Optimale Fussballprognosen

Niels Bohr soll einmal folgendes gesagt haben: „Prognosen sind schwierig, besonders wenn sie die Zukunft betreffen“. Dies gilt nicht nur für Finanzmarktprognosen, sondern auch für Fussballprognosen. Wir entwickeln eine Fussballprognoseformel, welche versucht, anhand vergangener Spielergebnisse optimale Prognosen zu erstellen. Dabei wird der Schwierigkeit zwischen Signal und Noise zu unterscheiden Rechnung getragen. Unsere Prognose für das Eröffnungsspiel Russland : Saudi-Arabien lautet: 2:1

Erfahren Sie ausserdem, welche Spielbedingungen zu einem zukünftigen Fussballweltmeisterschaftsendspiel Bhutan gegen Peru führen könnten und wie die Fussballprognoseformel von Stephen Hawking *lautete.*

Hauptprognoseproblem von Fussballergebnissen: Es liegen zu wenige Beobachtungen vor

Fussballprognosen sind notorisch schwierig, insbesondere für ausgewiesene Nichtexperten, zu denen der Autor zählt! Warum? Das Hauptproblem besteht darin, dass es wenig verwendbare Beobachtungen gibt (Das zweite grosse Problem, welches im „Gefühlsgesteuerten – auch hormongesteuerten Prognosebias besteht, wird hier nicht behandelt. Dieser Bias macht jedoch gerade das Fussballspielen interessant. Fans dürfen Hoffnung auf den WM-Titel gegen alle Vernunft und statistische Evidenz haben und dann den Hormoncocktail während des Spieles geniessen. Eigentlich handelt es sich hier streng genommen nicht mehr um Prognosen sondern um Wunschdenken).

Betrachten wir das Eröffnungsspiel dieser WM: Russland gegen Saudi-Arabien:

Es gibt nur eine Beobachtung vom 06.10.1993. Damals gewann Saudi-Arabien 4:2 gegen Russland. Wäre 4:2 eine sinnvolle Prognose für das Eröffnungsspiel?

Nein! Warum nicht:

1. Diese Beobachtung ist so alt, dass sie praktisch keine verwertbaren Informationen enthält, da sich die WM-Mannschaften 2018 aus komplett neuen Spielern zusammensetzen. Allenfalls die grundsätzliche Fähigkeit von Russland und Saudi-Arabien Nationalmannschaften aufzustellen, könnte durch das veraltete Spielergebnis grob gemessen werden, weshalb das Ergebnis nicht völlig wertlos sein dürfte.
2. Selbst wenn das Spiel vom 06.10.1993 repräsentativ für das Eröffnungsspiel der WM 2018 wäre, bestünde das Problem, dass es sich nur um eine einzige Beobachtung handelt. Mit nur einer Beobachtung lassen sich Mittelwerte jedoch nicht seriös abschätzen. Nur wenn man sehr viele Begegnungen Russlands gegen Saudi-Arabien aus jüngster Zeit hätte, wäre die Prognose einfach. Man würde einfach die arithmetischen Durchschnitte als erwartungstreue Schätzer verwenden.

Lösung des Problems: Optimales Schrumpfen gegen leichter messbare Grössen

Die mathematische Statistik kennt Methoden, die optimale Schätzergebnisse liefern, wenn mehr als drei Mittelwerte (hier Fussballergebnisse) gleichzeitig geschätzt werden sollen. Bei der WM 2018 spielen 32 Mannschaften mit die 64 Spiele absolvieren sodass 64 Fussballergebnisse *gleichzeitig geschätzt werden müssen*. **Somit müssen alle Informationen aller Länderspiele der Vergangenheit (wenn diese nicht zu alt sind) verwendet werden. Die Grundidee hinter diesen Methoden besteht**

darin, die Prognosen in Richtung von besser messbaren Grössen zu schrumpfen. Die durchschnittliche Torschusskraft einer durchschnittlichen WM-Mannschaft ist beispielsweise aufgrund der hohen Anzahl an Beobachtungen viel zuverlässiger messbar, als die Torschusskraft von „Exotenbegegnungen", wie Russland – Saudi-Arabien.

Um das Problem zu vereinfachen, betrachten wir im Folgenden nur das Prognoseproblem *„wieviel Tore wird Russland gegen Saudi-Arabien schiessen?"* Wir negieren bis auf weiteres auch die Beobachtungen der Direktbegegnungen Russlands und Saudi-Arabiens.

„Aus dem Himmel" fällt jetzt folgende vorläufige Formel (1), die im Folgenden motiviert wird:

(1) $Torprognose^{Russland\ schiesst\ gegen\ Saudi\ Arabien} = \emptyset Tore_{\emptyset Mannschaft} + c_{Russland}(\emptyset Tore_{Russland} - \emptyset Tore_{\emptyset Mannschaft})$

Wir möchten die Anzahl Tore prognostizieren, die Russland gegen Saudi-Arabien erzielen wird ($Torprognose^{Russland\ gegen\ Saudi\ Arabien}$).

Diese Prognose ist ein gewichteter Durchschnitt zwischen der Anzahl Tore, die eine durchschnittliche Mannschaft im Durchschnitt gegen einen durchschnittlichen Gegner schiesst ($\emptyset Tore_{\emptyset Mannschaft}$ = gut messbar, da viele Beobachtungen) und der Anzahl Tore, die Russland im Durchschnitt geschossen hat ($\emptyset Tore_{Russland}$. =weniger gut messbar, da weniger Beobachtungen).

Wäre c=1 bestünde die Prognose nur aus der durchschnittlichen Anzahl Tore, die Russland in der Vergangenheit erzielen konnte.

Wäre c=0 bestünde die Prognose nur aus der Anzahl Tore, die eine durchschnittliche Mannschaft im Durchschnitt in der Vergangenheit erzielte.

Intuitiv ist klar, dass c umso näher bei 1 sein sollte, desto präziser die durchschnittliche Anzahl Tore der russischen Mannschaft geschätzt werde kann. Die Präzision dieser Schätzung ist im Wesentlichen eine Funktion der Standardabweichung der "Anzahl Tore der russischen Mannschaft". Je mehr Beobachtungen über russische Länderspiele vorliegen, je konstanter die Spielleistung der russischen Mannschaft, desto genauer kann die russische Torschusskraft geschätzt werden und desto näher liegt c bei 1.

Intuitiv sollte ebenfalls klar sein, dass c umso näher bei 0 sein sollte, desto identischer die wahre unbekannte Spielstärke aller Mannschaften ist. Ist diese komplett identisch und wären die Fussballresultate ein reines Zufallsprodukt, müsste c exakt bei 0 sein (und die Prognose würde lediglich der Prognose der Anzahl Tore dienen, die Prognose würde also immer auf "unentschieden" lauten).

Tore, die Russland schiesst und Tore, die Saudi-Arabien erduldet, müssen beide berücksichtigt werden

Bis jetzt wurden nur Spiele berücksichtigt, bei denen Russland mitgespielt hat, das heisst, es wurde nur die Torschusskraft der russischen Mannschaft berücksichtigt. Somit wurden Informationen aus Spielen der saudischen Mannschaft, das heisst die Verteidigungskraft der saudischen Mannschaft nicht berücksichtigt. Dies soll im Folgenden geschehen, denn jedes Tor, das Russland schiesst, ist ein Tor, welches Saudi-Arabien erleiden muss.

Wenn Ihnen das nicht klar ist, stellen Sie sich einfach vor, dass Sie die Anzahl Tore, die Russland gegen Saudi-Arabien schiessen wird, prognostizieren können, in dem Sie von Russlands

Torschusskraft ausgehen (Formel 1) oder indem Sie von Saudi-Arabiens Torabwehrkraft ausgehen (Formel 2), oder, was optimal ist, beides berücksichtigen, was zu Formel (3) führt:

(2) $Torprognose^{Saudi\ Arabien\ steckt\ von\ Russland\ ein} =$
$\emptyset Tore_{\emptyset Mannschaft} + c_{Saudi\ Arabien}(\emptyset Tore_{Saudi\ Arabien\ steckt\ ein} - \emptyset Tore_{\emptyset Mannschaft})$

(3) $Torprognose^{Russland\ gegen\ Saudi\ Arabien} =$
$\gamma\ Torprognose^{Russland\ schiesst\ Tore\ gegen\ Saudi\ Arabien} +$

$$(1-\gamma)\ Torprognose^{Saudi\ Arabien\ steckt\ von\ Russland\ ein}$$

Der Gewichtungskoeffizient γ muss zwischen „0“ und „1“ liegen. Je genauer (ungenauer) die „Torschusskraft Russlands“ (=„Torverhinderungskraft Saudi-Arabiens“) gemessen werden kann, desto näher liegt γ bei „1“. Der fussballpolitisch korrekte Schrumpfungsfaktor würde natürlich genau „0.5“ betragen. Wie meistens, würde jedoch folgendes Phänomen auftreten: Alles, was politisch korrekt ist, ist in der Regel hochgradig ineffizient (hier gemessen am Ziel, möglichst treffsichere Prognosen zu erstellen).

In Formel (3) wurden bisher die Ergebnisse der Direktbegegnungen Russland – Saudi Arabien nicht explizit berücksichtigt. Dies ist natürlich ebenfalls nicht optimal, auch wenn die konditionierte Torschusskraft Russlands, wenn Saudi-Arabien der Gegner ist, nur sehr schwierig gemessen werden kann, da es nur eine einzige Beobachtung gibt (so dass eher Noise gemessen wird). Die vollständige Fussballprognoseformel für die Anzahl Tore, die Russland gegen Saudi-Arabien schiessen wird, könnte somit (vereinfacht) wie folgt lauten:

(4) $Torprognose^{Russland\ gegen\ Saudi\ Arabien} =$
$(1-\tau-\psi)(Russische\ Tore^{Länderspiel\ Russland\ gegen\ Saudi\ Arabien}) +$
$\psi\ Torprognose^{Russland\ schiesst\ Tore\ gegen\ Saudi\ Arabien} +$
$\tau\ Torprognose^{Saudi\ Arabien\ steckt\ von\ Russland\ ein}$

Setzt man (1) und (2) in (4) ein, erhält man (5).

Gleichung (5) berücksichtigt nun alle vorhandenen Informationen vergangener Spiele von Russland und Saudi-Arabien. Dieser Ansatz versucht, einen optimalen Kompromiss zwischen Erwartungstreue (Bias) und Schätzgenauigkeit, letztendlich zwischen Signal und Noise zu erreichen.

(5) $Torprognose^{Russland\ gegen\ Saudi\ Arabien} =$
$(1-\tau-\psi)(Russische\ Tore^{Länderspiel\ Russland\ gegen\ Saudi\ Arabien})$
$+\ \psi(\emptyset Tore_{\emptyset Mannschaft} + c_{Russland}(\emptyset Tore_{Russland} - \emptyset Tore_{\emptyset Mannschaft}))$
$+\ \tau(\emptyset Tore_{\emptyset Mannschaft} + c_{Saudi\ Arabien}(\emptyset Tore_{Saudi\ Arabien\ steckt\ ein} - \emptyset Tore_{\emptyset Mannschaft}))$

Wie lautet nun die konkrete Prognose?

Russland hat gegen Saudi-Arabien im letzten Duell 2 Tore erzielt.
$Russische\ Tore^{Länderspiel\ Russland\ gegen\ Saudi\ Arabien} = 2$

Bei der WM 2014 wurden 171 Tore bei x Spielen erzielt. Es fanden 64 Spiele statt. Damit lautet das durchschnittliche Spielergebnis ($\emptyset Tore_{\emptyset Mannschaft}$) = 1.34 Mit anderen Worten, eine durchschnittliche Mannschaft hat im Durchschnitt bei der WM 2014 1.34 Tore geschossen. Die russische Mannschaft hat im Durchschnitt pro Länderspiel gemäss Wikipedia 1091/640 = 1.7 Tore geschossen: $\emptyset Tore_{Russland}$ =1.7 (0.95 einkassieren müssen).

Die saudische Mannschaft hat im Durchschnitt pro Länderspiel 1.6 Tore „einkassieren müssen": $\emptyset Tore_{Saudi\ Arabien\ steckt\ ein} = 1.6$ (1.13 Tore geschossen).

Mit den hier verwendeten Schrumpfkoeffizienten $c_{Russland}$= 0.8, $c_{Saudi\ Arabien}$ = 0.8, $\psi = 0.48, \tau = 0.48$ lautet die Prognose für die Anzahl Tore, die Russland gegen Saudi-Arabien erzielen wird, somit 4% * 2 + 48% * (1.34 + 0.8 * (1.7-1.34)) + 48% * (1.34 + 0.8 * (1.6-1.34)) = 1.6

Eine analoge Berechnung[1] für die Anzahl Tore, die Saudi-Arabien gegen Russland erzielen wird, führt zu 1.22 Toren.

Diese Methode führt somit zur Prognose Russland gegen Saudi-Arabien 1,6 : 1,2.

Gerundet erhält man die Prognose Russland : Saudi-Arabien 2:1.

Wir haben nun alle Informationen vergangener Spiele ausgenützt, um zu einer möglichst guten Prognose zu kommen[2]. Trotzdem ist diese sehr unsicher und das ist auch gut so, denn Fussballspiele sind genau deshalb so spannend, weil sie so schlecht prognostizierbar sind!

Dem Leser mag aufgefallen sein, dass hier ein rein statistischer Ansatz präsentiert wurde und alle physikalischen Einflüsse komplett ausser Acht gelassen wurden (Space & Time). Die Fussballprognoseformel von Prof. Stephen Hawking berücksichtigt diese Dimensionen explizit. So gibt es beispielsweise einen Heimvorteil[3], da der Jetlag geringer ausgeprägt ist, die Temperaturdifferenz (hinsichtlich Luftdruck, Luftfeuchtigkeit, Sauerstoffgehalt, Regen, Sauerstoffgehalt der Atemluft) tiefer ist usw. Diese Einflussfaktoren sprechen alle dafür, dass Russland Saudi-Arabien besiegen wird. Würden alle Qualifikationsspiele und die nächste Fussballweltmeisterschaft im Mount-Everest Basislager ausgetragen werden, müsste man vermutlich Tibet gegen Bolivien (oder doch Bhutan gegen Peru?) als Endspielkombination prognostizieren.

Exkurs I für interessierte Leser: der optimale Schrumpfkoeffizient

Man kann zeigen, dass sich für den optimalen Schrumpfkoeffizienten c eine exakte Formel herleiten lässt. Diese lautet:

$$c_{Russland} = 1 - \frac{(k-3)\sigma^2_{Russland}}{\sum_{i=1}^{k} -(\emptyset Tore_{Russland} - \emptyset Tore_{\emptyset Mannschaft})^2}$$

Was besagt nun obige Formel?

Je genauer die unbekannte wahre Torschusskraft einer Mannschaft geschätzt werden kann, desto grösser ist c und desto stärker demgemäss das Gewicht, der durchschnittlichen Anzahl Tore, welche Russland in der Vergangenheit erzielt hat. Wäre die Varianz der Anzahl Tore von Russland null, wäre c genau „1". Mit anderen Worten, hätte Russland immer genau gleich viel (Eigen-)Tore geschossen (auf die Nachkommastelle genau), wäre der Vertrauensgrad in den historischen Durchschnitt von Russland so hoch, dass es keinerlei Grund gäbe, Informationen, die aus dem Torschussdurchschnitt aller Mannschaften resultieren, überhaupt zu berücksichtigen. Die Schrumpfungstendenz, mit der die optimale Torprognose in Richtung des individuellen Tordurchschnitts der jeweiligen Mannschaft gezogen wird, ist also individuell verschieden und eine Funktion der Messgenauigkeit, d.h. der Varianz dieses individuellen Durchschnitts.

Je grösser die aufsummierte quadrierte Abweichung zwischen dem Durchschnitt aller Durchschnitte (=durchschnittliche Anzahl Tore, die in einem WM-Spiel geschossen werden) und allen individuellen Durchschnitten der Mannschaften ist, desto unterschiedlicher sind die unbekannten, wahren Torschusskapazitäten aller Mannschaften und demzufolge, desto grösser ist c. Je ähnlicher die historischen Durchschnitte aller Mannschaften sind, desto weniger stark wird der historische Durchschnitt von Mannschaft i in der optimalen Torprognosegleichung berücksichtigt und desto stärker wird der Durchschnitt aller Durchschnitte berücksichtigt (der aufgrund der geringen Varianz auch sehr treffsicher geschätzt werden kann). Desto näher liegt c bei null.

Exkurs II für interessierte Leser: das Steinsche Schrumpfparadoxon

Jetzt ist es Zeit, das Stein's Paradoxon zu skizzieren: http://www.naftaliharris.com/blog/steinviz/

Wenn mehrere Mittelwerte gleichzeitig geschätzt werden sollen und die durchschnittliche quadratische Abweichung der Prognose vom Ergebnis im Erwartungswert minimiert werden soll, ist die optimale Prognose auch dann nicht das arithmetische Mittel, wenn gegen etwas völlig Schwachsinniges (z.B. die Anzahl der Legosteine, die in Zürich in Kabadosen aufbewahrt werden oder die Anzahl verkaufter Döner im Aargau) geschrumpft wird. Wenn also gleichzeitig die Fussballergebnisse und die Anzahl der Döner prognostiziert werden soll, und das Ziel ist, den Erwartungswert der mittleren quadratischen Abweichung zu minimieren, ist die optimale Schätzung der Fussballergebnisse ein gewichteter Durchschnitt zwischen den durchschnittlichen Fussballergebnissen und der durchschnittlichen Anzahl verkaufter Döner im Aargau einerseits und die optimale Schätzung der verkauften Döner im Aargau ein gewichteter Durchschnitt zwischen den durchschnittlich verkauften Dönern und den Fussballergebnissen andererseits. Das heisst mit anderen Worten, dass jeder zusätzlich verkaufte Döner im Aargau einen marginalen Einfluss auf die optimale Fussballprognose hat, wenn das Ziel ist, den Erwartungswert der durchschnittlichen quadratischen Abweichung der gemeinsamen Prognose zu minimieren. Ist das nicht absurd? Einige Mathematiker und Statistiker vergleichen denn auch das Steinsche Schrumpfparadoxon mit der Mondlandung und erinnern sich genau, was sie damals getan haben…