

Bedingte Wahrscheinlichkeiten – Autokorrektur und Autovervollständigung

Mona Hitzenauer



© Ridofranz / iStock / Getty Images Plus

Und wieder einen Termin verpasst. Schnell noch eine Nachricht geschrieben und – oops – vertippt. Ein „t“ zu viel, und aus der Ankündigung, später noch nachzukommen, wird der Satz „Ich komme später noch nacht!“. Aber damit noch nicht genug, denn die Autokorrektur verschlimmbessert das zu „Ich komme später noch nicht!“ und ärgert uns mal wieder. Doch wie passiert das?

Lassen Sie Ihre Schüler und Schülerinnen erforschen, wie Autokorrektur und Autovervollständigungsprogramme arbeiten und warum dafür bedingte Wahrscheinlichkeiten eine große Rolle spielen.

Bedingte Wahrscheinlichkeiten – Autokorrektur und Autovervollständigung

Oberstufe (grundlegend)

Mona Hitzenauer

Hinweise	1
M1 Autokorrektur – Mengendiagramme	3
M2 Autokorrektur – Vierfeldertafel	5
M3 Autovervollständigung – bedingte Wahrscheinlichkeiten	6
M4 Übungsaufgaben	9
Lösungen	10

© RAABE 2023

Die Schülerinnen und Schüler lernen:

wie Autokorrektur bzw. Autovervollständigungsprogramme arbeiten und erkennen, warum sie bei jeder Nachricht, die sie mit dem Smartphone abschicken, ohne es zu merken von grundlegenden Konzepten der Wahrscheinlichkeitsrechnung profitieren.

Damit erkennen die Jugendlichen den Nutzen und die Anwendung der Mathematik, der sie in ihrem Schulalltag begegnen. Die Schülerinnen und Schüler lernen in dieser Thematik Vierfeldertafeln kennen, gehen mit Mengendiagrammen um und bestimmen verschiedene Ereigniswahrscheinlichkeiten und bedingte Wahrscheinlichkeiten. Sie unterscheiden dabei scharf zwischen den Wahrscheinlichkeiten $P(A \cap B)$, $P_A(B)$ und $P_B(A)$.

Autokorrektur – Vierfeldertafel

M2

In der Tabelle kann man statt Mengendarstellungen auch die entsprechenden absoluten oder relativen Häufigkeiten sowie Wahrscheinlichkeiten eintragen. Solche Tabellen nennt man Vierfeldertafeln, da besonders die vier Felder bzw. Zellen in der Mitte oft von Interesse sind.

Aufgabe 1

Das Autokorrektur-Programm hat nun folgende relativen Häufigkeiten berechnet: Das Wort „noch“ hat insgesamt eine relative Häufigkeit von 0,010 %, beim Wort „nicht“ beträgt sie 0,015 %. Die Kombination „noch“ mit „nicht“ hat eine relative Häufigkeit von 0,009 %.

Ergänzen Sie in der Vierfeldertafel die relativen Häufigkeiten h in Prozent.

	noch	nicht	
nicht			
nicht			

Folgern Sie aus der Tabelle eine Aussage über die relative Häufigkeit der Wortkombination „noch nicht“. Begründen Sie damit, warum das Programm bei der Korrektur „noch nicht“ geblieben ist.

Aufgabe 2

Verallgemeinern Sie die Aufgabe, indem Sie zwei Ereignisse A und B und deren Wahrscheinlichkeiten (bzw. Gegenwahrscheinlichkeiten) betrachten. Das gleichzeitige Auftreten entspricht der Schnittmenge der Ereignismengen.

Stellen Sie damit eine allgemeine Vierfeldertafel auf.

Leiten Sie aus den Zeilen bzw. Spalten allgemeine Formeln ab.

M3 Autovervollständigung – bedingte Wahrscheinlichkeiten

Neben dem gleichzeitigen Auftreten von Wortkombinationen testen die meisten Tipp-hilfen, wie wahrscheinlich es ist, dass nach einem Wort ein anderes auftaucht. Das nennt man **bedingte Wahrscheinlichkeit**. Die Bedingung – das erste getippte Wort – ist dabei schon sicher eingetreten.



Achtung: Die Wahrscheinlichkeit, dass ein Ereignis B eintritt, wenn ein Ereignis A sicher zuvor eingetreten ist, nennt man bedingte Wahrscheinlichkeit $P_A(B)$.

Wir nehmen an, das Programm findet insgesamt folgende Wortkombinationen in den Texten:

schon so	in der	so nicht	so ein
gar nicht	schon noch	nicht das	so wie
heute ist	aber das	nur noch	nicht so
so nicht	gar nicht	doch schon	noch so
schon wieder	nicht so	nicht so	gar nicht
sogar mit	nicht nur	sogar schon	so wie
jetzt bin	sogar heute	aber nicht	nicht nur
wieder ein	gar kein	so nicht	nicht so
in den	so nicht	heute nicht	nicht gar

Aufgabe 1

Bestimmen Sie die relative Häufigkeit des gemeinsamen Auftretens der Wörter „gar“ und „nicht“. Interpretieren Sie diese relative Häufigkeit $h(\text{gar} \cap \text{nicht})$ im Sachzusammenhang.

Aufgabe 2

Suchen Sie nun aus der Tabelle alle Wortkombinationen heraus, die „nicht“ enthalten und bestimmen Sie auf deren Grundlage dann die relative Häufigkeit der Wortkombinationen mit den Wörtern „gar“ und „nicht“. Bezeichnen Sie diese relative Häufigkeit mit $h_{\text{nicht}}(\text{gar})$. Beschreiben Sie, was diese relative Häufigkeit aussagt.

Entscheiden Sie begründet, ob das Programm bei der Eingabe „nicht“ die Wortkombination „gar nicht“ vorschlagen sollte.

Aufgabe 3

Kehren Sie zur Gesamtliste aller Wortkombinationen zurück. Suchen Sie nun die Kombinationen mit „gar“ heraus und bestimmen Sie auf dieser Grundlage die relative Häufigkeit der Wortkombinationen mit den Wörtern „gar“ und „nicht“. Bezeichnen Sie diese relative Häufigkeit mit h_{gar} (nicht).

Beschreiben Sie, was diese relative Häufigkeit aussagt.

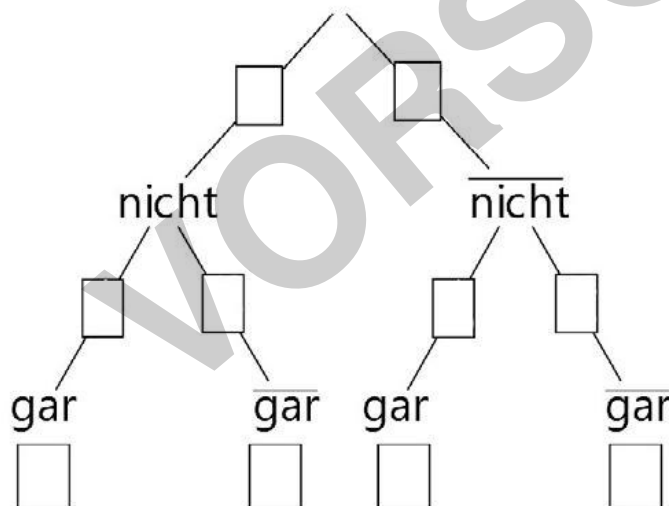
Entscheiden Sie auch hier mit Begründung, ob das Programm bei der Eingabe „gar“ die Wortkombination „gar nicht“ vorschlagen sollte.

Aufgabe 4

Untersuchen Sie die Tabelle der Wortkombinationen erneut. Bestimmen Sie, welche Wortkombination das Programm bei der Eingabe von „nicht“ vorschlagen sollte.

Aufgabe 5

Ergänzen Sie die Baumdiagramme auf dieser und der nächsten Seite. Verwenden Sie dabei Ihre Ergebnisse aus den Aufgaben 1) bis 3). Die ermittelten relativen Häufigkeiten können Sie hierbei als Wahrscheinlichkeiten verwenden.



Grafik: Mona Hitznauer

