Maren HATTEBUHR, Karlsruhe, Norbert HENZE, Karlsruhe &
Martin FRANK, Karlsruhe

Rekorde in zufälligen Permutationen – Material zur statistischen Analyse von Temperaturrekorden (Sek. II)

Der Klimawandel wird rege in der Öffentlichkeit diskutiert. Vorhersagen zur globalen Erwärmung werden teilweise heftig angezweifelt, auch weil potentielle Gegenmaßnahmen starke, ggf. als negativ empfundene, Auswirkungen auf Wirtschaft und Privatleben hätten. Die Öffentlichkeit scheint unsicher, ob die in den letzten Jahrzehnten beobachteten Klimarekorde aller Art bereits stichhaltige Indizien für einen Klimawandel sind oder lediglich temporäre Wetterphänomene darstellen. Ziel ist es, in einem gesellschaftlich und politisch höchst relevanten Kontext kritisches Denken zu fördern: Lernende analysieren und reflektieren Meldungen über den Klimawandel. So nehmen sie als mündige Bürger mit theoretisch fundiertem Wissen an einer gesellschaftsrelevanten Diskussion teil.

Ausgangspunkt ist die Aussage „Jedes der letzten drei Jahrzehnte (bezogen auf die Jahre von 1980 bis 2010) war an der Erdoberfläche sukzessive wärmer als alle vorangegangenen Jahrzehnte seit 1850." (vgl. IPCC, 2013). Darauf aufbauend erforschen Schülerinnen und Schüler in Rahmen einer Doppelstunde selbstständig die Frage, ob die auftretenden Rekorde der globalen Oberflächentemperatur in den Jahren von 1850 bis 2019 durch zufällige Schwankungen erklärt werden können. Die statistische Analyse ist in drei Teilschritte gegliedert, wobei jeweils der Modellierungskreislauf nach Blum (1985) komplett durchlaufen wird.

Schritt 1: Mathematische Modellierung von Temperaturrekorden

Zunächst machen sich die Lernenden mit dem öffentlich zugänglichen realen Datensatz HadCRUT4 vertraut. Dieser wurde durch die Climatic Research Unit (University of East Anglia) in Kooperation mit dem Hadley Centre (UK Met Office). Der Datensatz beschreibt die durchschnittliche globale jährliche Erdoberflächentemperaturanomalie (vgl. Morice et al., 2012; Hattebuhr, 2021). Zur Analyse der Aussage des IPCC werden die Temperaturdaten in Dekaden zusammengefasst, innerhalb derer jeweils Mittelwert gebildet wird. Desweiteren wird der Begriff Rekordjahr bzw. -periode eingeführt: Rekordperioden sind Perioden, in denen die (durchschnittliche) Temperaturanomalie größer ist als in allen Perioden zuvor (seit Beginn der Messdatenaufzeichnung). Die Schülerinnen und Schüler mathematisieren den Begriff

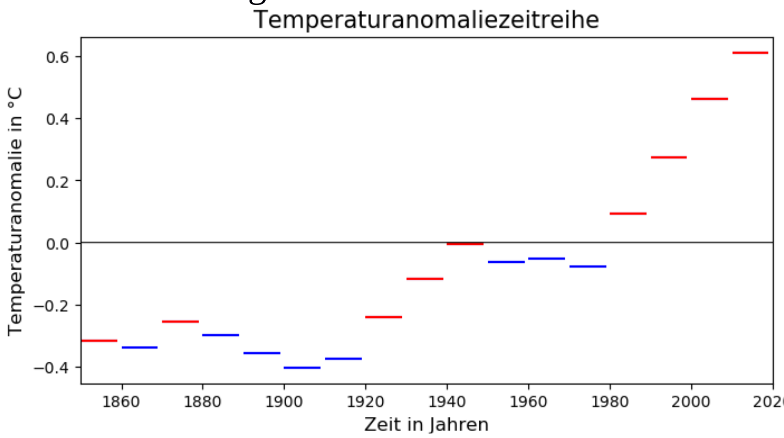

Abb. 1: Temperaturperioden (blau) mit Rekorden (rot)

In: Kerstin Hein, Cathleen Heil, Silke Ruwisch & Susanne Prediger (Hrsg.). Beiträge zum Mathematikunterricht 2021. Münster: WTM Verlag. https://doi.org/10.37626/GA9783959871846.0
Online unter https://eldorado.tu-dortmund.de/handle/2003/30630

und wenden ihn auf die Datenreihe an. Die Daten (blau) werden entsprechend des aufgestellten Algorithmus visualisiert und die Rekordperioden rot eingefärbt (vgl. Abb. 1). Diese Rückmeldung dient den Lernenden zur selbstständigen Validierung des (mathematischen) Ergebnisses. Damit können sie die These des IPCC-Berichts nicht nur bestätigen, sondern sogar erweitern: Jedes der letzten vier Jahrzehnte war an der Erdoberfläche wärmer als alle vorangegangenen Jahrzehnte seit 1850. Das ist zuvor (im Rahmen der vorliegenden Daten) noch nie aufgetreten."

Schritt 2: Anzahl der Rekorde in rein zufälligen Permutationen (Laplace-Modell) und ihr Erwartungswert

In einem zweiten Schritt untersuchen die Schülerinnen und Schüler die Verteilung von Rekorden unter einem Laplace-Modell. Dieser lässt sich auf zwei Niveaustufen erarbeiten. Die Temperaturanomaliewerte der Dekaden werden einfachheitshalber durch die Zahlen 1 bis 17 repräsentiert. Die Lernenden untersuchen, wie groß die Wahrscheinlichkeit ist, dass beim zufälligen Vertauschen (Permutieren) von 17 Zahlen insgesamt $k = 1, \dots, 17$ Rekorde auftreten. Auf dem ersten Niveau erzeugen sie dazu mehrfach zufällige Permutationen und bestimmen jeweils die Anzahl der Rekorde. Steht R_{17} für die zufällige Rekordanzahl, so nähert sich nach dem Gesetz der großen Zahlen die relative Häufigkeit des Ereignisses $R_{17} = k$ immer weiter der Wahrscheinlichkeit $P(R_{17} = k)$ an, die in Abbildung 2 dargestellt ist. Dieses mathematische Ergebnis gilt es nun zu interpretieren: Die Wahrscheinlichkeit, dass mindestens neun Rekorde auftreten, liegt bei 0,05%. Das Auftreten von Rekordperioden wie im aktuellen Datensatz ist sehr unwahrscheinlich und – da hier ein Laplace-Modell angenommen wurde – kaum durch den Zufall erklärbar. Für das Auftreten von 10 bis 17 Rekorden sind die Wahrscheinlichkeiten bereits so gering, dass sie im Säulendiagramm nicht mehr sichtbar sind. Am häufigsten treten mit einer Wahrscheinlichkeit von etwa 29% drei Rekorde auf. Diese Anzahl an Rekorden liegt in der Nähe des Erwartungswertes der Rekordverteilung unter einem Laplace-Modell.

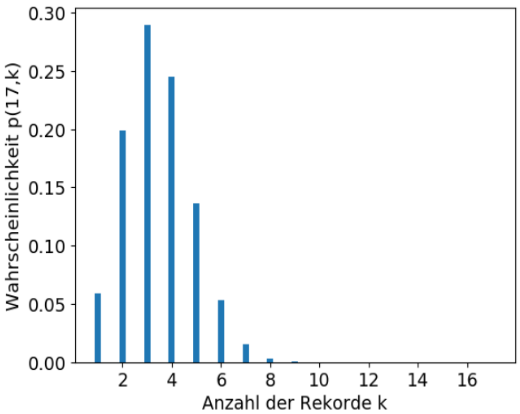

Abb. 2: Wahrscheinlichkeiten der Anzahl der Rekorde

Auf der zweiten Niveaustufe ist über das Schulwissen hinausgehende Mathematik nötig. Diese wird eingeführt und problembezogen angewandt. Damit ist dient dieses Niveau vor allem der Differenzierung und Forderung aller Mathematik-Begeisterten. Zunächst erkunden die Schülerinnen und Schüler, wie wahrscheinlich es ist, dass in einer rein zufälligen Vertauschung der Zahlen von 1 bis 17 an der j-ten Stelle ein Rekord auftritt (Ereignis A_j). Die Wahrscheinlichkeiten können schrittweise bestimmt werden: An der ersten Stelle steht – nach Definition – ein Rekord, es gilt also $P(A_1) = 1$. Zur Bestimmung von $P(A_2)$ greift eine Symmetrieüberlegung: Entweder steht an der ersten Stelle die größere der beiden Zahlen, dann ist an der zweiten Stelle kein Rekord, oder an der zweiten Stelle steht die größere Zahl, dann liegt ein Rekord an der zweiten Stelle vor. Beide Fälle sind gleich wahrscheinlich, da der reine Zufall wirkt. Es gilt $P(A_2) = 1/2$. So kann nun für alle weiteren Ereignisse argumentiert werden, und wir folgern allgemein $P(A_j) = 1/j$ für alle $j = 1, \dots, 17$. Das eigentliche Interesse gilt aber der Wahrscheinlichkeit, dass sich bei einer zufälligen Permutation der Zahlen 1 bis 17 mindestens neun Rekorde einstellen. Wir schreiben allgemein für die Wahrscheinlichkeit, dass sich in einer rein zufälligen Permutation der Zahlen von 1 bis n genau k Rekorde ergeben $P(R_n = k) =: p(n, k)$. Die Lernenden betrachten zunächst den Fall, dass *genau ein Rekord* auftritt. Dieser kann nur eintreten, wenn die größte Zahl (hier: 17) an der ersten Stelle steht. Aus Symmetriegründen gilt daher $p(17,1) = 1/17$. Als nächstes beleuchten die Schülerinnen und Schüler den Fall, dass es *genau 17 Rekorde* gibt, d. h., dass an jeder Stelle ein Rekord steht. Dieser Fall tritt genau dann ein, wenn die Zahlen von 1 bis 17 der Reihe nach aufsteigend angeordnet sind. Diese Wahrscheinlichkeit ist durch $p(17,17) = 1/17!$ gegeben. Der gesuchte Fall kann nun über einen rekursiven Ansatz erhalten werden, in dem $p(17,9)$ über die beiden zuvor erörterten Fälle ausgedrückt wird. Hier werden die Schülerinnen und Schüler angeleitet, zunächst den Fall zu betrachten, dass *an der letzten Stelle ein Rekord* auftritt. Mit dem obigen Wissen ist diese Wahrscheinlichkeit durch $P(A_{17}) = 1/17$ gegeben. Außerdem sind nun $k - 1 = 8$ weitere Rekorde auf die vorherigen Plätze zu verteilen. Diese Wahrscheinlichkeit, dass sich bei einer zufälligen Permutation der Zahlen von 1 bis 16 genau 8 Rekorde ergeben, ist $p(16,8)$. Im nächsten Fall betrachten die Schülerinnen und Schüler die komplementäre Situation, dass *an der letzten Stelle kein Rekord* auftritt. Hierfür ist die Wahrscheinlichkeit gegeben durch $P(\bar{A}_{17}) = 16/(16 + 1)$. Alle 9 Rekorde müssen auf den 16 Plätzen davor verteilt werden: Die Wahrscheinlichkeit dafür beträgt $p(16,9)$. Unter Anwendung des Satzes von der totalen Wahrscheinlichkeit können nun diese aufgezeigten Fälle zusammengefasst werden zu

$$P(R_{17} = 9) = p(17,9) = \frac{1}{17} \cdot p(16,8) + \frac{16}{16 + 1} \cdot p(16,9) = 0{,}00052.$$

Die benötigte Unabhängigkeit der Ereignisse kann gezeigt werden (vgl. Henze et al., 2021). Das theoriegeleitete Ergebnis bestätigt das experimentell gewonnene Ergebnis des ersten Niveaus. Um nun den Erwartungswert zu bestimmen ist normalerweise Wissen über die Verteilung des Auftretens der Rekorde nötig. Dieses kann geschickt umgangen werden, indem man zunächst die Anzahl der Rekorde R_{17} durch die Indikatorvariablen $1\{A_j\}$ ausdrückt und anschließend die Additivität der Erwartungswertbildung ausnutzt. Es folgt $E(R_{17}) = 1 + 1/2 + \cdots + 1/17$. Auf einem Informationsblatt wird diese Herangehensweise erläutert, und die Lernenden wenden ihre neu erworbenen Kenntnisse zur Abschätzung des Erwartungswerts an. Sie erhalten mit dem Wert von $E(R_{17}) \approx 3{,}44$ eine gute Prognose für den Durchschnitt der erhaltenen Rekordanzahlen auf lange Sicht. Die

Realität mit neun Rekorden liegt weit, genauer gesagt mehr als vier Standardabweichungen, vom Erwartungswert entfernt.

Abschließend steht für alle die Berechnung der Wahrscheinlichkeit im Blickpunkt, dass die vier letzten Jahrzehnte die wärmsten der letzten 170 Jahre sind. Auch in dieser Aufgabe wird ein rekursives Vorgehen zur Lösung genutzt.

Fazit

In allen Schritten ist die reale Situation Ausgangs- und Endpunkt der statistischen Analyse. Die selbstständig erforschten Methoden lassen sich auf die Analyse der ursprünglichen Datenreihe mit jährlichen Temperaturanomaliewerten übertragen. Alle Beteiligten erleben aktiv das wesentliche Vorgehen in der mathematischen Modellierung. Insgesamt lässt sich diese Problemfrage nach Maaß (2010) in die Kategorie der authentischen – sowohl in Bezug auf die Fragestellung, als auch auf die angewandten Methoden – mathematischen Modellierung einordnen. Die Ergebnisse führen nicht nur zu einer fundierten Einschätzung des Themas Klimawandel, sondern zeigen auch die Relevanz von Mathematik für unsere Gesellschaft: Schulmathematische Inhalte erhalten eine Bedeutung und können damit Heranwachsende motivieren, sich mit Mathematik zu beschäftigen. Das für eine Doppelstunde ausgelegte digital umgesetzte Material eignet sich durch seine Differenzierungsmöglichkeiten (verschiedene Niveaus, Zusatzaufgaben, …) und individuellen, automatischen Rückmeldungen zu den Lösungseingaben der Schülerinnen und Schüler auch für ein heterogenes Lernklientel. Es ist online über einen Webbrowser auf einer Cloud-Plattform des Karlsruher Instituts für Technologie (KIT) zugänglich und kann daher direkt sowohl im Präsenz- als auch Fernunterricht eingesetzt werden (s. www.cammp.online/214.php).

Literatur

Blum, W. (1985). Anwendungsorientierter Mathematikunterricht in der didaktischen Diskussion? In Kahle et al. (Hrsg.), Mathematische Semesterberichte. Zur Pflege des Zusammenhangs zwischen Schule und Universität (Bd. 32, 2, S. 195–232). Göttingen: Vandenhoeck & Ruprecht.

Hattebuhr, M. (2021): Gibt es den Klimawandel wirklich? – Statistische Analyse von realen Temperaturdaten. In: Neue Materialien für einen realitätsbezogenen Mathematikunterricht 9 (ISTRON), Frank, M. & Roeckerath, C. (Hrsg.). Springer-Verlag GmbH, Heidelberg (im Druck).

Henze, N., Müller, K. & Schilling, J. (2021): Stochastik rezeptfrei unterrichten – Anregungen für spannende Lehre über den Zufall. Springer Spektrum, Heidelberg.

IPCC (2013). Climate Change 2013 – The Physical Science Basis. New York, USA: Cambridge University Press.

Maaß, K. (2010). Classification Scheme for Modelling Tasks. JMD, 31(2), 285–311.

Morice, C. P., Kennedy, J. J., Rayner, N. A. & Jones, P. D. (2012). Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 dataset, J. Geophys. Res. and 117, D08101. https://www.metoffice.gov.uk/hadobs/hadcrut4/; Zugegriffen: 03. April 2020. doi:10.1029/2011JD017187