +6, siehe Tabelle 1. Dort ist ein spezieller Ausgang rot hervorgehoben. In diesem Beispiel hätte Mannschaft A sieben Punkte und eine Tordifferenz von 3 erzielt. Sie wäre damit Gruppensieger vor D mit fünf Punkten geworden. Mit Hilfe der Skellam-Verteilung kann man nun die Wahrscheinlichkeit berechnen, mit der die sechs Spiele genauso ausgehen wie im Beispiel in Tabelle 1.

Es gibt 13 Möglichkeiten für jedes Spiel und insgesamt $13^6=4826809$ Kombinationen für jede Gruppe, für die das R-Programm auf dem iMac etwa 10 Minuten benötigte. Manche Ausgänge sind dann zwar nicht eindeutig, z. B. wenn alle Spiele unentschieden ausgehen, aber dies kommt nur mit einer Wahrscheinlichkeit kleiner als 0,0001 vor. In diesen Fällen haben wir die Mannschaft höher platziert, die in der Rangliste besser steht. Wir haben auch die Anzahl der betrachteten Tordifferenzen variiert und gesehen, dass sich die Ergebnisse z. B. bei Tordifferenzen von -5 und +5 nur in der dritten Nachkommastelle unterscheiden.

Für Liga-Prognosen wie z. B. die Bundesliga [5] mit 306 Spielen taugt der Ansatz allerdings selbst dann nicht, wenn man nur Toto-Tipps betrachten will, da man dann 3^{306} , also etwa 10^{145} Kombinationen. Aber vielleicht lassen sich Qualifikationsgruppen damit exakt prognostizieren.

In den KO-Spielen wird die Spielprognose einfacher, da es keine Unentschieden mehr geben kann und das exakte Ergebnis egal ist. Hier kann man mit den Toto-Wahrscheinlichkeiten oder Gewinnwahrscheinlichkeiten rechnen. Da man für jede Gruppe für jede Mannschaft die Wahrscheinlichkeit kennt, dass sie Erster oder Zweiter wird, hat man für jedes Achtelfinale 16 verschiedene Paarungen zu berücksichtigen. Daraus kann man für jedes Achtelfinale für 8 Mannschaften die Wahrscheinlichkeit bestimmen, dass genau diese Mannschaft weiterkommt. Für jedes Viertelfinale hat man 64 verschiedene Möglichkeiten und bekommt die Wahrscheinlichkeiten für 16 Mannschaften. Dies gibt für jedes Halbfinale 256 Möglichkeiten und für das Finale 1024.

Das Ergebnis ist überraschend: Trägt man die Leistungsstärke der Mannschaften z. B. aus dem SPI gegen die berechnete Wahrscheinlichkeit, Weltmeister zu werden, doppelt logarithmiert auf, so erhält man einen sehr guten Zusammenhang, siehe Abbildung 4.

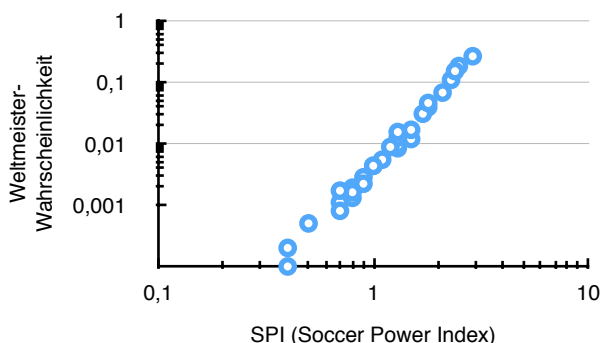


Abbildung 4: Leistungsstärke und berechnete Weltmeister-Wahrscheinlichkeit nach SPI (doppelt logarithmisch)

Mit Hilfe einer Regressionsgeraden lässt sich mit einem Bestimmtheitsmaß $R^2 = 0.985$ folgender Zusammenhang bestimmen

$$\log(P_{WM}) \approx 3.7734 \log(LS) - 2.3763 \text{ bzw. } P_{WM} \approx \frac{LS^{3.7734}}{237.84} \quad (13)$$

Auswertung des WM-Tippspiels

Auswertung nach Geschlecht

Von den insgesamt 2744 Teilnehmern machten 466 keinerlei Angaben bezüglich ihres Geschlechtes und werden deswegen in der folgenden Auswertung vernachlässigt. Von den restlichen 2278 Tippern waren 1273 Personen männlich und 1003 Personen weiblich.

Bei Betrachtung der Geschlechtsverteilung fällt auf, dass diese nur bei Kindern und Jugendlichen annähernd ausgeglichen ist. Dies hängt vermutlich damit zusammen, dass viele Schulen geschlossen mitgemacht haben. In den älteren Personengruppen liegt der Anteil der Männer wie erwartet höher. In den Kategorien „Aktivität“ und „Interesse“ ist die Tendenz, dass Männer sich allgemein mehr für Fußball begeistern deutlich sichtbar: Daraus resultiert auch, dass die Männer eher in der besseren Leistungsklasse, nämlich den „Experten“ vertreten sind, dafür aber weniger bei den „Anfängern“. Wertet man die Tipps von Männern und Frauen aus der Vorrunde getrennt aus, fällt auf, dass die durchschnittliche Punktzahl, die pro Spiel erreicht worden ist, bei Männern bei ca. 1,48 Punkten, bei Frauen hingegen bei ca. 1,42 Punkten liegt.

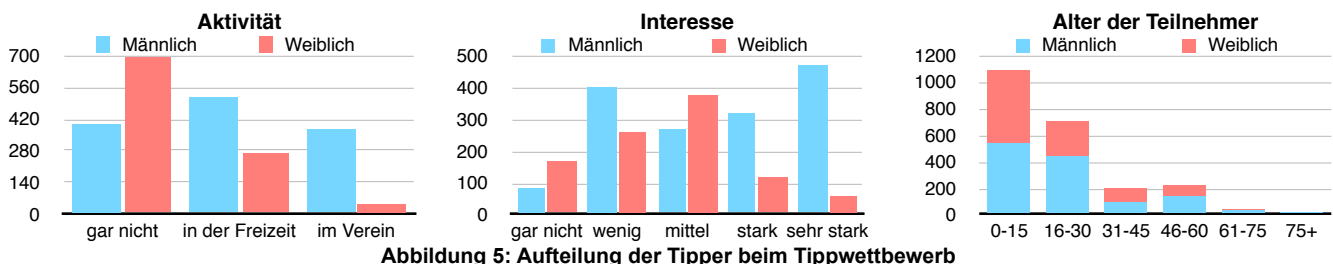


Abbildung 5: Aufteilung der Tipper beim Tippwettbewerb

Auswertung nach Leistungsklassen

Wir haben die Teilnehmer nach Fussballinteresse in drei Gruppen eingeteilt und für jede Gruppe von R eine Wahrscheinlichkeitsdichte für die Gesamtpunktzahl schätzen lassen, siehe Abbildung 6. Alle Verteilungen haben nur ein Maximum und sind mit geringen Abweichungen ziemlich ähnlich zu einer Normalverteilung. Außerdem haben wir mit dem Wilcoxon-Rangsummentest [26] die Daten von R auf signifikante Unterschiede prüfen lassen. Im Ergebnis waren alle Unterschiede aufgrund des hohen Stichprobenumfangs hoch signifikant, wenn sie auch in Abbildung 6 eher gering aussehen.

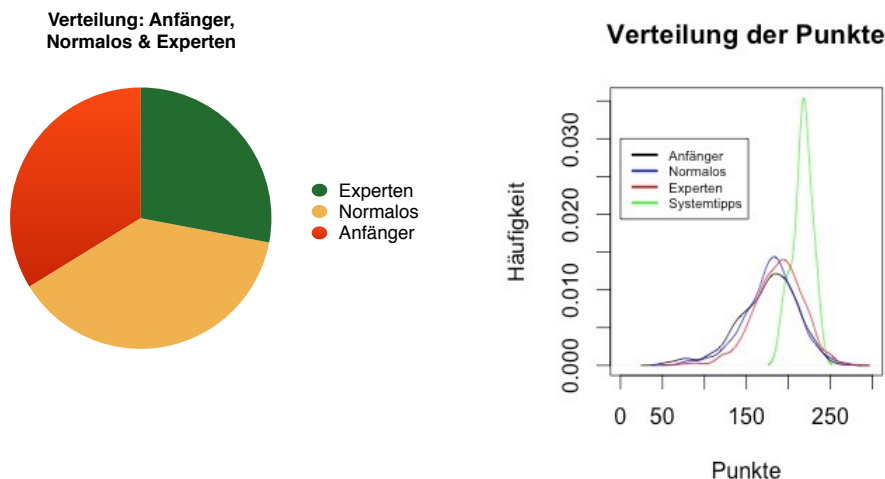


Abbildung 6: Aufteilung der Tipper und Verteilung der Gesamtpunkte beim Tippwettbewerb

Auswertung nach Alter

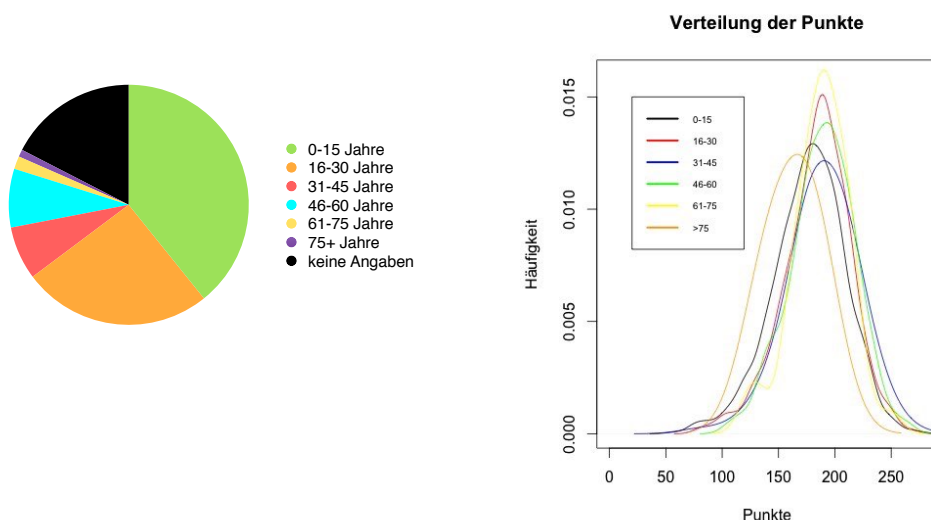


Abbildung 7: Aufteilung der Teilnehmer in 7 verschiedene Altersklassen und Verteilung beim Tippwettbewerb

Um unseren Tippwettbewerb nach Alter auswerten zu können, teilten wir die Teilnehmer in insgesamt sieben Alters-Kategorien ein. Vergleicht man die durchschnittlichen Punktzahlen pro Spiel der Vorrunde, sieht man, dass die Kategorie 75+ mit ca. 1,506 Punkten vorne liegt. Dies liegt allerdings daran, dass unsere Systemtipps in dieser Gruppe vertreten sind. Danach folgen die Kategorien 61-75 (ca. 1,485 Punkte), 31-45 (ca. 1,478 Punkte), 16-30 (ca. 1,464 Punkte) und 0-15 (ca. 1,439 Punkte) dicht hintereinander. Mit einer Durchschnittspunktzahl von nur 0,905 Punkten liegen die Personen ohne Altersangaben deutlich zurück. Es scheint also nicht wirklich vom Alter abhängig zu sein, wie gut man tippt. Prüft man diese Angaben allerdings mit dem Wilcoxon-Test [26], sieht man, dass die jungen Tipper statistisch signifikant schlechter waren als die alten Tipper.

Auswertung der Systemtipps

Auch für die Systemtipps haben wir eine Wahrscheinlichkeitsdichte mit R schätzen lassen, siehe die grüne Kurve in Abbildung 6. Es fällt die deutlich steilere Form der Dichte auf, die einen wesentlich größeren Mittelwert und eine deutlich geringere Streuung aufweist als die anderen Kurven.

Außerdem haben wir Systemtipps untereinander noch einmal qualitativ verglichen, indem wir die vorhergesagten Platzierungen der Mannschaften verglichen haben, siehe Tabelle 2. Dabei konnten wir auch Tipps einbeziehen, für die wir kein Modell hatten und die deswegen nicht an unserem Tippwettbewerb teilnehmen konnten. Dort sieht man farblich, welche Mannschaften in Wirklichkeit besser (dunkelgrün) oder schlechter (rot) abgeschnitten haben als von den Tipps prognostiziert. Grün steht für exakt richtige Tipps und orange für etwas schlechtere Tipps. Wir haben ausgezählt, wie viel richtige Mannschaften die Tipps in den Stufen der KO-Runden vorhergesagt haben, die Unterschiede waren aber minimal. Fett gekennzeichnet sind die Tipps, die signifikant besser waren als der Zufall. Insgesamt am besten hat die „Weisheit der Vielen“ abgeschnitten, was aber daran liegt, dass die meisten Teilnehmer Deutsche waren, die aufgrund der „Lokalbrille“ häufig Deutschland ins Endspiel getippt haben. Nur dieser Systemtipp sowie der von Tolan haben Deutschland als Weltmeister vorhergesagt.

Zusätzlich haben wir den Tipp der Plattform Kicktipp [17] nach der Methode „Weisheit der Vielen“ ausgewertet. Hier zeigten sich aber kaum Unterschiede zu den Systemtipps unseres Tippwettbewerbs. Die Kicktippdatei enthielt insgesamt 28864 Tipps, darunter aber nur 10785 Tipps vollständige. Aus diesen ermittelten wir für jedes Spiel sowohl den Roh-Tipp (häufigstes Ergebnis), aber auch den Tendenz-Tipp (d. h. häufigstes Ergebnis bei dem häufigsten Tototipp eines Spiels). Der Kicktipp-Tippwettbewerb unterscheidet sich insofern von unserem Wettbewerb, dass jedes Spiel erst getippt wird, wenn die Gegner bekannt sind. Für die Vorrunde macht dies keinen Unterschied, aber in der Endrunde waren die Kicktipp-Teilnehmer dadurch im Vorteil, sodass wir die Endrunde nicht vergleichen konnten. Der Roh-Tipp erreichte in der Vorrunde eine Punktzahl von 89 Punkten und der Tendenz-Tipp 92. Damit lag er nach der Vorrunde knapp hinter unserem Tendenz-Tipp, der nach der Vorrunde sogar auf Platz 39 der Gesamtwertung lag.

	Tolan [23]	Ludwig [24]	Dormagen [19]	SPI [21]	bwin [28]	Elo [12]	FIFA [11]	Marktwert[15]	WdV [16]
WM	Deutschland	Brasilien	Brasilien	Brasilien	Brasilien	Brasilien	Spanien	Spanien	Deutschland
Vize	Niederlande	Argentinien	Spanien	Argentinien	Argentinien	Spanien	Deutschland	Deutschland	Brasilien
HF	England	Deutschland	Deutschland	Deutschland	Deutschland	Deutschland	Brasilien	Brasilien	Spanien
HF	Bosnien	Spanien	Argentinien	Spanien	Spanien	Argentinien	Portugal	Argentinien	Argentinien
VF	Ghana	Uruguay	Portugal	Chile	Belgien	Niederlande	Argentinien	Frankreich	Portugal
VF	Brasilien	Italien	Kolumbien	Frankreich	Italien	England	Schweiz	Belgien	Italien
VF	Elfenbeinküste	Frankreich	Frankreich	Kolumbien	Frankreich	Portugal	Uruguay	Portugal	Belgien
VF	Argentinien	Niederlande	Uruguay	Uruguay	England	Kolumbien	Kolumbien	Italien	England
AF	Algerien	Portugal	Chile	England	Portugal	Mexico	Italien	England	Kroatien
AF	Japan	Japan	England	Belgien	Niederlande	Uruguay	Belgien	Kolumbien	Niederlande
AF	Italien	Kolumbien	Mexico	Mexiko	Bosnien	Griechenland	Griechenland	Kroatien	Griechenland
AF	Russland	Russland	Belgien	Elfenbeinküste	Mexiko	Bosnien	Chile	Niederlande	Frankreich
AF	Frankreich	Belgien	Schweiz	Bosnien	Kolumbien	Ecuador	Bosnien	Schweiz	Schweiz
AF	Schweiz	Schweiz	Russland	Ecuador	Schweiz	Frankreich	Frankreich	Bosnien	Nigeria
AF	Chile	Nigeria	Elfenbeinküste	Portugal	Russland	Belgien	Kroatien	Elfenbeinküste	Russland
AF	Kroatien	Kroatien	Bosnien	Russland	Japan	Russland	Russland	Russland	Kolumbien
AF	8	10	10	9	9	10	10	8	10
VF	4	5	5	5	5	5	4	5	4
HF	2	3	3	3	3	3	2	3	3
F	1	1	0	1	1	0	1	1	1
WM	1	0	0	0	0	0	0	0	1

Tabelle 2: Qualitative WM-Prognose der Systemtipps sowie Vergleich mit dem Ergebnis

Diskussion

Zur Vorbereitung wurden die Zusammenhänge zwischen den verschiedenen Prognosemodellen ermittelt (s. Abbildung 2). Viele Zusammenhänge waren durch die Literatur bekannt, aber einige mussten wir selbst entwickeln, wie z.B. die Berechnung der Leistungsstärke über die Wettquote oder den Zusammenhang zwischen Gewinnwahrscheinlichkeit und Toto-Tipp (Formeln (10), (11), (12)). Damit wurde ein geschlossener Zusammenhang zwischen den Tipps ermittelt, der in dieser Form in der Literatur nicht vorkommt.

Außerdem beschränkt sich die Literatur auf die Simulation von WM-Ergebnissen ([1], [2], [19], [23], [24]). Dies wurde durch persönliche Kommunikation mit den Wissenschaftlern der Universität Frankfurt am Main sowie der TU Berlin bestätigt [25]. Dagegen haben wir eine Möglichkeit gefunden, mit der wir den WM-Verlauf numerisch berechnen können. Der Vorteil gegenüber einer Simulation liegt insbesondere darin, dass weniger Rechenzeit benötigt wird und die Ergebnisse bis auf Rundungsfehler exakt sind. Natürlich ergeben sich die gleichen Resultate und man muss auch zugeben, dass dieser Ansatz nur für kleine Gruppengrößen und wenige KO-Spiele, wie das bei WM, EM oder Champions League der Fall ist, funktioniert. Insbesondere ist es uns mit diesem Ansatz gelungen, einen direkten Zusammenhang zwischen der Leistungsstärke einer Mannschaft und der Wahrscheinlichkeit, Weltmeister zu werden, herzustellen (Formel (13)). Dies hat sich als nützlich für die Bestimmung der WM-Tipps aus Wettquoten herausgestellt.

Zur Auswertung der Tipprunde wurden die Teilnehmer anhand ihrer selbst angegebenen Daten in drei Gruppen eingeteilt: Anfänger, Normalos und Experten. In der Auswertung kam heraus, dass die Experten im Schnitt nur acht Punkte vor den Normalos lagen und diese nur weitere drei Punkte vor den Anfängern. Statistisch gesehen sind diese Effekte wegen des hohen Stichprobenumfangs zwar signifikant, allerdings ist der Abstand relativ gering, wie man auch an den geschätzten Dichtefunktionen in Abbildung 6 sieht. Bei Frauen und Männern tritt dieser Effekt ebenfalls auf. Die Frauen schnitten acht Punkte schlechter ab als die Männer. Beim Vergleich der jungen und der alten Teilnehmer war das Ergebnis ähnlich. Den Tippwettbewerb hat aber eine Schülerin gewonnen! Hieran sieht man, wie viel

Zufall im Spiel ist. Das ist gerade das Spannende an Tippspielen. Keiner kennt vorher die Ergebnisse und selbst Teilnehmer, die sich weniger für Fußball interessieren, können den Wettbewerb gewinnen.

Die Systemtipps haben im Mittel über alle Klassen am Besten abgeschlossen, aber den Tippwettbewerb nicht gewonnen. Die meisten Systemtipps sind auf den Plätzen 100 bis 200 gelandet und waren dutzende Punkte hinter den besten Tippern. Dies lässt sich mit Abbildung 6 erklären. In dieser Abbildung sind die mit R geschätzten Dichtefunktionen der Punkte von Anfängern, Normalos, Experten und Systemtipps aufgetragen. Man erkennt, dass die Systemtipps im Schnitt deutlich besser sind als die anderen drei Gruppen. Auch ist die Streuung viel kleiner als bei den anderen Gruppen, d. h. sie unterscheiden sich nicht so stark. Die Systemtipps sind zwar im Mittel besser als die anderen, aber je größer die Runde ist, mit der getippt wird, desto wahrscheinlicher ist es, dass zufällig Tipper besser abschneiden als der beste Systemtipp. In unserem Fall waren dies 86 Tipper, die besser waren als jeder Systemtipp. Dies liegt an der großen Streuung der Punkteverteilung der menschlichen Tipper. Je mehr mitmachen, desto größer wird die Wahrscheinlichkeit, dass zufällig einer der menschlichen Tipper ein besonders gutes Ergebnis erzielt. Diesen Effekt kann man sich auch veranschaulichen, wenn man sich 10 Münzwürfe mit einer gezinkten Münze vorstellt, sagen wir mal Kopf kommt mit Wahrscheinlichkeit 0,8 vor. Der Systemtipp kennt das Modell und würde damit im Schnitt 8 von 10 Würfeln richtig vorhersagen. Aber z. B. bei 1000 Tippern, die zufällig tippen, hätte im Schnitt einer alle Würfe richtig vorhergesagt und etwa 10 hätten neun Würfe richtig vorhergesagt. Das heißt, wenn man sich auf Systemtipps verlässt, wird man in der Regel bei Tipprunden gut abschneiden, aber bei großen Tipprunden wird man nur selten gewinnen.

Zu Beginn des Tippspiels wurden von den Teilnehmern das Alter, das Geschlecht, das Interesse, die sportliche Aktivität und der Lieblingsverein erhoben. Wenn man in R versucht, mithilfe dieser Daten eine Regression bezüglich der Gesamtpunkte im Tippspiel zu ermitteln, um zu überprüfen, wie wichtig die Daten für die Gesamtpunktzahl sind, erhält man nur geringe Bestimmungsmaße für die ermittelte lineare Gleichung. Das heißt, die von uns erhobenen Daten haben keinen großen Einfluss auf das Ergebnis im Ranking bzw. die Gesamtpunktzahl. In Abbildung 8, die die den Leistungsklassen zugrundeliegende Punktzahl der Teilnehmer den erreichten Gesamtpunktzahlen gegenüberstellt, kann man dies auch optisch gut erkennen: man kann kaum einen sinnvollen Zusammenhang zwischen den beiden Variablen herstellen.

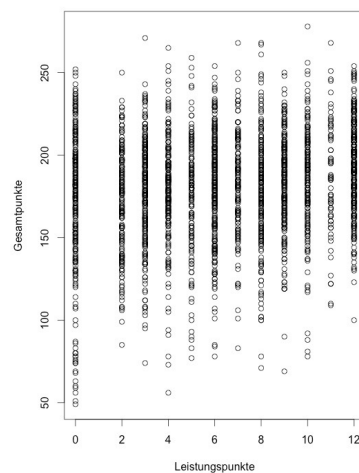


Abbildung 8: Zusammenhang zwischen Leistungsklasse der Teilnehmer und Gesamtpunktzahl

Entweder haben andere Merkmale, die wir nicht erhoben haben, einen größeren Einfluss, oder die Gesamtpunktzahl ist sehr stark vom Zufall abhängig. Dies würde sich mit den Ergebnissen von Heuer decken [2], der für die Prognose einzelner Fussballspiele bzw. für die Bundesligatabelle ebenfalls einen sehr hohen Zufallseinfluss nachgewiesen hat.

Zusammenfassung

Zur wissenschaftlichen Auswertung haben wir zunächst die Beziehung zwischen den verschiedenen Prognosemodellen ermittelt. Teilweise haben wir dabei neue Ergebnisse erzielt, die bisher in der Literatur nicht bekannt waren. Außerdem haben wir eine Methodik entwickelt, mit der wir den Verlauf der WM berechnen konnten, anstatt wie bisher in der Literatur nur zu simulieren. Damit hatten wir die Voraussetzung für eine mathematisch fundierte Auswertung geschaffen.

Damit haben wir für unser Tippspiel die folgenden Ergebnisse erzielt:

- Die Experten haben statistisch signifikant besser abgeschnitten als die Normalos und diese wiederum besser als die Anfänger. Allerdings waren die Effekte relativ gering. Im WM-Tippspiel waren die Experten nur etwa acht Punkte besser als die Normalos
- Ebenso haben die Männer statistisch signifikant besser abgeschnitten als die Frauen, aber auch hier war der Effekt mit ungefähr acht Punkten klein.
- Die älteren Tipper (über 30 Jahre) haben statistisch signifikant besser abgeschnitten als die jüngeren Tipper, allerdings war der Effekt auch hier gering. Die „alten Hasen“ hatten im Schnitt nur neun Punkte mehr als die Jugendlichen.
- Die Systemtipps haben im Mittel deutlich besser abgeschnitten als die menschlichen Tipper, allerdings konnten wir nachweisen, dass auch der beste Systemtipp realistischweise einen Tippwettbewerb mit sehr vielen Tippnern nicht gewinnen wird.
- Unter den Systemtipps gab es quantitativ und qualitativ relativ wenige Unterschiede, von zufälligen Schwankungen abgesehen.
- Besonders gut haben die Tipps vom Typ „Weisheit der Vielen“ unter Berücksichtigung der Tendenz abgeschnitten. Dies liegt aber wahrscheinlich daran, dass die Teilnehmer überwiegend Deutsche waren und damit Deutschland favorisiert haben. Da Deutschland Weltmeister wurde, hat sich dies ausgezahlt.
- Allerdings hat sich heraus gestellt, dass die von uns von den Teilnehmer erhobenen Daten nicht dazu geeignet waren, um das Abschneiden der Teilnehmer einigermaßen verlässlich vorherzusagen.
- Es gab keine wesentlichen Unterschiede zwischen der Auswertung der Vorrunde und des kompletten Tippwettbewerbs.

Insgesamt hat sich ergeben, dass die Ergebnisse von WM-Tippspielen ähnlich wie Bundesliga-Tipprunden einen hohen Zufallsanteil aufweisen. Verlässt man sich auf Systemtipps statt auf die eigene Intuition, dann wird man in den Tipprunden mit hoher Wahrscheinlichkeit zu den besten 20% der Teilnehmer gehören, aber bei großen Tipprunden trotzdem eine äußerst geringe Gewinnwahrscheinlichkeit haben

Abkürzungen und Symbole

Die Symbole wurden in Anlehnung an Heuer [2] gewählt, um eine Vergleichbarkeit der Formeln und Ergebnisse zu erleichtern.

α	Irrtumswahrscheinlichkeit (ist hier in der Regel 5%)
T_A	Tore von Mannschaft A im Spiel (Zufallsgröße)
$\lambda_A = E(T_A)$	erwartete Tore von Mannschaft A
LS_A	Leistungsstärke von Mannschaft A (mittlere Tordifferenz auf neutralem Platz)
TS_A	mittlere Torsumme aller Spiele mit Beteiligung von A
TS	mittlere Torsumme aller Spiele
HV_A	Heimvorteil Mannschaft A
MW_A	Marktwert von A
MW	mittlerer Marktwert aller Mannschaften

Quellen- und Literaturverzeichnis

- [1] Tolan, M.: Manchmal gewinnt der Bessere - Die Physik des Fussballspiels, Piper, 2011
- [2] Heuer, A.: Der perfekte Tipp, Wiley-VCH, Weinheim, 2012
- [3] Wollschläger, D.: R Kompakt, Springer, 2013
- [4] Griesel, H. et al.: Elemente der Mathematik, Schroedel, Braunschweig, 2009

- [5] Braband, M.: Fußball ist ein Würfelspiel oder: Mit welcher Wahrscheinlichkeit schafft Eintracht Braunschweig den Klassenerhalt? Jugend Forscht Projekt Nr. 123374, März 2014
- [6] Wikipedia: http://de.wikipedia.org/wiki/Statistischer_Test (15.12.14)
- [7] Wikipedia: <http://de.wikipedia.org/wiki/Poisson-Verteilung> (15.12.14)
- [8] Wikipedia: http://en.wikipedia.org/wiki/Skellam_distribution (07.01.2015)
- [9] The R project for statistical computing, <http://www.r-project.org> (07.01.2015)
- [10] RStudio, <http://www.rstudio.com> (07.01.2015)
- [11] FIFA-Weltrangliste, de.fifa.com/fifa-world-ranking/ranking-table/men/index.html (07.01.2015)
- [12] Wikipedia: World Football Elo Rankings, en.wikipedia.org/wiki/World_Football_Elo_Ratings (07.01.2015)
- [13] Wikipedia: Elo-Zahl, de.wikipedia.org/wiki/Elo-Zahl (07.01.2015)
- [14] Surowiecki, J.: Die Weisheit der Vielen. Warum Gruppen klüger sind als der Einzelne („The wisdom of crowds“). Goldmann, München 2007
- [15] Transfermarkt, www.transfermarkt.de (07.01.2015)
- [16] Tippwettbewerb zur WM 2014 der Cäcilien- und St. Marienschule Wilhelmshaven, www.tippwettbewerb.de (07.01.2015)
- [17] Kicktipp Tippspiel, www.kicktipp.de (07.01.2015)
- [18] Rojas, R.: Wer gewinnt die Fußballweltmeisterschaft?, 28.5.2014 <http://www.heise.de/tp/artikel/41/41867/1.html> (31.05.2014)
- [19] Dormagen, D.: Development of a Simulator for the FIFA Worldcup 2014, Bachelorarbeit FU Berlin, 13.05.2014 http://www.inf.fu-berlin.de/inst/ag-ki/rojas_home/documents/Betreute_Arbeiten/Bachelor-Dormagen.pdf (31.05.2014)
- [20] Errechnet: Brasilien wird Weltmeister, DIE WELT, 27.5.2015, <http://www.welt.de/sport/article128441272/Errechnet-Brasilien-wird-Fussball-Weltmeister.html> (31.05.2014)
- [21] ESPN: Soccer Power Index, <http://www.espnfc.com/spi/rankings?cc=5739> (31.05.2014)
- [22] Spiegel Online, Physiker Hawking stellt Erfolgsformel für Englands Kicker vor, <http://www.spiegel.de/sport/fussball/fussball-wm-in-brasilien-stephen-hawking-und-die-erfolgsformel-a-972206.html>, (31.05.2014)
- [23] Brasilien, Spanien oder wir! DIE ZEIT Nr. 23/2014, 28.5.2014, S. 18
- [24] DIE ZEIT: Meisterchancen, Infografik Nr. 260 vom 12.6.2014, basierend auf Simulationen von Prof. Ludwig, Universität Frankfurt am Main
- [25] E-Mail-Kommunikation mit Prof. Ludwig, Prof. Rojas sowie Herr Dormagen (B. Sc.)
- [26] Wikipedia: Wilcoxon-Mann-Whitney-Test, de.wikipedia.org/wiki/Wilcoxon-Mann-Whitney-Test, (03.01.2015)
- [27] Kicker-Sonderheft WM 2014, Olympia-Verlag, Nürnberg, 2014
- [28] bwin Sportwetten, bwin.de (07.01.2015)

Wir möchten uns bei folgenden Unterstützern bedanken:

- Unsere Lehrer Herr Eberhardt und Hr. Maul haben die Arbeit betreut und Korrektur gelesen.
- Prof. Heuer, Prof. Ludwig und Prof. Rojas haben uns bereitwillig Fragen zu ihren Arbeiten beantwortet.
- Insbesondere mit Herrn Dormagen ergab sich ein intensiver Austausch zu dem Simulationsmodell aus seiner Bachelor-Arbeit.
- Die Firma Kicktipp GmbH hat uns die Daten ihrer öffentlichen Tipprunde zur Verfügung gestellt.
- Natürlich wären diese Ergebnisse nicht ohne die zahlreichen Teilnehmer am Tipp-Wettbewerb möglich gewesen.

Gastzugang für www.tippwettbewerb.de:

Benutzername: christoph@friedrich-whv.de

Passwort: gastjufo2015