FUSSBALL-BUNDESLIGA UND STOCHASTIK-UNTERRICHT

H.K. STRICK

1.Einführung

In der Spielzeit 1980/81 fielen in der Fußball-Bundesliga insgesamt 1039 Tore; das sind (bei 306 Spielen insgesamt) im Mittel pro Spiel 3,395 Tore. 628 Tore wurden von den Heimmannschaften erzielt, 411 von den Gastmannschaften (Mittelwerte 2,052 bzw. 1,343 Tore pro Spiel).

Lassen sich hieraus Aussagen über die Resultate einzelner Spiele gewinnen ?

Mit diesen Informationen durch einen Fußball-begeisterten Schüler versehen wurde in einem Stochastik-Leistungskurs der Einstieg in die POISSON-Verteilung versucht.
Den Schülern war bis dahin im wesentlichen nur die Binomialverteilung als Verteilung bekannt. Es lag für die Schüler also nahe, einen Binomialansatz zu versuchen.

Der Begriff des Erfolgs bei BERNOULLI-Versuchen läßt sich unschwer mit dem Torerfolg einer Mannschaft identifizieren. Was aber sind bei einem Fußballspiel die Mißerfolge, was die Stufen des BERNOULLI-Versuchs ? Was schließlich ist die Erfolgswahrscheinlichkeit ?
Nach einigen Versuchen (etwa: Anzahl der Stufen = Anzahl der Torschüsse) kommt die Idee auf, jede Minute des 90minütigen Fußballspiels als eine Stufe anzusehen. 'Erfolge' sind dann die Minuten, in denen ein Tor fällt, 'Mißerfolge' die Minuten, in denen kein Tor fällt.
Als Erfolgswahrscheinlichkeit ergibt sich aus dem o.a. Mittelwert: $p = \frac{3,395}{90} = 0,0377$ (0,0228 bzw. 0,0149 für Torerfolge der Heim- bzw. Gastmannschaft).

Die Wahrscheinlichkeit für ein Spiel mit 0 Toren ist demnach:

$$P(X=0) = \binom{90}{0}\cdot 0,0377^0\cdot 0,9623^{90} = 0,0315$$

... mit 1 Tor:

$$P(X=1) = \binom{90}{1}\cdot 0,0377^1\cdot 0,9623^{89} = 0,1111$$

... mit 2 Toren:

$$P(X=2) = \binom{90}{2}\cdot 0,0377^2\cdot 0,9623^{88} = 0,1937 \quad \text{usw.}$$

Ein Fußballspiel als 90stufigen BERNOULLI-Versuch zu interpretieren, erscheint nicht erst fragwürdig, wenn man bedenkt, daß es Spiele gibt, bei denen innerhalb einer Minute auch 2 Tore fallen. Die Interpretation eines Fußballspiels als $90 \cdot 60 = 5400$-stufigen BERNOULLI-Versuch führt allerdings zu etwa gleichen Wahrscheinlichkeiten für Spiele mit 0,1,2,.. Toren.

Es ist: $P(X=0) = 0{,}0335$
$P(X=1) = 0{,}1139$
$P(X=2) = 0{,}1935$ usw.

2.POISSON-Verteilung als Näherung

Die Erhöhung des Stichprobenumfangs brachte bei vorgegebenem Mittelwert eine weitere Verkleinerung der Erfolgswahrscheinlichkeit p mit sich. Es erscheint daher sinnvoll, für große Stichprobenumfänge n und kleine Erfolgswahrscheinlichkeiten p die POISSON-Verteilung als Näherungsverteilung einzuführen.

Aus $P(X=0) = \binom{n}{0} \cdot p^0 \cdot (1-p)^n = \left(1 - \frac{\mu}{n}\right)^n \approx e^{-\mu}$ und

$$P(X=k) = \frac{(n-k+1)p}{k(1-p)} \cdot P(X=k-1) \approx \frac{\mu}{k} \cdot P(X=k-1) \text{ für } k=1,2,3,\ldots$$

gewinnt man die Formel:

$$P(X=k) \approx \frac{\mu^k}{k!} e^{-\mu} \qquad \text{(vgl. Lehrbücher zur Stochastik).}$$

3.Anpassung durch die POISSON-Verteilung

Mit den Werten $\mu=3{,}395$ bzw. $\mu_H=2{,}052$ bzw. $\mu_G=1{,}343$ erhält man die folgenden Wahrscheinlichkeiten für Spiele mit k Toren :

k	Wahrscheinlichkeit für k Tore		
	insgesamt	.. der Heimmannschaft	.. der Gastmannschaft
0	0,034	0,128	0,261
1	0,114	0,264	0,351
2	0,193	0,270	0,235
3	0,219	0,185	0,105
4	0,186	0,095	0,035
5	0,126	0,039	0,010
6	0,071	(≥6) 0,019	(≥6) 0,003
7	0,035		
≥8	0,022		

Tab. 3.1: Wahrscheinlichkeitsverteilung: Anzahl der Spiele mit k Toren

Bei 306 Spielen ergeben sich aus der Tabelle 3.1 die folgenden Erwartungswerte für die Anzahl der Spiele in der Spielzeit mit k Toren:

k	Erwartungswert der Anzahl der Spiele mit k Toren		
	insgesamt	.. der Heimmannschaft	.. der Gastmannschaft
0	10,3	39,3	79,9
1	34,8	80,7	107,3
2	59,1	82,8	72,0
3	66,9	56,6	32,3
4	56,8	29,1	10,8
5	38,6	11,9	2,9
6	21,6	(≥6) 5,6	(≥6) 0,8
7	10,6		
≥8	6,1		

Tab. 3.2: Erwartungswerte der Anzahl der Spiele mit k Toren

Die Anpassung der tatsächlichen Verteilung der Anzahl der Spiele mit 0,1,2,... Toren an die aufgrund des POISSON-Ansatzes ist erstaunlich:

k	Empirische Verteilung: Anzahl der Spiele mit k Toren		
	insgesamt	.. der Heimmannschaft	.. der Gastmannschaft
0	13	44	79
1	29	74	107
2	60	87	71
3	65	51	33
4	61	31	11
5	39	14	2
6	26	(≥6) 5	(≥6) 1
7	6		
≥8	7		

Tab. 3.3: Empirische Verteilung: Anzahl der Spiele mit k Toren

Man beachte, daß die 'theoretische' Verteilung allein aus der Information über die Anzahl der insgesamt gefallenen Tore berechnet wurde.

4. Faltung der POISSON-Verteilung

Welcher Zusammenhang besteht nun zwischen der (theoretischen) Verteilung für die Anzahl der Tore insgesamt und den Verteilungen für die Anzahl der Tore der Heim- bzw. Gastmannschaft ?

Es müßte gelten:

$$P(X=0) = P(X_H=0 \wedge X_G=0) = P(X_H=0) \cdot P(X_G=0),$$

$$P(X=1) = P(X_H=0 \wedge X_G=1 \quad \vee \quad X_H=1 \wedge X_G=0)$$

$$= P(X_H=0) \cdot P(X_G=1) + P(X_H=1) \cdot P(X_G=0),$$

$$P(X=2) = P(X_H=0) \cdot P(X_G=2) + P(X_H=1) \cdot P(X_G=1) + P(X_H=2) \cdot P(X_G=0), \text{ usw.}$$

Ein numerischer Vergleich zeigt die Übereinstimmung.

In der Tat gilt allgemein:

$$P(X_H=0) \cdot P(X_G=k) + P(X_H=1) \cdot P(X_G=k-1) + \ldots + P(X_H=k) \cdot P(X_G=0)$$

$$= \sum_{i=0}^{k} P(X_H=i) \cdot P(X_G=k-i)$$

$$= \sum_{i=0}^{k} e^{-\mu_H} \cdot \frac{\mu_H^{\,i}}{i!} \cdot e^{-\mu_G} \cdot \frac{\mu_G^{\,k-i}}{(k-i)!}$$

$$= e^{-(\mu_H+\mu_G)} \sum_{i=0}^{k} \frac{\mu_H^{\,i} \, \mu_G^{\,k-i}}{i! \, (k-i)!}$$

$$= \frac{1}{k!} \, e^{-(\mu_H+\mu_G)} \cdot \sum_{i=0}^{k} \binom{k}{i} \mu_H^{\,i} \, \mu_G^{\,k-i}$$

$$= \frac{1}{k!} \, e^{-(\mu_H+\mu_G)} \cdot (\mu_H + \mu_G)^k$$

Andererseits ist dies gerade gleich

$$P(X=k) = e^{-\mu} \cdot \frac{\mu^k}{k!} \quad \text{,da} \quad \mu = \mu_H + \mu_G .$$

5. Aussagen über einzelne Spielergebnisse

Von der Erfahrung aus Abschnitt 4 bestärkt, wagen wir weitere Aussagen:

Spiele mit 1 Tor können nur 1:0 oder 0:1 enden. Die Wahrscheinlichkeit für das Resultat '1:0' ist – gemäß POISSON-Ansatz

$$P('1:0') \approx 0{,}264 \; 0{,}261 = 0{,}0689 \; .$$

Bei 306 Spielen bedeutet dies einen Erwartungswert von 21,1 Spielen. Entsprechend ergibt sich für das Resultat '0:1':

$$P('0:1') \approx 0{,}128 \cdot 0{,}351 = 0{,}0449$$

oder ein Erwartungswert von 13,7 Spielen. Die Summe der beiden Erwartungswerte ergibt gemäß Abschnitt 4 wieder insgesamt 34,8 Spiele mit 1 Tor.

In den nachfolgenden Tabellen 5.1 und 5.2 sind entsprechende Rechenergebnisse für andere Spielresultate aufgeführt sowie zum Vergleich die empirische Verteilung der Resultate aus der Spielzeit 80/81:

Anzahl der Tore der Heimmannschaft \ Anzahl der Tore der Gastmannschaft	0	1	2	3	4	5
0	10,3	13,7	9,2	4,1	1,4	0,4
1	21,1	28,4	19,0	8,5	2,8	0,8
2	21,6	29,0	19,4	8,7	2,9	0,8
3	14,8	19,9	13,3	5,9	2,0	0,6
4	7,6	10,2	6,8	3,1	1,0	0,3
5	3,1	4,2	2,8	1,3	0,4	0,1

Tab. 5.1: Theoretische Verteilung der Spielresultate

Anzahl der Tore der Heimmannschaft \ Anzahl der Tore der Gastmannschaft	0	1	2	3	4	5	6
0	13	8	9	10	3	0	1
1	21	31	9	8	4	1	0
2	20	33	27	4	2	1	0
3	13	15	14	7	2	0	0
4	8	13	8	2	0	0	0
5	4	7	1	2	0	0	0
6	0	0	1	0	0	0	0
7	0	2	2	0	0	0	0

Tab. 5.2: Empirische Verteilung der Spielresultate 80/81

War man in Abschnitt 3 über die Güte der Anpassung erstaunt, dann wird man sich über die zum Teil erheblichen Abweichungen wundern.

Wie läßt sich dies erklären ?

In den Rechnungen von Abschnitt 4 wurde zum Beispiel geschlossen:

$$P(X=1) = P(X_H=1) \cdot P(X_G=0) + P(X_H=0) \cdot P(X_G=1)$$

$$\overset{!}{=} P(X_H=1 \wedge X_G=0 \quad \vee \quad X_H=0 \wedge X_G=1) \; .$$

dh. die Wahrscheinlichkeiten wurden multipliziert, ohne daß geprüft wurde, ob die entsprechenden Ereignisse unabhängig sind !
Aus dem Vergleich der Tabellen 5.1 und 5.2 ergibt sich, daß die Voraussetzung der Unabhängigkeit der Torerfolge der Heim- bzw. Gastmannschaft nicht zulässig ist.
An einem Beispiel sei die Interpretation der Abweichungen der Werte gewagt: Es ist plausibel, daß ein Spielstand von 1:2 für die Gastmannschaft die Heimmannschaft zu erhöhten Anstrengungen veranlaßt, die dann über die theoretische Erwartung hinaus zum Spielstand 2:2 führt. (Die Gesamtzahl der Spiele mit 1:2 und 2:2 ist im theoretischen wie im empirischen Falle etwa gleich.)

6.Aussagen über die Gesamtzahl der Siege und Unentschieden

Nach den Überlegungen von Abschnitt 5 ist das Ergebnis der folgenden Untersuchung nicht mehr verwunderlich. Aus Tabelle 5.1 könnte man Erwartungswerte für die Anzahl der Heimsiege, Unentschieden und Heimniederlagen gewinnen:
Für die Unentschieden ergibt sich aus den Diagonalwerten ein Erwartungswert von ca. 65. Für die Heimniederlagen ist der Erwartungswert ca. 76 (Summe aller Erwartungswerte oberhalb der Diagonale der Unentschieden). Hieraus ergibt sich ein Erwartungswert von ca. 165 Heimsiegen. Die empirischen Werte der Spielzeit 80/81 waren: 164 Heimsiege (!), 80 Unentschieden und 62 Heimniederlagen.

7.Untersuchungen zur Abschlußtabelle der Spielzeit

Legt man nun die Gesamtzahl der Heimsiege, Unentschieden und Heimniederlagen zugrunde, dann kann man hiermit versuchen, einen Bezug zur Bundesliga-Abschlußtabelle herzustellen.
Die Anteile von 53,6% Heimsiegen, 26,1% Unentschieden und 20,3% Heimniederlagen entsprechen in etwa auch den Anteilen früherer Spielzeiten. Setzt man diese Anteile als Wahrscheinlichkeiten für Heimsiege, Unentschieden und Heimniederlagen einer Mannschaft an, dann läßt sich die Abfolge von 17 Heimspielen als Zufallsversuch mit 17 Stufen und den Wahrscheinlichkeiten von 0,536 für den Gewinn von 2 Punkten bzw. von 0,261 für den Gewinn von 1 Punkt für jede Stufe interpretieren. Entsprechend hat man bei den Auswärtsspielen eine Wahrscheinlichkeit von 20,3% für den Gewinn von 2 Punkten und ebenfalls von 26,1% für den Gewinn von 1 Punkt.

Aus der Wahrscheinlichkeit für k_1 Heimsiege, k_2 Unentschieden und k_3 Heimniederlagen

$$P(k_1,k_2,k_3) = \binom{17}{k_1,k_2,k_3} \cdot 0{,}536^{k_1} \cdot 0{,}261^{k_2} \cdot 0{,}203^{k_3} \quad \text{bzw.}$$

$$P(l_1,l_2,l_3) = \binom{17}{l_1,l_2,l_3} \cdot 0{,}203^{l_1} \cdot 0{,}261^{l_2} \cdot 0{,}536^{l_3}$$

für l_1 Siege, l_2 Unentschieden und l_3 Niederlagen bei Auswärtsspielen erhält man die Wahrscheinlichkeitsverteilung für den Gewinn von $k = 2k_1+k_2$ Punkten aus Heimspielen bzw. $l = 2l_1+l_2$ Punkten aus Auswärtsspielen und hieraus wiederum die Wahrscheinlichkeitsverteilung für die Gesamtpunktzahl der Spielzeit.

Die hierzu notwendigen Rechnungen lassen sich mit dem Computer schnell durchführen. Die Tabellen 7.1, 7.2 und 7.3 enthalten die beschriebenen Wahrscheinlichkeitsverteilungen für Punkte aus Heim-, Auswärtsspielen, sowie die Gesamtverteilung.

Der Deutsche Meister der Spielzeit 1980/81 erreichte 53 Punkte, was der theoretischen Tabelle nach nur selten vorkommt. (Auch die 45 Punkte, die als bisherige Minimalpunktzahl in der Spielzeit 75/76 für den Gewinn der Deutschen Meisterschaft genügten, liegen in der Tabelle 7.3 im Bereich der 'seltenen' Ereignisse.) Entsprechend selten wäre nach der theoretischen Tabelle auch die Tatsache, daß eine Mannschaft mit 26 Punkten noch den Klassenerhalt erreicht, wie dies in der Spielzeit 80/81 der Fall war (3 von 18 Mannschaften steigen ab - das sind 16,7%).

Die starken Abweichungen zwischen theoretischer Verteilung und Bundesliga-Abschlußtabelle erklären sich aus der unterschiedlichen Spielstärke der Mannschaften - der Ansatz für die Rechnung, 'Wahrscheinlichkeiten' für Heimsiege, Unentschieden und Heimniederlagen für alle Mannschaften, erweist sich als nicht zulässig. Das benutzte Modell stellt sich als unzulänglich heraus.

8.Abschließende Bemerkung

In diesem Aufsatz wurde versucht, mit Hilfe von mathematischen M o d e l l e n die Realität zu beschreiben. Die Schwierigkeiten, die sich in den Abschnitten 5, 6 und 7 ergaben, bieten eine gute Möglichkeit, die Problematik der Modellkonstruktion im Unterricht zu thematisieren.

Auszug aus dem Computer-Ausdruck:

Tabelle 7.1: Wahrscheinlichkeitsverteilung für Punkte aus Heimspielen

Punktzahl	Wahrscheinlichkeit	kumulierte Wahrscheinlichkeit
0 : 11	$\} < 0.5\ 10^{-4}$	< 0,0005
12	0,001	0,0014
13	0,002	0,0036
14	0,005	0,008
15	0,009	0,017
16	0,016	0,033
17	0,028	0,061
18	0,043	0,105
19	0,063	0,167
20	0,084	0,251
21	0,103	0,353
22	0,116	0,470
23	0,121	0,591
24	0,115	0,705
25	0,099	0,804
26	0,077	0,882
27	0,054	0,936
28	0,033	0,969
29	0,018	0,987
30	0,008	0,9956
31	0,003	0,9988
32	0,001	0,9998
33 34	$< 0{,}5\cdot 10^{-4}$	1,0000

Tabelle 7.2: Wahrscheinlichkeitsverteilung für Punkte aus Auswärtsspielen

ergibt sich aus Tabelle 7.1 bei Betrachtung der Verlustpunkte (Gewinnpunkte = 34 - Verlustpunkte).

Tabelle 7.3: Gesamtverteilung (Heim- und Auswärtsspiele)

Punktzahl	Wahrscheinlichkeit	kumulierte Wahrscheinlichkeit
0 : 19	$\} < 0{,}5\cdot 10^{-3}$	< 0,0007
20	0,001	0,0016
21	0,002	0,0032
22	0,003	0,006
23	0,005	0,011
24	0,008	0,020
25	0,013	0,033
26	0,019	0,052
27	0,028	0,080
28	0,037	0,117
29	0,048	0,165
30	0,059	0,225
31	0,070	0,295
32	0,078	0,373
33	0,084	0,457
34	0,086	0,543
35	0,084	0,627
36	0,078	0,705
37	0,070	0,775
38	0,059	0,835
39	0,048	0,883
40	0,037	0,920
41	0,028	0,948
42	0,019	0,967
43	0,013	0,980
44	0,008	0,989
45	0,005	0,994
46	0,003	0,997
47	0,002	0,9984
48	0,001	0,9992
49 : 68	$\} < 0{,}5\cdot 10^{-3}$	> 0,9996

```
10 REM *** THEORETISCHE BUNDESLIGA-TABELLE ***
20 DIM H(34),G(34),W(68)
30 PRINT "EINGABE WAHRSCHEINL.F.HEIMSIEG,UNENTSCHIEDEN;ANZ.D.SPIELE"
40 INPUT P,Q,A
50 REM *** BERECHNUNG D.WAHRSCHEINL.MIT POLYNOMIALANSATZ ***
60 F=1 \ F1=1 \ R=1-P-Q
70 FOR I=0 TO A
80 FOR J=0 TO A-I
90 K=A-I-J \ L=2*I+J
100 H(L)=F*P^I*Q^J*R^K+H(L)
110 F=F*K/(J+1)
120 NEXT J
130 F1=F1*(A-I)/(I+1)
140 F=F1
150 NEXT I
160 REM *** EINZELVERTEILUNG ***
170 PRINT "EINZELVERTEILUNG" \ PRINT  \ PRINT
180 PRINT ,"HEIMSPIELE",,"AUSWAERTSSPIELE"
190 PRINT "PUNKTZAHL","W-VERTEILUNG","KUMUL.W-VERT.","W-VERTEILUNG",
200 PRINT "KUMUL.W-VERT."
210 PRINT
220 S=0 \ T=0
230 FOR L=0 TO 2*A
240 G(L)=H(2*A-L)
250 S=S+H(L)
260 T=T+G(L)
270 PRINT "    ";L,H(L),S,G(L),T
280 NEXT L
290 REM *** BERECHNUNG GESAMTVERTEILUNG ***
300 PRINT  \ PRINT  \ PRINT  \ PRINT "GESAMTVERTEILUNG" \ PRINT  \ PRINT
310 PRINT "PUNKTZAHL","W-VERTEILUNG","KUMUL.W-VERTEILUNG"
320 FOR L=0 TO 2*A
330 FOR M=0 TO 2*A
340 N=L+M
350 W(N)=W(N)+G(L)*H(M)
360 NEXT M
370 NEXT L
380 REM *** AUSDRUCK GESAMTVERTEILUNG ***
390 S=0
400 FOR N=0 TO 4*A
410 S=S+W(N)
420 PRINT "    ";N,W(N),S
430 NEXT N
440 END
```