Mathematik der Natur - Tiefgreifende Konsequenzen aus dem Lemma von Borel-Cantelli

Kristallmann Philipp

May 27, 2023

Abstrakt

Das in diesem Abschnitt dargelegte fundamentale Resultat, mutmaßlich erstmals von Borel und Cantelli entdeckt, findet leider in heutiger Zeit wenig Beachtung. Allerdings habe ich bereits während meinem Studium der Mathematik festgestellt, dass es sich hierbei um ein sehr wichtiges, tiefgründiges und weitreichendes Ergebnis handelt, welches näher untersucht werden sollte. Das könnte ein grundlegender Ansatzpunkt für weitere Erkenntnisse in Philosophie, Physik, Mathematik, Astronomie, Rechtswesen und weiteren Bereichen sein. Mit diesem Abschnitt möchte ich euch ein paar Ideen vorstellen, welche ich seit dem erstmaligen Kennenlernen gewonnen habe.

Vorbereitung und grundlegende Aussagen

Im Folgenden sei Z eine Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$. Es sei mit B_n eine unendliche Folge von Ereignissen[1] bezeichnet. Unabhängigkeit ist genau das was es sagt: Zwei Ereignisse (oder Events) passieren unabhängig voneinander im wörtlichen Sinne. Die große Leistung der Wahrscheinlichkeitstheoretiker bestand darin, diesen Umstand mathematisch formal korrekt zu benennen. Diese Definition der Unabhängigkeit zweier Mengen ist folgendermaßen gegeben: Seien B_1 und B_2 zwei Ereignisse auf dem oben besagten Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$. Dann gilt, dass die Wahrscheinlichkeit des Schnittes der beiden Ereignisse gleich dem Produkt der Wahrscheinlichkeiten der jeweiligen Ereignisse ist. Formal bedeutet

[1] Ein Ereignis ist eine Menge von möglichen Ergebnissen aus dem Wahrscheinlichkeitsraum Ω. Zum Beispiel ist bei dem einfachen, einmaligen Würfelwurf der Ereignisraum Ω gleich der Menge aller möglichen Augenzahlen von 1 bis 6. Ein Ereignis davon wäre zum Beispiel $\{1, 2, 3\}$ oder $\{$"Die Menge aller geraden Zahlen"$\} = \{2, 4, 6\}$.

das:

$$P[B_1 \cap B_2] = P[B_1] \cdot P[B_2].$$

Durch einfache Überlegungen sind wir im Stande, diese Definition auf beliebig viele Ereignisse zu erweitern. Eine Menge an Ereignissen heißt paarweise unabhängig, wenn zwei beliebig gewählte Ereignisse aus der Gesamtheit aller möglichen Ereignisse[2] unabhängig sind.

Wesentliches Resultat

Das Lemma von Borel-Cantelli besagt nun in vereinfachender und für unsere Zwecke ausreichender Form folgendes:

Lemma 1

- *Falls die Summe der Wahrscheinlichkeiten $P[B_n]$ über alle natürlichen Zahlen n endlich ist, so gilt: Die Wahrscheinlichkeit für das Ereignis $\{B_n$ trifft ∞ oft auf$\}$ beträgt 0. Formal bedeutet dies:*

  $\sum_{n=1}^{\infty} P[B_n] < \infty \Rightarrow P[\{B_n \textit{ trifft } \infty \textit{ oft auf}\}] = 0.$

- *Falls die Summe der Wahrscheinlichkeiten für paarweise unabhängige Ereignisse B_n unendlich groß ist, so gilt: Die Wahrscheinlichkeit für das Ereignis $\{B_n$ trifft ∞ oft auf$\}$ beträgt 1. Formal bedeutet dies:*

  $\sum_{n=1}^{\infty} P[B_n] = \infty \Rightarrow P[\{B_n \textit{ trifft } \infty \textit{ oft auf}\}] = 1.$

Die 1. Aussage können wir so interpretieren: Falls wir nur eine endliche Summe der Wahrscheinlichkeiten bezüglich der unendlichen Folge an Ereignissen B_n haben, dann gibt es auch nur endlich viele Mengen aus der Folge, welche eine positive Wahrscheinlichkeit aufweisen. Die restlichen Ereignisse sind alles Nullmengen[3] oder Mengen mit stark abfallendem Wert, sodass diese in der Summe endlich sind. Also können die Ereignisse nur endlich oft auftreten. Dieses Resultat ist intuitiv einleuchtend.

[2] Das ist genau der Raum $\mathcal{A}$, die Menge aller Ereignisse. Für Zufallsvariablen mit endlicher Anzahl Ergebnisse wird in der Regel ohne große Einschränkungen die sogenannte Potenzmenge (also alle möglichen Teilmene des Ergebnisraumes Ω) genommen. Hier könnte man auch schreiben: Beliebige Ereignispaare.

[3] Nullmengen sind Mengen, welche mit Wahrscheinlichkeit 0 auftreten. Mathematisch formal bedeutet das: Sei N so eine Nullmenge, dann gilt: $P[N] = 0$.

Genauso verhält es sich mit der zweiten Aussage, jedoch ist es komplizierter, dieses Resultat auch zu beweisen[4]. Hier ist der Zusatz der "paarweise unabhängigen" Mengen entscheidend. Dieser kann durch weiter abschwächende Annahmen verallgemeinert werden. Für die meisten Fälle und die hier betrachteten Beispiele reicht das jedoch aus.

Konsequenzen aus diesem Satz

Dieses Resultat scheint zuerst nicht beeindruckend. Allerdings sind aufgrund der Allgemeinheit dieser Aussage auch weitreichende und tiefgründige Schlussfolgerungen möglich. Anhand einiger Beispiele möchte ich das genauer verdeutlichen.

Kein statistischer Test ist 100% genau!

Aus der Medizin wissen wir, dass dort medizinische Tests Anwendung finden. Vielen Leuten fällt die Unterscheidung schwer, ob man tatsächlich krank ist oder nicht, oder nur der Test positiv oder negativ[5] ausschlägt. Daher ergeben sich 4 mögliche Kombinationen: Entweder ist der Patient krank und der Test ist positiv; oder er ist krank und der Test ist negativ; oder er ist gesund und der Test ist positiv; oder er ist gesund und der Test ist negativ. Dieser Sachverhalt kann leicht als Mengenschreibweise geschrieben werden, in der Art:

$$P = \{\text{"Test ist positiv"}\}$$
$$N = \{\text{"Test ist negativ"}\}$$
$$K = \{\text{"Patient ist tatsächlich krank"}\}$$
$$G = \{\text{"Patient ist tatsächlich gesund"}\}$$

[4] Den Beweis möchte ich hier nicht erbringen, da ein wenig ausgeholt werden muss. Ich empfehle für den interessierten Leser das Standardlehrwerk "Wahrscheinlichkeitstheorie" von Klenke. Alternativ ist es auch möglich, im Internet nach Vorlesungsskripten zu diesem Lemma zu recherchieren, da es sich hierbei um ein grundlegendes Resultat handelt.

[5] Falls ein Test positiv ausschlägt, so wird behauptet, dass der Patient krank ist. Allerdings bedeutet das noch nicht, dass dieser tatsächlich krank ist. Umgekehrt verhält es sich genauso: Ein negativer Test bedeutet noch lange nicht, dass der Patient tatsächlich gesund ist.

Testergebnis / Tatsächlicher Status	**Positiv (P)**	**Negativ (N)**	
Krank (K)	$P[K \cap P]$	$P[K \cap N]$	$P[K]$
Gesund (G)	$P[G \cap P]$	$P[G \cap N]$	$P[G]$
	$P[P]$	$P[N]$	1

Tabelle 1: Grundlegender Aufbau einer Vierfelder-Tafel. In ihr sind die jeweiligen abstrakten Wahrscheinlichkeiten eingetragen. Beispielsweise sind in der Zeile "Krank" in der ersten Spalte die Wahrscheinlichkeit abgebildet, dass der Patient sowohl krank als auch positiv getestet ist. ($P[K \cap P]$) als auch krank und negativ getestet ist ($P[K \cap N]$). Rechts von diesen beiden Zellen ist die Wahrscheinlichkeit dafür abgebildet, dass der Patient krank ist. Das ergibt nach elementarer Berechnung die Summe der beiden vorigen Wahrscheinlichkeiten (also $P[K \cap P] + P[K \cap N] = P[K]$).

In Tabelle 1 sind die daraus resultierenden Schnittereignisse sehr treffend in einer Vierfeldertafel zusammengefasst und in sinnvolle Form gebracht.
Zunächst halten wir einige Annahmen fest, welche sich logischerweise im Rahmen einer Vielzahl von statistischen Tests ergeben, und zwar:

1. Die Tests geben unabhängig voneinander die Ergebnisse aus[6] und

2. wir haben mit Wahrscheinlichkeit größer 0 und kleiner 1 sowohl Gesunde, Kranke, positive Tests als auch negative Tests[7].

Wir folgern aus dem Lemma, dass unter den obigen Annahmen eine beliebige Konstellation aus Testergebnissen und Gesundheitszustand mit 100 %-iger Wahrscheinlichkeit unendlich oft auftreten wird. Dem ist so, weil erstens die Ereignisse (also hier jeweils ein für sich betrachtetes Testergebnis) paarweise unabhängig zueinander sind und zweitens alle hier und in Tabelle 1 untersuchten Wahrscheinlichkeiten für jeden der möglichen Fälle (Ereignisse P, N, G und K) für sich genommen positiv sind.

[6] Anders gesagt: Ein beliebiger Test aus einer (möglicherweise unendlichen) Testreihe darf ein beliebig gewähltes aber anderes Testresultat nicht beeinflussen. Damit haben wir genau die Definition der paarweisen Unabhängigkeit erfüllt.

[7] Mathematisch formuliert bedeutet das: $0 \leq P[P], P[N], P[G], P[K] \leq 1$.

Geschichte wiederholt sich (nicht)!

Es steht seit geraumer Zeit die Frage im Raum, ob die Geschichte sich wiederholt oder nicht. Die einen behaupten, dass dem so sei. Deren Begründung liegt in den historischen Prozessen verankert, in denen zyklisch wiederholende Fakten herangezogen werden. Diese belegen dann den immer wieder neu aufkommenden Charakter der Geschichte.
Andere stellen sich hin und weisen auf die offensichtlichen Unterschiede in den aufgeführten Fakten hin. Sie stellen fest, das sich zwar die Geschichte in bestimmten Aspekten wiederholt, jedoch die Bedingungen andere waren als zuvor. Es lassen sich Details finden, welche darauf schließen könnten, dass es doch keine tatsächliche Wiederholung gibt. Vielmehr finde ein Entwicklungsprozess statt.
Welche dieser Ansichten ist nun die korrekte? Wiederholt sich die Geschichte oder nicht? Meine Antwort darauf ist folgende: Es kommt auf die Definitionen und die Sichtweise an!

Unter welchen Umständen sie sich wiederholt

Was genau meine ich damit? Die Befürworter der These, welche für die Wiederholung sprechen, legen die Ereignisse auf bestimmte Parameter fest. Zum Beispiel sagen sie: Wir betrachten das Ereignis, dass eine Frau ein Land führt (im Sinne der öffentlich sichtbaren Monarchin, Präsidentin usw.) und schauen in der Historie, wann dies genau geschah, und nach welcher Zeit sich das wiederholt. Daraus leiten wir ab, dass immer wieder Frauen an die Spitze kamen, also wiederholt sich die Geschichte. Diese Aussage ist soweit korrekt unter dem hier besprochenen Lemma und unter der Annahme, dass die Frauen unabhängig voneinander an die Macht kommen[8]. Die Wahrscheinlichkeit für das Ereignis ist klein, aber sicherlich positiv. Daraus kann man unter den Annahmen und nach Borel-Cantelli schließen, dass in Zukunft mit 100 % Wahrscheinlichkeit immer wieder Frauen Länder regieren werden.

[8] Diese Annahme ist durchaus plausibel, wobei auch in jüngerer Geschichte "weibliche Seilschaften" zu finden sind, die nicht als unabhängig wegzuargumentieren sind. Hier nehme ich trotzdem Unabhängigkeit an, insbesondere in Bezug mit den vorhin erläuterten Möglichkeiten der "Aufweichung", also Schwächung dieser Unabhängigkeits-Annahme für dieses Lemma.

Unter welchen Umständen sie sich nicht wiederholt

Bleiben wir bei dem Beispiel. Nun gibt es aber Gegner dieser Hypothese. Diese würden bei der eben geführten Argumentation behaupten, dass wir hier die Ereignisse zu stark einschränken. Denn zwar wiederholt sich das Ereignis, dass Frauen an die Macht kommen, allerdings waren sie beispielsweise unter teils völlig verschiedenen Umständen an sie gekommen, oder haben diese unter anderen Rahmenbedingungen verloren und so weiter. Allgemein gesagt war zwar das vorgegebene Event wiederkehrend, aber die Ausgangssituation und das Bild drum herum waren anders. Der Anhänger dieser Hypothese argumentiert mathematisch so, dass der Ereignisraum eigentlich ein Raum mit unendlich vielen verschiedensten Parametern ist. Somit betrachten wir nur wenige der unendlich[9] vielen möglichen Eigenschaften des historischen Ereignisses. Dadurch wird das hier geschilderte Event wegen der unendlichen Vielzahl an zu berücksichtigten Details zu einer Nullmenge und somit der andere Teil des Borel-Cantelli-Lemmas wirksam. Die Wahrscheinlichkeit mit der sich dieses konkrete Ereignis haargenau so wie in den letzten Malen zutragen wird, beträgt somit 0%.

Wer hat Recht?

Tatsächlich haben beide gewissermaßen Recht. Wenn wir für einen Augenblick die hier getätigten Annahmen für wahr befinden, dann stellen wir fest, dass beide Sichtweisen logisch schlüssig sind. Es handelt sich hier tatsächlich um die korrekte Definition und Sichtweise[10], wie wir die geschichtlichen Ereignisse auffassen. Entweder ist es ein von dem Gesamtprozess abgetrenntes Ereignis, oder wir gliedern es in den globalen[11] historischen Prozess mit all den insgesamt von allen Wesenheiten geschaffen Ereignissen und Informationen ein.

[9] Hier könnte man auch sagen überabzählbar vielen Ereignissen. Damit wird unter unserem Wahrscheinlichkeitsmaß P jedes bis auf abzählbar viele Ereignisse zu einer Nullmenge.

[10] Das wird übrigens noch astrologisch untermauert. Denn die Sternenkonstellationen wiederholen sich, wenn überhaupt (Das ist noch näher zu untersuchen und wäre eine interessante Fragestellung.), nach einer sehr langen Zeit. Allein dieser Umstand verhindert dass sich die Geschichte stets eins zu eins wiederholen kann. Vermutlich repetieren sich die Himmelskörper nie, sondern bilden immer ein eindeutiges und einmaliges Abbild.

[11] oder gar kosmischen historischen Prozess?

Statistische Ermittlungsverfahren

Dem aufmerksamen Leser mag sich die Frage stellen, wie wir diese Prozesse quantitativ erfassen könnten. Wie ist es möglich, dieses fundamentale theoretische Wissen in der Praxis anwenden zu können? Dazu widmen wir uns dem eben beschriebenen Beispiel der medizinischen Tests zu.

Ermittlung der Periodendauer und Frequenz der Testergebnisse

Kehren wir zurück zu der eben geschilderten Ausgangssituation der Tests. Die Frage, die sich als nächstes stellen könnte ist, wie lange wir darauf warten müssen und wie groß die Häufigkeit ist, bis sich das Ereignis wiederholt. Dies können wir mit elementarer Wahrscheinlichkeitstheorie ermitteln.
Nehmen wir dafür an, wir untersuchen die Dauer[12], bis der gesunde Patient wiederholt positiv getestet wird. Wir bezeichnen hier der Einfachheit halber die Wahrscheinlichkeit mit $P[G \cap P] = p$. Formal können wir diese Situation als Stoppzeit[13] $T_{G \cap P}$ auf das Ereignis $G \cap P$ auffassen und es ist die erwartete Periodendauer T gesucht, also:

$$T = \mathbb{E}[T_{G \cap P}].^{14}$$

Es ist bekannt, dass diese Stoppzeit geometrisch verteilt[15] ist. Durch

[12] Das entspricht in diesem Fall der durchschnittlichen Anzahl Tests, die wir benötigen, bis wir zum wiederholten Male das beschriebene Ereignis erreichen.

[13] Stoppzeiten sind zufällig verteilte Zeiten für das erste Auftreten eines Ereignisses. Formal für ganzzahlige Zeitschritte bedeutet dies: $T_B := \min\{t \in \mathbb{Z} : Z_t \in B\}$, wobei Z_t hier der zugrundeliegende stochastische Prozess ist.

[14] Der Erwartungswert $\mathbb{E}$ einer Zufallsvariable kann aufgefasst werden als der Wert, welcher nach unendlich ofter Simualtion der Zufallsvariable im Durchschnitt vorkommt. Alternativ können wir hier auch die formale Definition mittels Summen- oder Integralschreibweise verwenden. Dafür sei wieder auf die einschlägigen Stochastiklehrbüber oder -skripte verwiesen.

[15] Eine Zufallsvariable heißt geometrisch verteilt, wenn sich positive ganzzahlige Werte annimmt und die Wahrscheinlichkeit folgendermaßen berechnet werden kann:

$$P[T = n] = (1 - p)^{n-1} \cdot p$$

Hierbei ist n eine beliebige positive ganze Zahl und p die Wahrscheinlichkeit für das Eintreten eines Ereignisses im n-ten Versuch. In diesem Beispiel wäre p also die Wahrscheinlichkeit, dass der Patient gesund und der Test positiv ist.

Anwendung der Formel für den Erwartungswert erhalten wir:

$$T = \sum_{n=1}^{\infty} n(1-p)^{n-1}p = -p\frac{d}{dp}\sum_{n=0}^{\infty}(1-p)^n, = -p\left(-\frac{1}{p^2}\right) = \frac{1}{p}.^{16}$$

Wir sehen also: Die Periodendauer T ist umgekehrt proportional zur Wahrscheinlichkeit des Ereignisses. Wenn beispielsweise ein Patient mit 5% gesund ist und positiv getestet wird, dann benötigen wir im Schnitt $1/0,05 = 20$ Tests, bis wir erstmalig getroffen haben.
Wie groß ist jetzt die Frequenz? Hier gäbe es zwei mögliche Interpretationen. Die erstere wäre die **Frequenz der erwarteten Periodendauer**, $\overline{f} = 1/E[T]$, die andere wäre die **erwartete Frequenz** $\hat{f} = E[1/T]$.
Ersteres ist leicht aus obigen Überlegungen zu ermitteln. Bei 5 % beträgt diese nämlich $\overline{f} = 1/20 = 0,05Hz$[17]. Dagegen bei der erwarteten Frequenz müssen wir ein wenig mehr ausholen. Wie oben können wir hier wieder mit der Erwartungswertformel rechnen:

$$\hat{f} = E[1/T] = \sum_{n=1}^{\infty}\frac{1}{n}(1-p)^{n-1}\cdot p = \frac{p}{1-p}\sum_{n=1}^{\infty}\frac{1}{n}(1-p)^n$$

$$= \frac{p}{1-p}\sum_{n=1}^{\infty}\int_p^1 (1-r)^{n-1}dp = \frac{p}{1-p}\int_p^1\sum_{n=0}^{\infty}(1-r)^n dr$$

$$= \frac{p}{1-p}\int_p^1 \frac{1}{r}dr = \frac{p}{1-p}[\ln(1)-\ln(p)] = -\frac{p\cdot\ln(p)}{1-p}.^{18}$$

Bei einer Wahrscheinlichkeit von $p = 5\% = 0,05$ ergibt das ungefähr $\mathbb{E}[1/T] = 0,1576Hz$, also drei mal mehr als der vorhin ermittelte Wert für die erwartete

[16] Hier habe ich einen Trick angewendet, nämlich dass $n(1-p)^{n-1}$ als Ableitung nach p geschrieben werden kann. Anschließend nutze ich aus, dass hier die Summe (oder allgemein das Integral) mit dem Ableitungsoperator vertauscht werden kann. Das folgt aus einem Satz über Reihen. Danach rechne ich die sogenannte geometrische Reihe aus, welche eine eplizite Darstellung ermöglicht. Diese leiten wir am Ende mit dem eingangs extrahierten Ableitungsorperator ab und erhalten somit das Endergebnis.

[17] Hz ist die Abkürzung für Hertz, welche nach dem gleichnamigen Physiker benannt ist. Hierbei handelt es sich einfach um die Anzahl der Ereignisse pro Sekunde. Dementsprechend ist 1 Hz gleich einem Ereignis (Umläufe, Takte,...) pro Zeiteinheit. In ist der zugrundeliegende zeitliche Richtwert Sekunde, allerdings definiere ich hier diese als einen Durchlauf.

[18] Hier sind wir ähnlich wie in der vorigen Rechnung auf das Ergebnis gekommen. Der Unterschied besteht darin, dass wir statt dem Ableitungstrick den entsprechenden Integrationstrick angewendet haben. Integral und Summenzeichen dürfen wir dann nach Fubini anwenden, nachdem wir Summierbarkeit nachgewiesen haben.

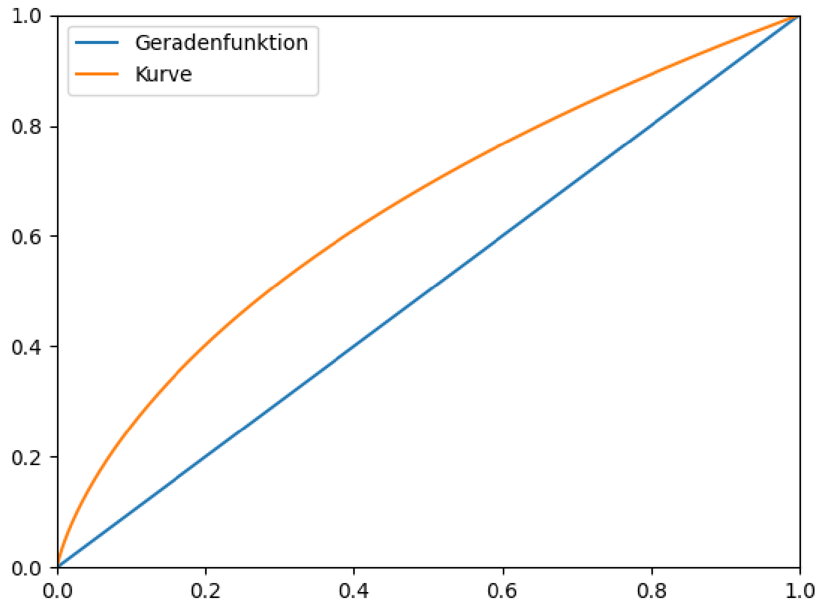

Abbildung 1: Hier sind die Frequenz der erwarteten Periodendauer (orange) und die erwartete Frequenz (blau) als Kurvenverläufe dargestellt. Auf der x-Achse ist die Wahrscheinlichkeit für das einmalige Auftreten des Ereignisses, p, dargestellt. Die y-Achse gibt die jeweilgen Werte für die Frequenz in Hz an. Diese Grafik wurde erstellt mit Python und dem Package matplotlib.

Frequenz.
Wie wir klar erkennen können, unterscheidet sich bereits in diesem einfachen Beispiel die Frequenz der erwarteten Periodendauer von der erwarteten Frequenz. Nachfolgend sind die beiden Kurvenverläufe in Abbildung 1 dargestellt.

Wie kommt der Unterschied in den Verläufen zustande? Das liegt daran, das wir bei der Frequenz der erwarteten Periodendauer einfach den Kehrwert bilden, wohingegen wir bei der erwarteten Frequenz zusätzlich noch die Schwankung der Stoppzeiten berücksichtigen. Das führt folglich zu einem anderen Funktionsverlauf.

Auslesen anhand von statistischen Daten

Haben wir umgekehrt eine statistische Messreihe von Stoppzeiten gegeben, so können wir einfach anhand dieser die zugrundeliegene Wahrscheinlichkeit des Ereignisses bestimmen. Bleiben wir bei dem obigen Beispiel mit dem medizinischen Tests und nehmen wir an, wir haben eine Messreihe gegeben(zum Beispiel waren bei 15 Test die folgenden Testnummern positiv: 2, 5, 9, 11, 14), welche angeben, wann der Test positiv ist. Hier sei also $p = P[\text{Test ist positiv}]$. Da die Ereignisse unabhängig voneinander sind, sind auch die Stoppzeiten voneinander unabhängig. Daher können wir in diesem Fall durch einfache Durchschnittbildung die erwartete Periodendauer und in diesem Fall die Wahrscheinlichkeit ermitteln.
Dafür müssen wir die Differenzen der jeweiligen Stoppzeiten bilden. Hier haben wir $T_1 = 2$, $T_2 = 5 - 2 = 3$, $T_3 = 9 - 5 = 4$, $T_4 = 11 - 9 = 2$ und $T_5 = 14 - 11 = 3$. Die geschätzte Periodendauer beträgt demnach $\hat{T} = 1/5 \sum_{i=1}^{5} T_i = 14/5 = 2,8$. Daher beträgt nach unserer Schätzung die Wahrscheinlichkeit für das Eintreten des Ereignisses $p = 1/2,8 = 0,3571 = 35,71\%$.

Übertragen auf zeitstetige Prozesse

In dem oben erwähnten Beispiel haben wir sogenannte diskrete Zeitprozesse betrachtet. Das sind Prozesse, dessen "Zeit" eine ganzzahlige, diskontinuierliche Zahl. Diese sind typischerweise geometrisch verteilt. Für zeitstetige Stoppzeiten, also Prozesse welche kontinuierliche Werte annehmen, ist eine übliche Verteilung für solche Zeiten die sogenannte Exponentialverteilung[19]. Jetzt stellt sich die Frage, was hier die Periodendauer und Frequenzen unserer Stoppzeit T_B, beispielsweise für ein Event B sind. Hier möchte ich nur eine beispielhafte Rechnung aufführen, welche die dahinterliegenden Prinzipien verdeutlichen soll.

$$
\begin{aligned}
T = \mathbb{E}[T_B] &= \int_0^\infty t\lambda \exp(-\lambda s) ds \\
&= [s \exp(-\lambda s)]_0^\infty - \int_0^\infty - \exp(-\lambda s) ds
\end{aligned}
$$

[19] Man schreibt, eine Zufallsvariable Z sei exponentialverteilt, falls $ZExp(\lambda)$ mit Parameter λ. Da es eine zeitstetige Zufallsvariable ist, können wir hier nicht für jeden Wert explizit eine Wahrscheinlichkeit zuordnen. Für die Verteilungsfunktion gilt $P[Z \leq t] = \int_0^t \lambda \exp(-\lambda s) ds = 1 - \exp(-\lambda t)$

$$= \frac{1}{\lambda}.$$[20]

Wir sehen: Die erwartete Periodendauer ist der Kehrwert des Parameters λ der Exponentialverteilung. Wir können daraus schlussfolgern, dass die erwartete Frequenz gleich diesem Parameter ist. Das bedeutet einerseits, dass wir aufgrund der Frequenz die erwartete Periodendauer ermitteln können und andererseits anhand der empirisch ermittelten Stoppzeiten die Frequenz und den Parameter schätzen können.

Fazit

In diesem Abschnitt haben wir das Lemma von Borel-Cantelli betrachtet. Es gibt zwei wesentliche Resultate: Zum einen treten voneinander unabhängige Ereignisse positiver Wahrscheinlichkeit mit Wahrscheinlichkeit 1 unendlich oft auf. Zum anderen treten Events mit positiver Wahrscheinlichkeit nur endlich oft auf, wenn deren Summe der Wahrscheinlichkeiten endlich ist.
Damit ist es beispielsweise möglich, Aussagen über statistische Tests zu treffen. Wir haben festgestellt dass unter bestimmten, durchaus plausiblen Annahmen sicher jede mögliche Konfiguration an Testergebnissen unendlich oft vorkommt.
Andererseits ist auch eine wichtige Erkenntnis aus diesem Lemma in der allgemeinen Betrachtung der Geschichte zu finden.
Es klärt sich in dieser Schrift ein Dilemma der Menschheit, nämlich ob sich die Geschichte wiederholt oder nicht. Hier stellt sich heraus, dass gewissermaßen beide Positionen korrekt sind und hier die richtige Betrachtungsweise entscheidend ist.
Entweder nehmen wir an, dass wir ein für sich geschlossenes und vom globalen historischen Prozess separiertes historisches Ereignis betrachten. In diesem Fall wiederholen sie sich wiederkehrend mit Wahrscheinlichkeit 1, unter der Voraussetzung dass diese unabhängig voneinander sind[21]. Das ist eine wichtige Einschränkung, weil das Vergessen der Geschichte (bewusst oder unbewusst) gleichbedeutend ist mit dem Entkoppeln (dem "unabhängig machen" des historischen Ereignisses mit der Gegenwart). Wenn also die Geschichte kollektiv in ausreichendem qualitativen und quantitativen Maße vergessen wird, so zieht das automatisch mit der Zeit eine Wiederholung der Geschichte im Rahmen des für sich gesetzten Ereignisses nach sich. Man hat

[20] Hier haben wir zunächst ausgenutzt, dass wir über alle möglichen Werte für T integrieren. Dann folgen partielle Integration und elementare Umrechungen.

[21] Es sei hier angemerkt, dass es eine sehr strikte Annahme ist, welche seltenst gezeigt werden kann, weil dem gewöhnlichen Beobachter Informationen fehlen.

dann sozusagen die Lehren nicht daraus gezogen[22]. Wann dieses Ereignis sich höchstwahrscheinlich wiederholt, kann anhand dieses Theorems nicht gesagt werden. Hier habe ich allerdings einige elementare statistische Methoden dargelegt, mit der wir ohne große Kenntnis eine grobe Abschätzung über die erwartete Dauer machen können. In diesem Zusammenhang spielt die Periodendauer der Stoppzeiten und die Frequenz eine wichtige Rolle.
Oder wir betrachten das Ereignis in möglichst großer Gesamtheit, im Extremfall als globaler (oder kosmischer) historischer Prozess. In diesem Fall ist die Geschichte in der ganzheitlichen Form vermutlich einzigartig und unterscheidet sich fundamental von den vorigen potenziell ähnlichen Ereignissen.
Daher ist es die Aufgabe eines jedem, die negativen geschichtlichen Ereignisse nicht vergessen zu lassen. Allein durch das Erinnern und ins Gedächtnis rufen kann bereits ein großer Beitrag dazu geleistet werden, dass sich die Konsequenzen nicht wiederholen[23]. Dazu könnte es sich anbieten, nicht die historischen Daten, Fakten und Individuen zu betrachten, sondern die Perspektive der objektiven historischen Ereignisse einzunehmen. Diese kann man katalogisieren und in einen kausalen (und somit zeitlichen) Zusammenhang stellen. Daraus können auch allgemeine historische Zyklen abgeleitet und somit ein grundlegenderes Verständnis über die geschichtlichen Entwicklungen gewonnen werden[24].
Wie allerdings die Informationen, also qualitativ in der Art und quantitativ, wie viel davon notwendig und/oder hinreichend ist um die Wiederholung der

[22]Jetzt ergibt das Zitat Kluchevskys Sinn, dass die Gesichte kein Lehrer sondern ein Aufseher ist. Nicht gelernte Lektionen werden "bestraft" und wiederholt, bis es sich "eingeprägt" hat. Man kann das unbewusst machen. Dann ist die Anzahl der Lektionen entsprechend hoch. Oder man kann bewusst anhand der gemachten Fehler lernen. Dann verringert sich die Anzahl der Lektionen, und das "Nachsitzen" fällt weg. Die Analogie zur Schule ist hier sehr treffend.

[23]Hier kommt mir das Zitat von Mohandas Karamchand Gandhi in den Sinn: "Zuerst ignorieren sie dich. Dann lachen sie dich aus. Dann bekämpfen sie dich. Und dann gewinnst du." Das könnte man wunderbar auf die geschichtlichen Ereignisse übertragen: Erst wird versucht, diese zu ignorieren, indem man diese unter den Tisch kehrt. Wenn das nicht mehr geht, wird versucht, diese zu verfälschen und ins Lächerliche zu ziehen. Gelingt auch das nicht mehr, werden aktiv physische Beweise beseitigt wie Dokumente oder Artefakte. Und erst danach hat sich das Ereignis in ihrer Essenz derart festgesetzt und verfestigt, dass es Bestandteil des kollektiven Bewusstseins geworden ist. Denke beispielsweise an große Ereignisse wie Atlantis oder Ägypten. Wie viele von den historischen Artifakten oder Relikten bleiben übrig? Und trotzdem gibt es ständig Menschen die Erinnerungen an diese Zeiten haben.

[24]Ob es bereits so eine Richtung in den Geschichtswissenschaften gibt, ist mir zum Zeitpunkt des Verfassens nicht bekannt. Wenn nicht, so könnte diese Sichtweise einen neuen Teilbereich öffnen.

Geschichte abzuwenden, ist in der Tat eine komplexere Aufgabenstellung[25].

Weiterführende Fragen

1. Wie lauten die entsprechende Periodendauer und hier in dieser Schrift besprochenen Frequenzen, wenn unsere Stoppzeit kontinuierliche Werte annimmt und sie exponentiell verteilt ist?

2. Untersuche für den unendlich oft wiederholten Würfelwurf, wie sein Ereignisraum und die Verteilungsfunktion P aussieht, wenn wir als Ereignismenge Ω die Potenzmenge[26] zugrundelegen. Wie groß sind die Periodendauer und Frequenzen dafür

   a) eine 6 zu würfeln,

   b) eine gerade Zahl zu würfeln,

   c) eine Primzahl[27] zu würfeln?

3. Recherchiere hinsichtlich der sogenannten Gammaverteilung. Wie sehen die Periodendauer und die jeweiligen Frequenzen aus, wenn unsere Stoppzeit T gammaverteilt ist?

4. Gibt es in der Gechichte historische Zyklen, welche durch immer wieder aufkommende Ereignisse, welche ineinander durch Kausalität übergehen? Versuche dazu zu recherchieren, und mithilfe der hier vorgestellten statistischen Methoden eine Schätzung darüber anzustellen, wann diese historischen Ereignisse sich wahrscheinlich wiederholen werden.

5. Wie müssten die geschichtlichen Information aufbereitet, kondensiert und präsentiert werden, damit zum einen möglichst viele negative Ereignisse so lange wie möglich verhindert und zum anderen diese möglichst effizient dargelegt werden können bei einer immer weiter wachsenden Anzahl an potenziell möglichen historischen Events?

[25] Siehe Fragestellung 5.

[26] Das sind alle mögliche Teilmengen unseres Ergebnisraumes Ω.

[27] Eine Primzahl ist eine Zahl, welche nur durch sich selbst und 1 teilbar ist. 1 zählt nicht zu den Zahlen. Bei dem Würfelwurf kommen hier nur die Zahlen 2, 3 und 5 in Frage, da 4 und 6 zusätzlich durch 2 teilbar sind.