Stochastik 3: W-Bäume, bedingte Wahrscheinlichkeiten und Vierfeldertafeln

1 40% aller zugelassenen PKW eines Autoherstellers sind innerhalb der letzten 5 Jahre hergestellt worden. Von diesen wurde die Abgas-Software bei 75% manipuliert. Von den älteren Autos wurde die Software nur in $10\,\%$ aller Fälle manipuliert.[1]
Es wird zufällig ein Auto des Herstellers ausgewählt. Es seien $J =$ »Auto ist jünger als 5 Jahre« und $M =$ »Software wurde manipuliert«.
Geben Sie die folgenden Wahrscheinlichkeiten an und deuten Sie sie im Sachzusammenhang:
$\mathrm{P}[M]$, $\mathrm{P}[J]$, $\mathrm{P}\left[\overline{M} \mid J\right]$, $\mathrm{P}\left[M \cap \overline{J}\right]$, $\mathrm{P}[M \cup J]$, $\mathrm{P}\left[M \mid \overline{J}\right]$.
Hinweis: Für viele dieser Wahrscheinlichkeiten ist keine Rechnung nötig.

2 Es seien A und B zwei Ereignisse.

(a) Stellen Sie die Situation als Wahrscheinlichkeitsbaum dar.

(b) Begründen Sie den sog. **Satz von der totalen Wahrscheinlichkeit** mit Hilfe des Baumes aus (a):

$$\mathrm{P}[B] = \mathrm{P}[A] \cdot \mathrm{P}[B \mid A] + \mathrm{P}\left[\overline{A}\right] \cdot \mathrm{P}\left[B \mid \overline{A}\right]$$

3 **Wirtschaftswissenschaften beliebtestes Studienfach bei Frauen und Männern**

Im Wintersemester 2009/10 sind $46{,}4\%$ der Studierenden an nordrheinwestfälischen Hochschulen Frauen; von diesen haben $16{,}5\%$ das Fach Wirtschaftswissenschaften belegt. Bei den Männern entschieden sich sogar $18{,}3\%$ für diesen Fachbereich.

(a) Stellen Sie die Informationen des Zeitungsartikels in Form eines Baumdiagramms dar unter Verwendung der Ereignisse

F : »Student ist eine Frau« und W : »Student belegt Wirtschaftswissenschaften«

(b) Geben Sie die im Artikel genannten Prozentzahlen als (bedingte) Wahrscheinlichkeiten an unter Verwendung der Ereignisse F und W.

(c) Bestimmen Sie $\mathrm{P}[W]$. [Kontroll-Ergebnis: $17{,}4648\%$]

(d) Zeigen Sie, dass $\mathrm{P}[F \mid W] \approx 43{,}84\%$. Interpretieren Sie diese Zahl im Sachzusammenhang.
Hinweis: Zeichnen Sie den inversen Baum, bei dem auf der ersten Stufe die Ereignisse W und $\overline{W}$ vorkommen und auf der zweiten Stufe die Ereignisse F und $\overline{F}$. Beachten Sie, dass die Pfad-Wahrscheinlichkeiten gleich bleiben!

[1] Diese Zahlen sind frei erfunden...

4 Die Schule *l'école du futur* hat 720 SchülerInnen, davon besuchen seit einem Jahr $5\,\%$ nachmittags die *Serlo Lab School*, wo sie selbständig Mathe lernen.[2]
Die *Vector*-Stiftung möchte Serlo weiterhin fördern, wenn die Ziele der Lab School erfüllt werden. Daher will die Stiftung wissen, ob die Ereignisse

L: SchülerIn besucht die Lab School.

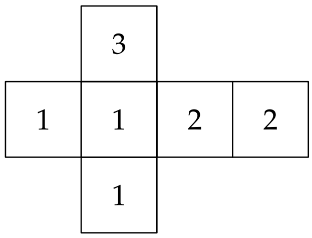
S: SchülerIn kann selbständig lernen und traut sich zu, die Schule zu schaffen.

stochastisch abhängig sind.
In einer Umfrage an der Schule ergibt sich folgendes Bild:

- 32 SchülerInnen, die in die Lab School gehen, schätzen sich als selbstständig ein und sagen, sie können Mathe meistern.
- 272 der SchülerInnen, die nicht die Lab School besuchen, wissen oft nicht weiter, wenn sie etwas nicht verstehen und haben Angst, wegen Mathe die Schule nicht zu schaffen.

Entscheiden Sie, ob die *Vector*-Stiftung Serlo weiterhin fördern sollte.

5 Bei einem Experiment werden zwei spezielle Würfel mit dem nebenstehenden Netz geworfen. Untersuchen Sie jeweils, ob die beiden Ereignisse A und B unabhängig sind.



(a) $A =$ »1. Würfel zeigt eine 3«, $B =$ »2. Würfel zeigt eine 2«

(b) $A =$ »1. Würfel zeigt eine 3«, $B =$ »Die Augensumme ist 5«

(c) $A =$ »Augensumme ist gerade«, $B =$ »Die Augensumme beträgt mindestens 4«

6 Begründen Sie anhand eines W-Baumes oder einer Vierfeldertafel:

(a) $\mathrm{P}\left[A \cup \overline{B}\right] = \mathrm{P}\left[A \cap B\right] + \mathrm{P}\left[A \cap \overline{B}\right] + \mathrm{P}\left[\overline{A} \cap \overline{B}\right]$

(b) $\mathrm{P}\left[A \cup B\right] = \mathrm{P}\left[A\right] + \mathrm{P}\left[B\right] - \mathrm{P}\left[A \cap B\right]$

(c) $\mathrm{P}\left[\overline{A} \cup \overline{B}\right] = \mathrm{P}\left[\overline{A} \cap \overline{B}\right] + \mathrm{P}\left[\overline{A} \cap B\right] + \mathrm{P}\left[A \cap \overline{B}\right]$

(d) $\mathrm{P}\left[A \cap \overline{B}\right] = 1 - \mathrm{P}\left[\overline{A} \cup B\right]$

7 Für zwei Ereignisse A und B gilt:

$$\mathrm{P}\left[A \cap B\right] = 0{,}3 \qquad \mathrm{P}\left[A \cup \overline{B}\right] = 0{,}8 \qquad \mathrm{P}\left[\overline{A} \cap \overline{B}\right] = 0{,}1$$

Zeichnen Sie die zugehörige Vierfeldertafel sowie die beiden möglichen Baumdiagramme. Bestimmen Sie alle dazu benötigten Wahrscheinlichkeiten.

8 Lösen Sie Aufgabe 7 erneut für

[2] de.serlo.org

(a) $\mathrm{P}[A \mid B] = 0{,}7$, $\mathrm{P}\left[\overline{A} \mid \overline{B}\right] = 0{,}4$, $\mathrm{P}[A \cap B] = 0{,}2$

(b) $\mathrm{P}\left[A \mid \overline{B}\right] = 0{,}6$, $\mathrm{P}\left[A \cup \overline{B}\right] = 0{,}5$, $\mathrm{P}[B] = 0{,}8$

(c) $\mathrm{P}[A \cup B] = 0{,}9$, $\mathrm{P}\left[\overline{A} \cup B\right] = 0{,}5$, $\mathrm{P}\left[A \cup \overline{B}\right] = 0{,}7$

Hinweis zu (c): Zeichnen Sie die Vierfeldertafel, stellen Sie ein Gleichungssystem für die gesuchten Wahrscheinlichkeiten auf und lösen Sie es mit dem TR.

9 Samuel bringt drei Tüten mit Kreppeln mit zur Fassenachts-Mathestunde. Tüte 1 enthält 8 ungefüllte Kreppel, Tüte 2 enthält 4 gefüllte und 4 ungefüllte Kreppel und Tüte 3 enthält 6 gefüllte und 2 ungefüllte Kreppel.

(a) Stellen Sie den Sachverhalt mit Hilfe eines Baum-Diagramms unter Verwendung von geeigneten Ereignissen dar.

(b) Nun sucht sich Marie blind eine Tüte aus und zieht einen Kreppel heraus.

(1) Berechnen Sie die Wahrscheinlichkeit dafür, dass der Kreppel gefüllt ist.

(2) Der Kreppel ist gefüllt. Berechnen Sie die Wahrscheinlichkeit dafür, dass er aus Tüte 2 stammt.

(c) Die folgende **6-Felder-Tafel** zeigt eine andere Möglichkeit, die geschilderte Situation darzustellen:

	gefüllt	ungefüllt	**gesamt**
Tüte 1	0	$\frac{1}{3}$	$\frac{1}{3}$
Tüte 2	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$
Tüte 3	$\frac{1}{4}$	$\frac{1}{12}$	$\frac{1}{3}$
gesamt	$\frac{5}{12}$	$\frac{7}{12}$	**1**

Erläutern Sie die Bedeutung der einzelnen Felder der Tabelle.

10 Über eine Internet-Plattform werden Wohnhäuser, Ferienhäuser und Eigentumswohnungen zum Verkauf angeboten. Da der Kauf einer Immobilie sehr kostspielig sein kann, kommt es manchmal zu der Unnannehmlichkeit, dass ein Kauf abgebrochen werden muss, da die Bank einen Kredit verweigert.
Im vergangenen Jahr lagen folgende Zahlen vor:

- $7\,\%$ aller angebotenen Immobilien waren Ferienhäuser.
- Bei 15% der Wohnhäuser musste der Kauf abgebrochen werden.
- Bei $23{,}8\,\%$ aller Fälle handelte es sich um Wohnhäuser, bei denen der Kauf nicht abgebrochen werden musste.

- Bei den Ferienhäusern musste kein einziger Kauf abgebrochen werden.
- Insgesamt wurden $6{,}15\,\%$ aller Käufe abgebrochen.

(a) Stellen Sie die beschriebene Situation als vollständigen Wahrscheinlichkeitsbaum oder als vollständige Mehrfeldertafel dar.

(b) Untersuchen Sie, ob die Tatsache, dass ein Kauf abgebrochen wird, stochastisch unabhängig davon ist, welche Art Immobilie gekauft werden sollte.

(c) Ein Kauf musste abgebrochen werden. Bestimmen Sie die Wahrscheinlichkeit dafür, dass es sich bei dem Objekt um ein Wohnhaus handelte.

(d) Ein Makler hat aktuell 7 Kunden, die daran interessiert sind, jeweils ein Wohnhaus zu kaufen. (Es handelt sich dabei um unterschiedliche Wohnhäuser.) Aufgrund der obigen Zahlen rechnet er damit, dass einer dieser Käufe abgebrochen werden wird.
Beurteilen Sie die Schätzung des Maklers sowie deren Verlässlichkeit.

11 Im Zuge der Abiturvorbereitungen haben die 90 Abiturient*innen der Marienschule schon vor einiger Zeit einen Abiball-Ausschuss und einen Abistreich-Ausschuss gebildet. 10% aller Abiturient*innen sind im Abiball-Ausschuss und $2{,}2\%$ sind in beiden Ausschüssen. Unter den Schüler*innen, die nicht im Abiball-Ausschuss sind, sind 21% im Abistreich-Ausschuss.

(a) Stellen Sie den geschilderten Sachverhalt als Wahrscheinlichkeitsbaum und als Vierfeldertafel dar.

(b) Bestimmen Sie die Anzahl der Mitglieder des Abistreich-Ausschusses.
[Zur Kontrolle: $\mathrm{P}\left[\text{»Mitglied im Abistreich-Ausschuss«}\right] = 21{,}1\%$]

(c) Eine zufällig gewählte Schülerin ist im Abistreich-Ausschuss. Bestimmen Sie die Wahrscheinlichkeit dafür, dass sie auch im Abiball-Ausschuss ist.

12 Eine Person gilt als *hochbegabt*, wenn der Intelligenzquotient über 130 Punkten liegt. Etwa $2\,\%$ der Bevölkerung sind hochbegabt. Dagegen wird eine Person mit einem IQ unter 85 als *lernbehindert* bezeichnet. Dies sind etwa $16\,\%$ der Bevölkerung.

(a) Hochbegabte Schulkinder fallen oft im Schulunterricht dadurch auf, das sie häufig stören oder unaufmerksam sind. Es seien

H: Das Kind ist hochbegabt
S: Das Kind stört häufig oder ist häufig unaufmerksam

Es ist bekannt, dass $80\,\%$ aller hochbegabten Schulkinder häufig im Unterricht stören oder unaufmerksam sind. Der Anteil aller Schulkinder, die weder hochbegabt sind noch häufig stören oder unaufmerksam sind, liegt bei $60\,\%$.[3]

[3] Diese Zahlen sind frei erfunden!

(1) Stellen Sie die beschriebene Situation als Wahrscheinlichkeitsbaum oder als 4-Feldertafel dar.

(2) Bestätigen Sie, dass $\mathrm{P}[S] = 39{,}6\,\%$.

(3) Untersuchen Sie die Ereignisse H und S auf stochastische Unabhängigkeit.

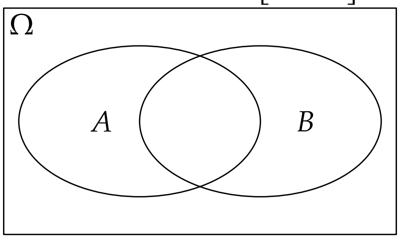
(4) Beurteilen Sie folgende Aussage eines Vaters anhand einer geeigneten Rechnung: »Da mein Kind häufig stört, ist es sehr wahrscheinlich hochbegabt.«

(b) Für eine wissenschaftliche Untersuchung werden Menschen mit Lernbehinderung benötigt. Dazu werden zufällig 20 Personen eingeladen. Erläutern Sie die folgende Rechnung und deuten Sie das Ergebnis $0{,}97$ im Sachzusammenhang.

(1) $0{,}84^{20} \approx 0{,}03$

(2) $1 - 0{,}03 = 0{,}97$

13 Begründen Sie die Pfadregel $\mathrm{P}[B] = \mathrm{P}[A \cap B] + \mathrm{P}\left[\overline{A} \cap B\right]$ anhand der folgenden Zeichnung:



Entscheiden Sie, ob die Formel auch gilt, wenn sich A und B nicht überschneiden.

14 Beweisen Sie: Wenn $\mathrm{P}[A] > 0$ ist, dann gilt

$$\mathrm{P}[B \mid A] = \frac{\mathrm{P}[A \cap B]}{\mathrm{P}[A]}$$

15 Beweisen Sie den sogenannten **Satz von Bayes**: Für zwei Ereignisse A und B mit $\mathrm{P}[A] > 0$ und $\mathrm{P}[B] > 0$ gilt

$$\mathrm{P}[A \mid B] = \mathrm{P}[B \mid A] \cdot \frac{\mathrm{P}[A]}{\mathrm{P}[B]}$$

16 Beweisen Sie die folgende Aussage für zwei Ereignisse A und B mit $\mathrm{P}[A] > 0$ und $\mathrm{P}B > 0$: Wenn A und B unabhängig sind, dann sind auch $\overline{A}$ und $\overline{B}$ unabhängig.

17 Folgender Bericht stammt von `apotheken-umschau.de` :

Viele falsch-positive Befunde bei hoher Spezifität und Sensitivität

Wie zuverlässig ein Test arbeitet, geben Hersteller mit Werten für die Spezifität und Sensitivität an. Die Sensitivität steht für die Erkennungsrate, also den Prozentsatz der Betroffenen, bei denen die Infektion tatsächlich erkannt wird. Ein Test mit einer Sensitivität von

95 Prozent identifiziert 95 von 100 Infektionen und 5 nicht. Für den Roche-Test *Elecsys Anti-Sars-CoV-2* beträgt die Sensitivität nach Angaben des Unternehmens 100 Prozent. Verwaltungsratspräsident Franz sprach am Montag von einem völlig "neuen Qualitätsniveau".
Die Spezifität sagt aus, wie viele Gesunde, die definitiv nicht mit dem Virus infiziert sind oder waren, von dem Test auch tatsächlich als gesund erkannt werden. Ein Test mit einer Spezifität von 95 Prozent liefert bei 5 von 100 nicht infizierten Menschen fälschlicherweise ein positives Ergebnis. Für *Elecsys Anti-Sars-CoV-2* gibt Roche die Spezifität mit $99{,}8$ Prozent an. Bei hoher Sensitivität und geringer Spezifität kann es viele falsch-positive Befunde geben.
Quelle: Apotheken-Umschau.de (27.09.2020)

Laut Wikipedia gibt es aktuell 23.373 Menschen in Deutschland, die an Corona erkrankt sind. In Deutschland leben insgesamt etwa $83{,}2$ Millionen Menschen.

(a) Ein Corona-Test habe eine Sensitivität und eine Spezifität von jeweils $95\,\%$.

(1) Stellen Sie die Situation als vollständigen W-Baum dar.

(2) Berechnen Sie die Wahrscheinlichkeit dafür, dass ein Test bei einer zufälligen Testperson positiv ist. [Zur Kontrolle: $\approx 5{,}027\,\%$]

(3) Berechnen Sie die Wahrscheinlichkeit dafür, dass eine Person wirklich krank ist, wenn sie ein positives Testergebnis erhält. [Zur Kontrolle: $\approx 0{,}531\,\%$]

(4) Berechnen Sie die Wahrscheinlichkeit dafür, dass eine Person wirklich gesund ist, wenn sie ein negatives Testergebnis erhält. [Zur Kontrolle: $\approx 99{,}9958\,\%$]

(b) Berechnen Sie die Wahrscheinlichkeiten aus der vorangegangenen Aufgabe noch einmal für den im Eingangstext beschriebenen Roche-Test.

(c) Man spricht von einem »falsch-positiven« Testergebnis, wenn der Test eine gesunde Person als positiv (also als krank) erkennt.
Beurteilen Sie den letzten Satz des Artikels: »Bei hoher Sensitivität und geringer Spezifität kann es viele falsch-positive Befunde geben.«