

Georg Berschneider und René L. Schilling

## Die Poisson-Verteilung, Fußballtore und das Gesetz der kleinen Zahlen

Vor dem 34. und letzten Spieltag 2017/18 der 2. Fußball-Bundesliga war für die SG »Dynamo Dresden« noch alles möglich, vom direkten Abstieg in die Drittklassigkeit bis zu einem guten Platz in der Tabellenmitte. Aufgrund des hohen öffentlichen Interesses erhielten die Autoren

Platz	Verein	Spiele	Tore	Punkte	$\varnothing$ Tore/Spiel ( $\lambda$ )
13	Dynamo Dresden	33	42:51	41	$\frac{42}{33} = 1,2727$
14	SV Darmstadt 98	33	46:45	40	$\frac{46}{33} = 1,3939$
15	Erzgebirge Aue	33	35:48	40	$\frac{35}{33} = 1,0606$
16	Eintracht Braunschweig	33	35:37	39	$\frac{35}{33} = 1,0606$
17	SpVgg Greuther Fürth	33	36:47	39	$\frac{36}{33} = 1,0909$

**Tabelle 1.** Ausschnitt der Tabelle der 2. Bundesliga nach dem 33. Spieltag, 2017/2018.

von der Sportredaktion der »Sächsischen Zeitung« die Anfrage, ob sie die Wahrscheinlichkeiten für die verschiedenen Tabellenendstände berechnen könnten. Wegen der engen Punktabstände spielte für die Prognose die Tordifferenz eine zentrale Rolle: basierend auf unserem Modell war in den Simulationen der Platz 14 mit 38 % am wahrscheinlichsten, während sowohl der direkte Abstieg als auch ein Platz im Mittelfeld nahezu ausgeschlossen werden konnten.

Für unsere Einschätzung simulierten wir alle ergebnisrelevanten Partien des letzten Spieltags, wobei wir die Wahrscheinlichkeit für die Tore einer Mannschaft mit Hilfe der *Poissonverteilung* ermittelten

$$P(\text{»Mannschaft } M \text{ erzielt } k \text{ Tore«}) = \text{Poi}(\lambda; k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots \quad (1)$$

Hier bedeuten  $k! = 1 \cdot 2 \cdot 3 \cdots k$ ,  $0! = 1$  und  $\lambda = \lambda_M$  ist die durchschnittliche Zahl der Tore pro Spiel, welche die Mannschaft  $M$  über die Saison hinweg erzielt hat (weitere Details in Abschnitt 4).

Mathematiker und (Natur-)Wissenschaftler verwenden die Formel (1) in der Bevölkerungsstatistik, zur Modellierung von Zerfallshäufigkeiten bei schwach radioaktiven Präparaten (Bateman 1910), der Biologie (Bestimmung der Anzahl der Hefezellen in der Maische der Guinness-Brauerei, Student 1907), der Modellierung von Telefonnetzwerken (Erlang 1909), der Personen- und Schadensversicherung (Gumbel 1937), von Daten geschichtlicher Ereignisse (Beginn von Kriegen, Richardson 1944) oder der Frage, ob die »fliegenden Bomben« (d. i. Hitlers Vergeltungswaffe V1) im Stadtgebiet von London in Clustern einschlugen (Clarke 1948).<sup>1</sup>

Was haben diese Beispiele mit Fußballtoren zu tun? Ein Blick auf Tabelle 1 zeigt, dass Tore »selten« sind, und das nicht nur im Tabellenkeller, sondern weltweit, vgl. Abschnitt 4. Der russisch-deutsche Mathematiker von Bortkewitsch (Bortkewitsch 1898) wies im Jahr 1898 darauf hin, dass sich die Poisson-Verteilung für die Modellierung seltener Ereignisse eignet. Er prägte den Begriff vom *Gesetz der kleinen Zahlen*, das er an etwas morbiden Beispielen – Anwendungen auf einige Daten der *Selbstmord- und Unfall-Statistik* – eingehend illustrierte; legendär ist sein Beispiel der *durch Schlag eines Pferdes im preußischen Heere Getöteten* (Tabelle 2).

**Tabelle 2.** Bortkewitsch (Bortkewitsch 1898, Kapitel II, §12). Todesfälle bei 10 Corps der preußischen Armee im Zeitraum von 20 Jahren 1875–1894. Spalte 2 der Tabelle verzeichnet die Zahl der »Corps-Jahre« (insgesamt  $20 \cdot 10 = 200$  Jahre), in denen 0,1,2,... Todesfälle vorkamen, die dritte Spalte enthält die Poisson-Approximation mit  $\lambda = 0,61$ .

Jahres- ergebnis	Zahl der Fälle, in denen das nebenstehende Jahresergebnis	
	vorgekommen ist	zu erwarten war
0	109	108,7
1	65	66,3
2	22	20,2
3	3	4,1
4	1	0,6
5 u. mehr	—	0,1

Doch wollen wir den Ereignissen nicht vorgreifen und zunächst einige mathematische Grundlagen diskutieren.

## 1 Approximation der Binomialverteilung

Siméon Denis Poisson (1781–1840) war ein französischer Mathematiker, der zu vielen Gebieten der Mathematik und Physik wichtige Beiträge leistete. Zur Wahrscheinlichkeitstheorie gelangte Poisson erst relativ spät, als er im Jahr 1829 statistische Untersuchungen zur Verteilung der Geschlechter bei Neugeborenen anstellte. Hier und später in Poissons Lehrbuch zur Wahrscheinlichkeitstheorie begegnen wir einer diskreten Wahrscheinlichkeitsverteilung, die man heute als einen wichtigen Beitrag Poissons<sup>2</sup> zur Stochastik ansieht.

Poisson war daran interessiert, gute Approximationen für die Binomialverteilung zu finden (Poisson 1841, §§73–77, S. 155–162). Eine Zufallsvariable  $X_n$  heißt *binomialverteilt* mit Erfolgswahrscheinlichkeit  $p$ , wenn gilt

$$\mathbb{P}(X_n = k) = B(n, p; k) = \binom{n}{k} p^k q^{n-k}, \quad q = 1 - p, \quad k = 0, 1, 2, \dots, n; \quad (2)$$

$X_n$  wird üblicherweise als Zahl der Erfolge in einem  $n$ -mal unabhängig ausgeführten Bernoulli-Experiment (»Münzwurf mit einer (ggf. gezinkten) Münze«) interpretiert. Der Erwartungswert ist  $\mathbb{E}X_n = n\mathbb{E}X_1 = np$  und die Varianz  $\mathbb{V}X_n = n\mathbb{V}X_1 = npq$ . Die exakte Berechnung der Wahrscheinlichkeit (2) ist für große Werte von  $n$  unangenehm<sup>3</sup> und daher sind gute Näherungsformeln von großer Bedeutung.

Weil  $X_n$  als Summe von  $n$  unabhängigen Bernoulli-verteilten Zufallsvariablen gesehen werden kann, ist es naheliegend, den zentralen Grenzwertsatz von de Moivre-Laplace zu verwenden und  $X_n$  durch eine Normalverteilung zu approximieren. Wir haben für  $n \rightarrow \infty$

$$\mathbb{P}(X_n \leq k) = \mathbb{P}\left(\frac{X_n - np}{\sqrt{npq}} \leq \frac{k - np}{\sqrt{npq}}\right) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(k-np)/\sqrt{npq}} e^{-x^2/2} dx, \quad (3)$$

gute Ergebnisse werden aber nur für  $n \geq 10$  und  $npq > 6$  erreicht – d. h. die Approximation ist nur für relativ »große« Werte  $pq$  interessant. Weil  $p \mapsto pq = p(1-p)$  eine nach unten geöffnete Parabel mit Maximum  $\frac{1}{4}$  an der Stelle  $p = \frac{1}{2}$  ist, heißt das, dass wir für  $p$  nahe Null oder Eins eine andere Asymptotik verwenden sollten.

Die Formel (3) geht auf Laplace (Laplace 1820, Buch II, Kapitel III, §16) zurück<sup>4</sup>, Poissons Lehrbuch (Poisson 1841, §§73ff., S. 155ff.) enthält einen von Laplace unabhängigen Beweis.

Für kleine Werte von  $p$  verwendet Poisson (Poisson 1841, §81, S. 171f.) folgende Näherung:

$$\mathbb{P}(X_n \leq k) \approx \sum_{i=0}^k \frac{\lambda^i}{i!} e^{-\lambda} \quad \text{für } n \rightarrow \infty \text{ und } np_n \rightarrow \lambda. \quad (4)$$

Die Bedingung  $np_n \rightarrow \lambda$  besagt insbesondere, dass  $p = p_n$  für  $n \rightarrow \infty$  klein ist. Interessanterweise wird die Formel für die Zähldichte der *Poissonverteilung*

$$\mathbb{P}(X_n = k) = \mathbb{P}(X_n \leq k) - \mathbb{P}(X_n \leq k-1) \approx \frac{\lambda^k}{k!} e^{-\lambda} \quad (5)$$

an keiner Stelle in Poissons Schriften explizit erwähnt.

## Die Poissonverteilung

Eine diskrete Zufallsvariable  $Z$  nennt man *Poisson-verteilt* zum Parameter  $\lambda > 0$ , wenn

$$\mathbb{P}(Z = k) = \text{Poi}(\lambda; k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, 3, \dots, \quad (6)$$

( $k! = 1 \cdot 2 \cdot \dots \cdot k$ ,  $0! = 1$ ) gilt. Mit Hilfe der Exponentialreihe

$$e^\lambda = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \quad (7)$$

sieht man sehr schnell, dass  $\sum_{k=0}^{\infty} \text{Poi}(\lambda; k) = 1$  gilt, und dass sowohl der Erwartungswert  $\mathbb{E}Z = \sum_{k=0}^{\infty} k \cdot \text{Poi}(\lambda; k)$  als auch die Varianz  $\mathbb{V}Z = \sum_{k=0}^{\infty} (k - \mathbb{E}Z)^2 \cdot \text{Poi}(\lambda; k)$  den Wert  $\lambda$  haben. Aus der Beziehung

$$\text{Poi}(\lambda; k) = \frac{\lambda}{k} \text{Poi}(\lambda; k-1), \quad k = 1, 2, \dots \quad (8)$$

folgt, dass  $\text{Poi}(\lambda; k)$  so lange wächst, bis  $k > \lambda$  wird, d. h.  $\text{Poi}(\lambda; k)$  erreicht sein Maximum im Intervall  $[\lambda - 1, \lambda]$ ; für ganzzahliges  $\lambda$  sind  $\text{Poi}(\lambda; \lambda - 1)$  und  $\text{Poi}(\lambda; \lambda)$  maximal.

Später werden wir noch benötigen, dass die Summe  $Z + W$  von zwei unabhängigen Poisson-verteilten Zufallsvariablen  $Z$  und  $W$  mit Parametern  $\lambda$  und  $\mu$  wiederum Poisson-verteilt ist mit Parameter  $\lambda + \mu$ .<sup>5</sup>

## Poissons Approximation der Binomialverteilung

Nach heutigen Maßstäben ist Poissons Herleitung der Beziehung (4) nicht mathematisch rigoros, trotzdem erklärt sie gut die Bedeutung und den Charakter der Poissonverteilung. Im folgenden fassen wir Poissons Argumentation zusammen.

Wir betrachten  $n = m + k$  unabhängige Wiederholungen eines Bernoulli-Experiments mit Erfolgswahrscheinlichkeit  $p$ . Weil die Anzahl der Erfolge  $X_n$  binomialverteilt ist, gilt

$$\sum_{i=0}^k \binom{n}{i} p^i q^{n-i} = \mathbb{P}(X_n \leq k) = \sum_{i=0}^k \binom{m+i-1}{i} q^{m-1} \cdot p^i \cdot q. \quad (9)$$

Die erste Gleichheit in (9) ist die Summe von Binomialwahrscheinlichkeiten. Die zweite Gleichheit ist nicht offensichtlich, aber man kann sie mit einer alternativen Aufteilung des Ereignisses »höchstens  $k$  Erfolge bei  $n$  Versuchen«  $\{X_n \leq k\}$  folgendermaßen zeigen: wir stellen uns die

Ergebnisse der  $n$  Experimente hintereinander aufgereiht vor. Der Ausdruck unter der Summe gibt die Wahrscheinlichkeit an, auf den ersten  $m+i-1$  Plätzen  $m-1$  Misserfolge und  $i$  Erfolge, auf Platz  $m+i$  einen Misserfolg, und auf den restlichen Plätzen  $m+i+1, \dots, n$  ein beliebiges Ergebnis zu erhalten. Weil dabei die Anzahl der Misserfolge über  $m$  hinaus unerheblich ist, tritt auf der rechten Seite die Wahrscheinlichkeit  $q$  eines Misserfolgs mit dem festen Exponenten  $m$  auf. Der Faktor  $\binom{m+i-1}{i}$  zählt, auf wieviele Arten wir  $i$  Erfolge auf die ersten  $m+i-1$  Plätze verteilen können. Die Summation über  $i=0, \dots, k$  ergibt dann eine Aufteilung<sup>6</sup> des Ereignisses  $\{X_n \leq k\}$ .

Wir nehmen nun an, dass  $p = k/n$  klein ist. Weil dann  $q = 1 - p = m/n \approx 1$  gilt, haben wir  $m \approx n$  und  $k$  kann gegenüber  $n$  vernachlässigt werden. Poisson macht (Poisson 1841, §83, S. 171) den Ansatz

$$pm \approx pn = \lambda, \quad m(m+1)p^2 \approx \lambda^2, \quad m(m+1)(m+2)p^3 \approx \lambda^3, \quad \dots$$

$$q = 1 - \frac{\lambda}{n}, \quad q^m = \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} \approx e^{-\lambda}. \quad (10)$$

Allerdings ist die asymptotische Gleichheit für  $q^m$  nicht offensichtlich (siehe Lemma 1), was Poisson stillschweigend übergeht. Indem wir diese Ausdrücke in die zweite Formel von (9) einsetzen, folgt

$$\mathbb{P}(X_n \leq k) \approx \left(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots + \frac{\lambda^k}{k!}\right) e^{-\lambda} \quad (11)$$

für die Wahrscheinlichkeit, dass ein Ereignis, dessen Wahrscheinlichkeit bei jedem Versuche der sehr kleine Bruch  $\frac{\lambda}{n}$  ist, nicht mehr, als  $k$  mal in einer sehr großen Anzahl  $n$  von Versuchen stattfindet (Poisson 1841, S. 171).

Dieselbe Approximation wurde in einem etwas anderen Zusammenhang bereits im Jahr 1718 von de Moivre (de Moivre 1756, S. 45, Problem V) verwendet.

## Der heute übliche Beweis der Poisson-Approximation

Poissons Interesse galt den kumulativen Formeln (3) und (4), was die Betrachtung der Summe (9) erklärt. Moderne Beweise zeigen direkt die Beziehung (5) und leiten daraus (4) her. Der wesentliche Beweisschritt ist, wie schon bei Poisson, die Asymptotik (10), die aber bewiesen werden muss, vgl. Lemma 1.

Ausgangspunkt ist wieder die Bemerkung, dass  $\lambda = np$  der Erwartungswert der Binomialverteilung  $B(n, p; \cdot)$  ist. Wenn  $\lambda$  konstant ist und  $n$  groß wird, muss notwendigerweise  $p = p_n = \lambda/n$  klein sein. Dann ist

$$B(n, p; k) = \frac{1}{k!} \cdot \underbrace{\frac{n}{n} \cdot \frac{n-1}{n} \cdot \dots \cdot \frac{n-k+1}{n}}_{\rightarrow 1} \lambda^k \left(1 - \frac{\lambda}{n}\right)^n \underbrace{\left(1 - \frac{\lambda}{n}\right)^{-k}}_{\rightarrow 1} \xrightarrow[n \rightarrow \infty]{k \text{ fest}} \frac{\lambda^k}{k!} e^{-\lambda}$$

und es folgt (5). In dieser Rechnung haben wir folgende Hilfsaussage verwendet.

**Lemma 1.** Wenn  $\lim_{n \rightarrow \infty} \lambda_n = \lambda$ , dann gilt  $\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda_n}{n}\right)^n = e^{-\lambda}$ .

*Beweis.* Es gilt

$$\left(1 - \frac{\lambda_n}{n}\right)^n = \exp \left[ n \ln \left(1 - \frac{\lambda_n}{n}\right) \right] = \exp \left[ -\lambda_n \cdot \frac{\ln \left(1 - \frac{\lambda_n}{n}\right)}{-\frac{\lambda_n}{n}} \right].$$

Nach Voraussetzung gilt  $\lambda_n \rightarrow \lambda$ , also  $x = x_n = -\lambda_n/n \rightarrow 0$ . Daher können wir den folgenden Grenzwert mit der Regel von de l'Hospital berechnen

$$\lim_{n \rightarrow \infty} \frac{\ln\left(1 - \frac{\lambda_n}{n}\right)}{-\frac{\lambda_n}{n}} = \lim_{x \rightarrow 0} \frac{\ln(1+x)}{x} = \lim_{x \rightarrow 0} \frac{\frac{1}{1+x}}{1} = 1,$$

insgesamt erhalten wir  $\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda_n}{n}\right)^n = e^{-\lambda}$ . □

## 2 Das Gesetz der kleinen Zahlen in der Schule

Wir wollen hier einige typische Rechenbeispiele angeben, wie die Poisson-Verteilung in der Schule verwendet werden kann.

**Beispiel 2** (Rosinenbrötchen). *Eine Bäckerei verkauft Rosinenbrötchen zu je 50 g. In 5 kg Teig werden 200 Rosinen gegeben, der Teig wird gründlich geknetet, in Brötchen zu 50 g aufgeteilt und gebacken. Wie groß ist die Wahrscheinlichkeit, dass ein zufällig herausgegriffenes Brötchen 0, 1, 2, 3, ... Rosinen enthält?*

*Lösung.* Aus 5 kg Teig erhalten wir 100 Brötchen zu je 50 g. Weil der Teig gut durchmischt ist, können wir annehmen, dass die Rosinen voneinander unabhängig in die Brötchen gelangen. Die Wahrscheinlichkeit, dass eine der 200 Rosinen sich in einem bestimmten Brötchen befindet ist  $\frac{1}{100}$ . Insgesamt haben wir 200 derartige »Versuche« und der Erwartungswert ist  $200 \cdot \frac{1}{100} = 2$  Rosinen je Brötchen. Wir erhalten mit der Poisson-Approximation für  $\lambda = 2$  und  $k = 0, 1, 2, \dots$

$$\mathbb{P}(\text{»}k\text{ Rosinen in einem zufällig ausgewählten Brötchen«}) \approx \frac{2^k}{k!} e^{-2}.$$

Die exakte Lösung ist durch die Binomialverteilung  $B(200, \frac{1}{100}; k)$  gegeben. Die folgende Tabelle zeigt, wie gut die Approximation ist.

$k$	0	1	2	3	4	5	6
Binomial	0,1340	0,2707	0,2720	0,1814	0,0902	0,0357	0,0117
Poisson	0,1353	0,2707	0,2707	0,1804	0,0902	0,0361	0,0120

**Beispiel 3** (Verbrauchertest). *Ein Verbrauchermagazin kauft bei einem Bäcker 100 Rosinenbrötchen und untersucht, ob die Brötchen tatsächlich Rosinen enthalten. In 37 Brötchen werden keine Rosinen gefunden. Kann man schätzen, wieviele Rosinen der Bäcker in den Teig gegeben hat?*

*Lösung.* Wenn  $m$  Rosinen im Teig waren, dann ist  $\lambda = m/100$  die durchschnittliche Zahl von Rosinen je Brötchen. Daher ist die Wahrscheinlichkeit in einem bestimmten Brötchen keine Rosine zu finden  $\text{Poi}(\lambda; 0) = e^{-\lambda}$ , und wir können  $m$  folgendermaßen schätzen:

$$37 = 100 \cdot e^{-m/100} \implies m = -100 \cdot \ln \frac{37}{100} \approx 99,4,$$

d. h. der Bäcker hat etwa 100 Rosinen für 100 Brötchen in den Teig gegeben.

**Beispiel 4** (Geburtstage). *An einer Schule gibt es 500 Schüler. Wie groß ist die Wahrscheinlichkeit dafür, dass genau  $k$  Schüler am 1. Weihnachtstag Geburtstag haben?*

*Lösung.* Der Einfachheit halber nehmen wir an, dass ein Jahr 365 Tage hat und dass die Geburtstage gleichmäßig über das Jahr verteilt sind.<sup>7</sup> Für jeden *fest vorab bestimmten* Tag erwarten wir  $\frac{500}{365} = 1,36986 \dots \approx 1,3699$  Geburtstage. Die exakte Verteilung ist  $B(500, \frac{1}{365}; k)$ , mit der Poisson-Approximation  $Poi(1,3699; k)$  erhalten wir

$k$	0	1	2	3	4	5	6
Binomial	0,2537	0,3484	0,2388	0,1089	0,0372	0,0101	0,0023
Poisson	0,2541	0,3481	0,2385	0,1089	0,0373	0,0102	0,0023

Das klassische Geburtstagsproblem lässt sich auch mit der Poisson-Approximation behandeln.

**Beispiel 5** (Klassisches Geburtstagsproblem). *Wie groß ist die Wahrscheinlichkeit, dass in einer Gruppe von  $n$  Personen mindestens 2 am selben Tag Geburtstag haben?*

*Lösung.* Der Einfachheit halber nehmen wir an, dass ein Jahr 365 Tage hat und dass die Geburtstage gleichmäßig über das Jahr verteilt sind.<sup>7</sup> Wir können mit  $n$  Personen  $\binom{n}{2} = \frac{1}{2}n(n-1)$  verschiedene Pärchen bilden, und im Durchschnitt haben  $\lambda = \binom{n}{2} \frac{1}{365}$  Paare am gleichen Tag im Jahr Geburtstag. Mit der Poisson-Approximation ist die Wahrscheinlichkeit für  $k$  Geburtstagspaare  $Poi(\lambda; k)$ . Weil wir die Wahrscheinlichkeit  $p$  suchen, dass *mindestens ein* solches Pärchen auftritt, erhalten wir mit der Gegenwahrscheinlichkeit

$$p = 1 - \mathbb{P}(\text{»kein Geburtstagspaar«}) = 1 - Poi\left(\binom{n}{2} \frac{1}{365}; 0\right) = 1 - \exp\left[-\frac{n(n-1)}{2 \cdot 365}\right].$$

Für  $n = 30$  ist  $p \approx 0,6963$ . Indem wir die Formel nach  $n$  auflösen, folgt

$$n(n-1) = -730 \cdot \ln(1-p) \quad \text{und} \quad n-1 \leq \sqrt{-730 \cdot \ln(1-p)} \leq n$$

(verwende  $(n-1)^2 \leq n(n-1) \leq n^2$ ); mit  $p = 1/2$  sehen wir, dass ab  $n \geq 22,49$  Personen die Wahrscheinlichkeit eines Mehrfachgeburtstags über 50 % liegt.

Schließlich geben wir noch einige Varianten dieser Beispiele an.

**Beispiel 6.** a) Eine Schreibkraft vertippt sich bei 1000 Anschlägen durchschnittlich einmal. Wie groß ist in einem 20-seitigen Manuskript die erwartete Zahl  $N_k$  der Normseiten (zu je 1500 Anschlägen) mit  $k = 0, 1, 2, 3, \dots$  Tippfehlern? Wie groß ist die erwartete Zahl  $N$  der Seiten mit mindestens einem Fehler? Antwort:  $\lambda = 1500/1000$  (durchschnittliche Fehlerzahl pro Seite) und damit  $N_k = 20 \cdot Poi(\lambda; k)$  und  $N = 20 - N_0$ .

b) Eine Fluggesellschaft weiß aus Erfahrung, dass 5 % aller gebuchten Passagiere den Flug nicht antreten (sog. *no-shows*). Daher wird ein Flugzeug mit 100 Sitzen systematisch überbucht, d. h. es werden 104 Tickets verkauft. Wie groß ist die Wahrscheinlichkeit, dass jeder Fluggast einen Sitzplatz erhält? Antwort:  $n = 104$ ,  $p = 0,05$ ,  $\lambda = np = 5,2$ ;  $\mathbb{P}(X \geq 104 - 100) = 1 - \mathbb{P}(X < 4) = 1 - Poi(\lambda; 0) - \dots - Poi(\lambda; 3)$ .

c) In  $100\text{cm}^3$  Lösung sind 200 Viren. Es werden 100 Ratten mit je  $1\text{cm}^3$  Lösung geimpft. Ein einziger Virus führt sicher zur Erkrankung des Tieres. Mit wievielen Krankheitsfällen  $N$  muss man rechnen? Antwort: das ist eine Spielart der Rosinenaufgabe.  $N = 100 \cdot (1 - Poi(2; 0))$  (Gegenwahrscheinlichkeit!).

### 3 Wie gut ist die Näherung wirklich?

Die Tabellen in den Beispielen 2 und 4 legen nahe, dass die Poisson-Approximation (5) für  $n \rightarrow \infty$  sehr schnell konvergiert und die Approximationsfehler über alle  $k = 1, \dots, n$  hinweg *summiert* klein bleiben. Tatsächlich kann man den Abstand zwischen den Verteilungen abschätzen, was zu einem neuen Beweis für (5) führt. Das folgende elegante Kopplungsargument geht auf (Hodges & Le Cam 1960) zurück, wir folgen der Darstellung in (Schilling 2018, Kapitel 8).

**Lemma 7.** *Es seien  $Y_1, \dots, Y_n$  unabhängige, identisch verteilte Bernoulli Zufallsvariablen mit Erfolgswahrscheinlichkeit  $p$ . Für  $X_n := Y_1 + \dots + Y_n$  und  $\lambda = np$  gilt*

$$\sum_{k=0}^{\infty} \left| \mathbb{P}(X_n = k) - \frac{\lambda^k}{k!} e^{-\lambda} \right| \leq 2np^2. \tag{12}$$

Lemma 7 zeigt, dass die (Terme unter der) Summe (12) die Größenordnung  $np^2 = \lambda p$  haben. Wenn wir  $p = p_n$  wählen, können wir (4) beweisen. Für  $k \leq n$  gilt

$$\left| \mathbb{P}(X_n \leq k) - \sum_{i=0}^k \text{Poi}(\lambda; i) \right| \leq \sum_{i=0}^k |\mathbb{P}(X_n = i) - \text{Poi}(\lambda; i)| \leq 2np_n^2.$$

Wenn wir  $\lim_{n \rightarrow \infty} np_n = \lambda$  annehmen, dann folgt  $\lim_{n \rightarrow \infty} 2np_n^2 = \left( \lim_{n \rightarrow \infty} 2np_n \right) \cdot \left( \lim_{n \rightarrow \infty} p_n \right) = 2\lambda \cdot 0 = 0$ . Der Beweis von (5) wird ganz ähnlich geführt.

*Beweis von Lemma 7 durch »Kopplung«.* Wir konstruieren Paare  $(Y_i, Z_i)$ ,  $i = 1, \dots, n$ , von Zufallsvariablen, so dass  $Y_i$  eine Bernoulli-Verteilung mit Erfolgswahrscheinlichkeit  $p$  und  $Z_i$  eine Poissonverteilung mit Parameter  $p$  hat. Diese Vorgaben bestimmen zunächst nur die Randverteilungen des Vektors  $(Y_i, Z_i)$ , die gemeinsame Verteilung soll durch die folgende Verteilungstabelle (»Kopplung« der Randverteilungen) gegeben sein:

$Z_i \backslash Y_i$	0	1	2	...	k	...	$\mathbb{P}(Y_i = \cdot)$
0	$1 - p$	0	0	...	0	...	$1 - p$
1	$e^{-p} - 1 + p$	$pe^{-p}$	$\frac{p^2}{2!}e^{-p}$	...	$\frac{p^k}{k!}e^{-p}$	...	$p$
$\mathbb{P}(Z_i = \cdot)$	$e^{-p}$	$pe^{-p}$	$\frac{p^2}{2!}e^{-p}$	...	$\frac{p^k}{k!}e^{-p}$	...	

**Tabelle 3.** Die Tabelle gibt die gemeinsame Verteilung der Zufallsvariablen  $Y_i$  und  $Z_i$  an, beispielsweise gilt  $\mathbb{P}(Y_i = 1, Z_i = 2) = \frac{p^2}{2!}e^{-p}$ .

Weiterhin sollen die Zufallsvariablen  $Y_1, Y_2, \dots$  und  $Z_1, Z_2, \dots$  jeweils unabhängig sein.<sup>8</sup>

Wir definieren nun  $X_n := Y_1 + \dots + Y_n$ ,  $W_n := Z_1 + \dots + Z_n$  und  $\lambda = np$ . Offensichtlich ist  $W_n$  wieder Poisson-verteilt mit Parameter  $\lambda$ , und es gilt

$$\begin{aligned} |\mathbb{P}(X_n = k) - \mathbb{P}(W_n = k)| &= |\mathbb{P}(X_n = k, W_n \neq X_n) + \mathbb{P}(X_n = k, W_n = X_n) \\ &\quad - \mathbb{P}(W_n = k, X_n = W_n) - \mathbb{P}(W_n = k, X_n \neq W_n)| \\ &= |\mathbb{P}(X_n = k, W_n \neq X_n) - \mathbb{P}(W_n = k : W_n \neq X_n)| \\ &\leq \mathbb{P}(X_n = k, W_n \neq X_n) + \mathbb{P}(W_n = k, W_n \neq X_n). \end{aligned}$$

Indem wir über  $k = 0, 1, 2, \dots$  summieren, erhalten wir

$$\sum_{k=0}^{\infty} \left| \mathbb{P}(X_n = k) - \underbrace{\mathbb{P}(W_n = k)}_{=e^{-\lambda} \lambda^k / k!} \right| \leq \sum_{k=0}^{\infty} \mathbb{P}(X_n = k, W_n \neq X_n) + \sum_{k=0}^{\infty} \mathbb{P}(W_n = k, W_n \neq X_n) = 2\mathbb{P}(X_n \neq W_n),$$

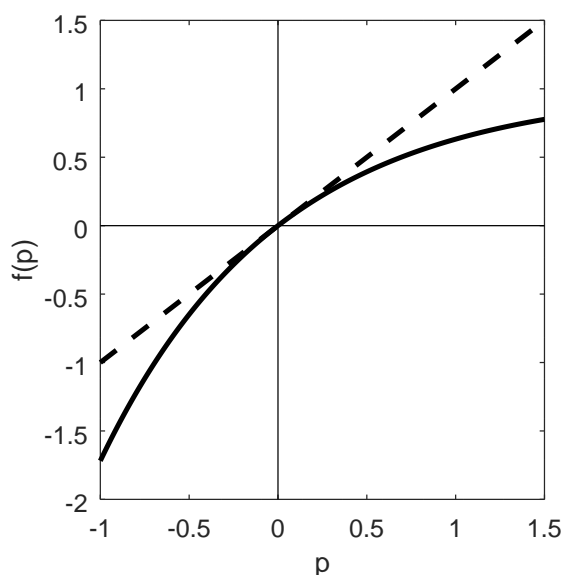
da  $\mathbb{P}(X_n = k) = 0$  für  $k > n$ . Also ergibt sich

$$\mathbb{P}(X_n \neq W_n) \leq \mathbb{P}\left(\bigcup_{i=1}^n \{Y_i \neq Z_i\}\right) \leq \sum_{i=1}^n \mathbb{P}(Y_i \neq Z_i) = \sum_{i=1}^n (1 - \mathbb{P}(Y_i = Z_i)).$$

Der Verteilungstabelle entnehmen wir, dass

$$\mathbb{P}(X_n \neq W_n) \leq \sum_{i=1}^n \left(1 - \underbrace{(1-p)}_{\mathbb{P}(Y_i=Z_i=0)} - \underbrace{pe^{-p}}_{\mathbb{P}(Y_i=Z_i=1)}\right) = np(1 - e^{-p}) \leq np^2.$$

Die letzte Ungleichung verdeutlicht man sich am besten an Hand von Abbildung 1. □

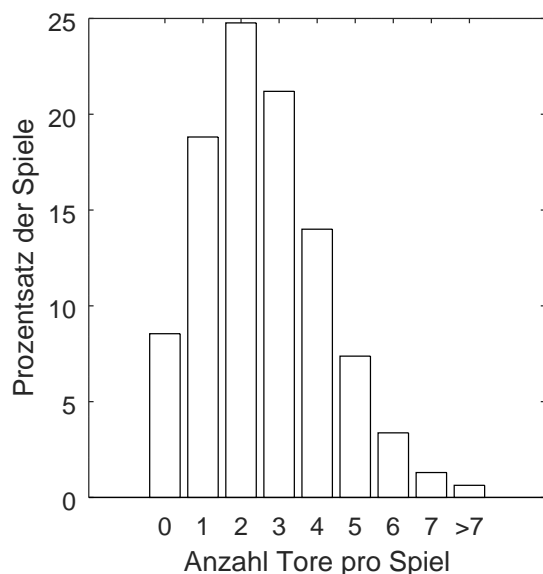


**Abbildung 1.** Die Funktion  $p \mapsto 1 - e^{-p}$  ist konkav (durchgezogene Linie) und  $p \mapsto p$  (gestrichelte Linie) beschreibt deren Tangente an der Stelle  $p = 0$ .

## 4 Eine Anwendung: Poisson-Fußball

Wir kommen nun auf unsere Anwendung der Poisson-Verteilung auf Fußballtore zurück. Untersucht man für den Zeitraum von 2000 bis 2017 die Zahl der erzielten Tore pro Spiel im professionellen Fußball in Deutschland (siehe (Tolan 2018)) oder rund um den Globus, vgl. Abbildung 2, so fällt die vergleichsweise geringe Anzahl – sowohl absolut, als auch im Verhältnis zur Dauer eines Fußballspiels – deutlich ins Auge. Basierend auf etwa 220000 Spielen ergibt sich eine mittlere Zahl von 2,6 Toren pro Spiel. Diese Beobachtung legt es nahe, Fußballtore als »seltene Ereignisse« aufzufassen. Daher können wir den Ausgang eines Fußballspiels durch ein Paar von Poisson-verteilten Zufallsvariablen  $X_H, X_G$  modellieren. Diese stehen für die erzielten Tore der Heim- bzw. Gastmannschaft. Wir nehmen vereinfachend an, dass die Zufallsvariablen unabhängig sind, d. h. der Gegner beeinflusst die Stärke einer Mannschaft nicht.





**Abbildung 2.** Anzahl der erzielten Tore pro Spiel. Basierend auf (Dubitzky *et. al.* 2017) (218916 Spiele von Profifußball-Ligen zwischen 2000 und 2017).

Zwar mag einem echten Fußballfan diese Annahme als wenig realistisch erscheinen, doch wie wir später sehen werden liefert dieses relativ einfache Modell bereits recht gute Ergebnisse.

Da bei der Poisson-Verteilung die Parameter  $\lambda_H, \lambda_G$  den Erwartungswerten der Zufallsvariablen  $X_H$  bzw.  $X_G$  entsprechen, können wir diese als Gradmesser für die »Torgefährlichkeit« oder »Spielstärke« der jeweiligen Mannschaft interpretieren. Zur Ermittlung der Torgefährlichkeit wurden in der Literatur bereits verschiedene Ansätze untersucht. Für Welt- und Europameisterschaften eignen sich Parameter auf Basis von Weltranglistenpunkten, wie der »FIFA/Coca-Cola-Weltrangliste« oder der »World Football ELO«-Rangliste, siehe (Gilch & Müller 2018; Leitner *et. al.* 2010). Ist man an Aussagen über den Spielausgang während einer laufenden Saison interessiert, so stellt die durchschnittliche Zahl von erzielten Toren je Spiel (also die erzielten Tore im Verhältnis zur Zahl der gespielten Partien) eine sinnvolle Kennzahl für die Spielstärke der Mannschaften dar.

Die Wahrscheinlichkeiten, dass die Heimmannschaft gewinnt bzw. verliert bzw. unentschieden spielt sind  $\mathbb{P}(X_H > X_G)$  bzw.  $\mathbb{P}(X_H = X_G)$  bzw.  $\mathbb{P}(X_H < X_G)$ . Zwar lassen sich diese Wahrscheinlichkeiten hinschreiben, etwa

$$\begin{aligned} \mathbb{P}(X_H > X_G) &= \sum_{k=0}^{\infty} \mathbb{P}(X_H > k) \mathbb{P}(X_G = k) = \sum_{k=0}^{\infty} \mathbb{P}(X_H > k) \frac{\lambda_G^k}{k!} e^{-\lambda_G} \\ &= \sum_{k=0}^{\infty} \sum_{i=k+1}^{\infty} \mathbb{P}(X_H = i) \frac{\lambda_G^k}{k!} e^{-\lambda_G} = \sum_{k=0}^{\infty} \sum_{i=k+1}^{\infty} \frac{\lambda_H^i \lambda_G^k}{i! \cdot k!} e^{-(\lambda_H + \lambda_G)}, \end{aligned}$$

aber bereits für ein Einzelspiel ist an einen geschlossenen Ausdruck nicht zu denken; eine Turniersituation ist noch viel aufwendiger. Daher bleibt eigentlich nur die Möglichkeit, die Wahrscheinlichkeiten mit dem Computer zu ermitteln. Dazu wird das Turnier oder die Saison im Computer hunderttausendfach simuliert und die sich daraus ergebenden relativen Häufigkeiten der Platzierungen können als Basis für eine Prognose des Abschneidens benutzt werden. Wir müssen dazu Realisierungen von Poisson-Zufallsvariablen auf dem Computer erzeugen, sogenannte *Poisson-Zufallszahlen*.

---

## Erzeugung von Poisson-Zufallszahlen

Auf die Problematik »zufällige« Zahlen im Computer mittels deterministischer (d. h. nicht-zufälliger) Verfahren zu erzeugen, wollen wir hier nicht eingehen. Wir verweisen dazu auf (Kneusel 2018; Kolonko 2008) und gehen davon aus, dass wir effiziente Methoden zur Erzeugung von auf dem Intervall  $(0, 1)$  gleichverteilten Zufallszahlen  $U$  zur Verfügung haben. Als Basis von Zufallszahlengeneratoren für weitere Verteilungen dient folgender Zusammenhang: Betrachtet man zu einer (beliebigen) Zufallsvariablen  $X$  die zugehörige Verteilungsfunktion  $F$ , so ist diese in der Regel zwar nicht *streng* monoton wachsend – und damit nicht invertierbar – aber zumindest liefert die Definition

$$F^{-1}(u) := \inf\{x \in \mathbb{R} : F(x) \geq u\}, \quad u \in (0, 1),$$

eine Funktion  $F^{-1} : (0, 1) \rightarrow \mathbb{R}$  welche für alle reelle Zahlen  $x$  und alle  $u \in (0, 1)$

$$F(x) \geq u \iff F^{-1}(u) \leq x, \quad (13)$$

sowie

$$F(F^{-1}(u)) = u, \quad u \in (0, 1),$$

erfüllt. Man nennt  $F^{-1}$  die *inverse Verteilungsfunktion* (von  $X$ ).

Ist nun  $U$  eine auf dem Intervall  $(0, 1)$  gleichverteilte Zufallsvariable, so besitzt die neue Zufallsvariable  $F^{-1}(U)$  dieselbe Verteilung wie  $X$ . In der Tat ist für beliebiges  $x \in \mathbb{R}$  die Verteilungsfunktion von  $F^{-1}(U)$  gegeben durch

$$\mathbb{P}(F^{-1}(U) \leq x) \stackrel{(13)}{=} \mathbb{P}(U \leq F(x)) = \mathbb{P}(U \in [0, F(x)]) = F(x).$$

Da sich für eine Poisson-Zufallsvariable  $X$  die Verteilungsfunktion  $F$  gemäß

$$F(x) = \sum_{k=0}^{\lfloor x \rfloor} \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=0}^{\lfloor x \rfloor} \text{Poi}(\lambda; k), \quad x \in \mathbb{R},$$

( $\lfloor x \rfloor$  bezeichnet den ganzzahligen Anteil von  $x$ ) ergibt und für die Einzelwahrscheinlichkeiten die Rekursion (8) gilt, liefert der folgende Algorithmus Poisson-Zufallszahlen, vgl. (Kolonko 2008, Beispiel 8.9b):

Setze  $p = \exp(-\lambda)$ ,  $F=p$ ,  $k=0$

Erzeuge eine auf  $(0,1)$  gleichverteilte Zufallszahl  $u$

WIEDERHOLE

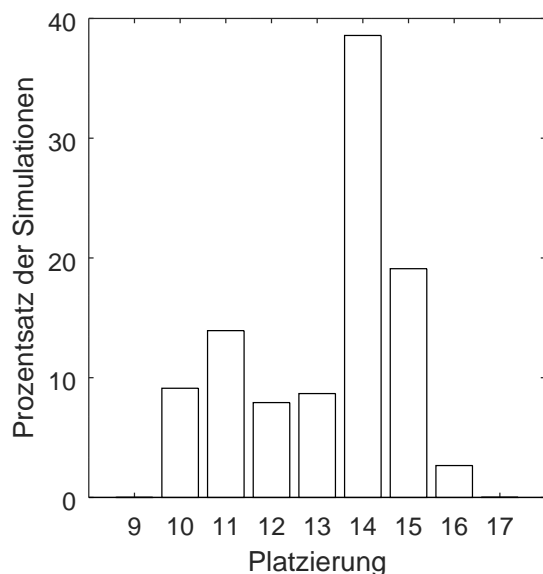
$p = p \cdot \lambda / (k+1)$ ,  $F = F + p$ ,  $k=k+1$

SOLANGE  $u > F$

Gebe  $x=k-1$  aus.

## Poisson-Fußball im Praxistest

Eingangs haben wir bereits die Situation der SG Dynamo Dresden vor dem letzten Spieltag der Saison 2017/2018 beschrieben. Basierend auf dem vorher beschriebenen Poisson-Modell haben wir die Spielausgänge der vier für den Abstieg relevanten Partien simuliert und daraus die Tabellenstände berechnet. Bei 100000 Wiederholungen ergab sich für die Platzierung der SG



**Abbildung 3.** Häufigkeitsverteilung der Platzierungen bei einer Simulation mit 100000 Wiederholungen. Die Platzierungen 9 und 17 haben eine Wahrscheinlichkeit von weniger als 0,5 %.

Dynamo das Häufigkeitsdiagramm in Abbildung 3. Die am häufigsten vorkommende Tabellenposition war der vierzehnte Platz (mit 38 %), auf welchem Dynamo die Saison auch tatsächlich beendete.

Es ist hier wichtig zu betonen, dass die Ergebnisse des Poisson-Modells stark von einer realistischen (Ein-)Schätzung der Torgefährlichkeiten  $\lambda$  abhängen. So ist, u. a., wichtig, dass über den betrachteten Zeitraum die Zusammensetzung und das Spielniveau der Mannschaften einigermaßen konstant bleiben. In der beschriebenen Situation führte die gute Datenlage zu einem sinnvollen Modell. Im Gegensatz dazu stellt sich die Lage bei einer WM-Endrunde sehr viel schwieriger dar.

## 5 Zur »Unvermeidbarkeit« der Poisson-Verteilung

Wir schreiben  $X_t$  für die Zahl, wie oft ein Ereignis  $E$  im Zeitintervall  $(0, t]$  eingetreten ist, z. B. ist für  $E$  »Tor für die Heimmannschaft«  $X_{45}$  die Zahl der Tore der Heimmannschaft in den 45 Minuten der ersten Halbzeit. Wir wollen zeigen, dass  $X_t$  unter relativ schwachen Annahmen einer Poisson-Verteilung folgen muss. Dazu nehmen wir an, dass

- (A1) zu jedem Zeitpunkt das Ereignis höchstens einmal auftreten kann;
- (A2) in jedem beschränkten Zeitintervall das Ereignis nur endlich oft eintreten kann;
- (A3) das Auftreten des Ereignisses in nicht-überlappenden Zeitintervallen unabhängig ist;
- (A4) das Auftreten des Ereignisses nicht vorhersagbar ist, d. h. zu jedem festen  $t$  gilt
 
$$\lim_{h \rightarrow 0} \mathbb{P}(X_{t+h} - X_t > 0) = \lim_{h \rightarrow 0} \mathbb{P}(X_t - X_{t-h} > 0) = 0.$$

Wir werden (A4) meist in Form von Gegenwahrscheinlichkeiten verwenden:

$$\lim_{h \rightarrow 0} \mathbb{P}(X_{t+h} - X_t = 0) = \lim_{h \rightarrow 0} \mathbb{P}(X_t - X_{t-h} = 0) = 1.$$

Im folgenden schreiben wir  $X(I) = X_t - X_s$  für die Zahl der Ereignisse in einem Intervall  $I = (s, t]$ . Beachte, dass  $X(I \cup J) = X(I) + X(J)$  ist, wenn  $I$  und  $J$  nicht-überlappende Intervalle sind.

Wir behaupten, dass  $\mathbb{P}(X_t - X_s = 0) > 0$  für beliebige  $s < t$  gilt. Dazu unterteilen wir das Intervall  $I = (s, t]$  in  $n$  nicht-überlappende Intervalle  $I_{i,n} = (s + \frac{i-1}{n}h, s + \frac{i}{n}h]$ ,  $h = t - s$ ,  $i =$

$1, 2, \dots, n$  und bemerken, dass  $X(I) = 0$  genau dann gilt, wenn  $X(I_{i,n}) = 0$  für alle  $i$  ist. Im Hinblick auf (A4) können wir  $n$  so groß wählen, dass  $\mathbb{P}(X(I_{i,n}) = 0) > 0$  für alle  $i = 1, \dots, n$  gilt. Dann folgt aber auch

$$\mathbb{P}(X(I) = 0) = \mathbb{P}\left(\bigcap_{i=1}^n \{X(I_{i,n}) = 0\}\right) \stackrel{(A3)}{=} \prod_{i=1}^n \mathbb{P}(X(I_{i,n}) = 0) > 0.$$

Wir können nun eine von Intervallen der Art  $I = (s, t]$  abhängige Funktion  $\lambda$  definieren

$$\lambda(I) = -\ln \mathbb{P}(X(I) = 0).$$

Auf Grund der Annahme (A4) gilt  $\lambda(\{t\}) = -\lim_{h \rightarrow 0} \ln \mathbb{P}(X(I_h) = 0) = -\ln 1 = 0$ , wobei  $I_h$  für  $(t-h, t]$  oder  $(t, t+h]$  steht.

Nun seien  $J = (0, t]$  und  $J = J_{1,n} \cup J_{2,n} \cup \dots \cup J_{n,n}$  eine Zerlegung in nicht-überlappende Intervalle. Wir wählen die  $J_{i,n}$  so, dass  $\lambda(J_{i,n}) = \frac{1}{n} \lambda(J)$  gilt. Das ist möglich, weil  $J \mapsto \lambda(J)$  monoton von der Länge des Intervalls abhängt und für jeden Punkt  $\lambda(\{s\}) = 0$  gilt.

Weil es in jedem Intervall nur endlich viele Ereignisse gibt (A2) und zu keinem Zeitpunkt zwei oder mehr Ereignisse simultan eintreten können (A1), gilt für

$$\Omega_n = \{X(J_{i,n}) = 0 \text{ oder } 1 \text{ für alle } i\} \quad \text{dass} \quad \lim_{n \rightarrow \infty} \mathbb{P}(\Omega_n) = 1.$$

Wir definieren mit Hilfe der  $X(J_{i,n})$  Bernoulli-Zufallsvariable:  $Y_{i,n} = \mathbb{1}_{\{X(J_{i,n})=1\}}$ , d. h.  $Y_{i,n}$  ist 1 bzw. 0, wenn  $X(J_{i,n}) = 1$  bzw.  $X(J_{i,n}) \neq 1$  ist; die Erfolgswahrscheinlichkeit ist  $\mathbb{P}(Y_{i,n} = 1) = \mathbb{P}(X(J_{i,n}) = 1) = 1 - e^{-\lambda(J_{i,n})} = 1 - e^{-\lambda(J)/n}$ .

Als Funktion von  $X(J_{i,n})$ , sind die  $Y_{i,n}$  wiederum unabhängig.<sup>9</sup> Insbesondere ist  $Y_n = Y_{1,n} + Y_{2,n} + \dots + Y_{n,n}$  binomialverteilt. Auf  $\Omega_n$  gilt  $X(J_{i,n}) = Y_{i,n}$ , weswegen wir

$$\{X(J) = Y_n\} = \{X(J_{1,n}) + X(J_{2,n}) + \dots + X(J_{n,n}) = Y_{1,n} + Y_{2,n} + \dots + Y_{n,n}\} = \Omega_n$$

haben. Aus der Additionsregel für Wahrscheinlichkeiten folgt

$$\begin{aligned} \mathbb{P}(X(J) = k) &= \mathbb{P}(\{X(J) = k\} \cap \Omega_n) + \mathbb{P}(\{X(J) = k\} \cap \Omega_n^c) \\ &= \mathbb{P}(\{Y_n = k\} \cap \Omega_n) + \mathbb{P}(\{X(J) = k\} \cap \Omega_n^c) \\ &= \mathbb{P}(Y_n = k) - \mathbb{P}(\{Y_n = k\} \cap \Omega_n^c) + \mathbb{P}(\{X(J) = k\} \cap \Omega_n^c). \end{aligned}$$

Die beiden letzten Terme lassen sich jeweils durch  $\mathbb{P}(\Omega_n^c) = 1 - \mathbb{P}(\Omega_n) \rightarrow 0$  abschätzen, den ersten Term können wir mit den Überlegungen aus Abschnitt 1 behandeln:

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}(Y_n = k) &= \frac{1}{k!} \lim_{n \rightarrow \infty} \frac{n}{n} \cdot \frac{n-1}{n} \cdot \dots \cdot \frac{n-k+1}{n} \left(e^{-\lambda(J)}\right)^{\frac{n-k}{n}} \left(\frac{1 - e^{-\frac{\lambda(J)}{n}}}{\frac{1}{n}}\right)^k \\ &= \frac{1}{k!} e^{-\lambda(J)} \lambda(J)^k. \end{aligned}$$

Es folgt, dass  $X(J) = X((0, t])$  Poisson-verteilt mit Parameter  $\lambda_t = \lambda((0, t])$  ist.

**Danksagung.** Wir danken unseren Kollegen Björn Böttcher (TU Dresden), Hans Fischer (KU Eichstätt) und Katharina Fischer (HTW Dresden) für die kritische Durchsicht des Manuskripts und zahlreiche Anmerkungen. Besonderer Dank gilt den Herausgebern des Sonderhefts »Stochastik« für die Einladung diesen Beitrag zu verfassen und für die stets gute Zusammenarbeit.

## Literatur

- H. Bateman: On the probability distribution of particles. *Philosophical Magazine and Journal of Science, 6th series.* **20** (1910) 704–707.
- L. von Bortkewitsch: *Das Gesetz der kleinen Zahlen*. Teubner, Leipzig 1898.
- R.D. Clarke: An application of the Poisson distribution. *Journal of the Institute of Actuaries* **72** (1946) 48.
- A. de Moivre: *The Doctrine of Chances*. A. Millar, London 1756 (3. Aufl.; 1. Aufl., London 1718)
- W. Dubitzky, P. Lopes, J. Davis, D. Berrar: *Open International Soccer Database*, osf.io/kqcye, 2017 (abgerufen am 24.09.2018).
- A.K. Erlang: Wahrscheinlichkeitsrechnung und Telefonverbindungen (in Dänisch). *Nyt Tidsskrift for Matematik B* **20** (1909) 33–39.
- E.J. Gumbel: Les centenaires. *Aktuárske Vedy, Praha* **7** (1937) 1–8.
- L.A. Gilch, S. Müller: On ELO based prediction models for the FIFA Worldcup 2018. arXiv:1806.01930 [stat.AP], 2018.
- J.L. Hodges Jr., L. Le Cam: The Poisson Approximation to the Poisson Binomial Distribution. *The Annals of Mathematical Statistics* **31** (1960) 737–740.
- R.T. Kneusel: *Random Numbers and Computers*. Springer, Cham 2018.
- M. Kolonko: *Stochastische Simulation : Grundlagen, Algorithmen und Anwendungen*, Vieweg+Teubner, Wiesbaden 2008.
- C. Leitner, A. Zeileis, K. Hornik: Forecasting sports tournaments by ratings of (prob)abilities: A comparison for the EURO 2008. *International Journal of Forecasting* **26** (2010) 471–481.
- P.S. Laplace: *Théorie analytique des probabilités*. Courcier, Paris 1820 (3. Aufl.; 1. Aufl., Paris 1812).
- P. Mühlbauer, K. Wörle: *Tafelwerk zur Stochastik*. Bayerischer Schulbuch-Verlag, München 1985.
- S.D. Poisson: Mémoire sur la proportion des naissances des filles et des garçons. *Mém. Acad. Sci. Paris* **8** (1829) 239–308.
- S.D. Poisson: *Lehrbuch der Wahrscheinlichkeitsrechnung und deren wichtigsten Anwendungen*. G.C.E. Meyer sen., Braunschweig 1841 (Franz. Originalfassung: Bachelier, Paris 1837).
- L.F. Richardson: The distribution of wars in time. *Journal of the Royal Statistical Society* **107** (1944) 242–250.
- R.L. Schilling: *Wahrscheinlichkeit*. De Gruyter, Berlin 2017.
- »Student«: On the error of counting with a haemocytometer. *Biometrika* **5** (1907) 351–360.<sup>10</sup>
- M. Tolan: *Manchmal gewinnt der Bessere. Die Physik des Fußballspiels*. Piper, München 2018

## Anmerkungen

<sup>1</sup>Die Zitate verweisen jeweils auf die historisch erstmalige Anwendung.

<sup>2</sup>Auf Poisson geht auch der Begriff *Gesetz der großen Zahlen* zurück (Poisson 1841, S. V), er hat als einer der ersten zwischen abstrakten Zufallsvariablen (*Erscheinungen jeglicher Art* (Poisson 1841, S. V)) und ihren Werten unterschieden.

<sup>3</sup>Mit handelsüblichen Taschenrechnern kann man maximal 69! berechnen; in der Schule verwendete Tafelwerke (z.B. (Mühlbauer & Wörle 1985)) enthalten meist nur die Werte für  $n = 3, \dots, 10, 15, 20, 25, 30, 50, 100, 200$  und einige  $p \in (0, 1)$  in Schritten von 0,01 bzw. 0,05.

<sup>4</sup>Ein Vorläufer dieser Formel findet sich bei de Moivre (de Moivre 1756, S. 243ff.) – daher der Name »Zentraler Grenzwertsatz von de Moivre-Laplace.«

---

<sup>5</sup>Das sieht man folgendermaßen:

$$\begin{aligned}\mathbb{P}(Z+W=m) &\stackrel{\text{unabh.}}{=} \sum_{k=0}^{\infty} \mathbb{P}(Z+k=m)\mathbb{P}(W=k) = \sum_{k=0}^m \mathbb{P}(Z=m-k) \frac{\mu^k}{k!} e^{-\mu} \\ &= \sum_{k=0}^m \frac{\lambda^{m-k}}{(m-k)!} e^{-\lambda} \frac{\mu^k}{k!} e^{-\mu} = \frac{1}{m!} \sum_{k=0}^m \binom{m}{k} \lambda^{m-k} \mu^k e^{-\lambda-\mu} = \frac{(\lambda+\mu)^m}{m!} e^{-(\lambda+\mu)}.\end{aligned}$$

In der zweiten Gleichheit verwenden wir, dass  $Z$  keine negativen Werte annehmen kann, d. h., dass  $k \leq m$  gelten muss.

<sup>6</sup>Die Bedingung » $m$ -ter Misserfolg findet an der Stelle  $m+i$  statt« führt für verschiedene Werte von  $i=0, \dots, k$  zu unterschiedlichen Mengen, d. h. wir erhalten eine paarweise disjunkte Zerlegung.

<sup>7</sup>Diese Annahme ist problematisch, da es saisonale Schwankungen bei Geburten gibt.

<sup>8</sup>Während man sich die Existenz jedes der zweidimensionalen Zufallsvektoren  $(Y_i, Z_i)$  noch ganz gut intuitiv vorstellen kann, ist die Existenz der Familie  $\{(Y_i, Z_i), i=1, 2, \dots\}$  mit den o.g. Unabhängigkeitseigenschaften nicht offensichtlich und bedarf einer Standard(produkt-)konstruktion, vgl. (Schilling 2018, Kapitel 6), die weit über den Schulstoff hinausgeht.

<sup>9</sup>Vgl. (Schilling 2018, Korollar 6.8); diese Tatsache kann man auch elementar mit der Definition der Unabhängigkeit nachrechnen, wenn man  $\{Y_{i,n}=1\} = \{X(I_{i,n})=1\}$  und  $\{Y_{i,n}=0\} = \{X(I_{i,n}) \neq 1\}$  beachtet.

<sup>10</sup>»Student« war das Pseudonym des Chemikers und Mathematikers William Sealy Gosset, der Angestellter der Guinness-Brauerei in Dublin war.