Frank Klinker

Stochastik

Teil 3: Bedingte Wahrscheinlichkeit, stochastische Unabhängigkeit und der Satz von Bayes

1 Bedingte Wahrscheinlichkeit und die Multiplikationsformel

1.1 Bedingte Wahrscheinlichkeit

Von einer **bedingten Wahrscheinlichkeit** spricht man, wenn Wahrscheinlichkeit eines Ereignis davon abhängt, ob ein anderes Ereignis vorher bereits eingetreten ist.

Das ist kein sonderlich besonderes Phänomen, sondern sehr oft die Regel – insbesondere bei mehrstufigen Zufallsexperimenten.

Beispiel 1 (Formulierungsabhängigkeit)**.** Beim Würfeln mit einem normalen Würfel betrachten wir die Ereignisse

$$G : \text{"Der Wurf ist gerade"} \text{ und } U : \text{"Der Wurf ist ungerade"}.$$

Es ist klar, dass $P(G) = P(U) = \frac{1}{2}$.

Würfelt man jedoch zunächst und erhält z. B. eine "2", dann ändert sich die Wahrscheinlichkeit des Ereignisses U zu Null.

Jetzt kann man in dem Beispiel einwenden, dass das ja klar sei, denn man hat ja eine "2" gewürfelt und man fragt ja schließlich vor dem Würfeln nach den Wahrscheinlichkeiten.

Dennoch: Die Wahrscheinlichkeit eine gerade oder ungerade Zahl zu würfeln hat sich geändert, nämlich abhängig davon, was vorher passiert ist!

Bezeichnung 2. Hängt die Wahrscheinlichkeit eines Ereignisses B von einem anderen Ereignis A ab, dann schreiben wir

$$P(B|A)$$

und nennen die das **die Wahrscheinlichkeit von B unter der Bedingung A.**[1]

Adresse: Eduard-Spranger-Berufskolleg, 59067 Hamm
E-Mail: mail@frank-klinker.de
Version: 7. November 2025

[1]In der Literatur findet man auch die Bezeichnungen $P(B|A) = P_A(B) = P(B/A)$.

Beispiel 3 (Standardbeispiel)**.** In einer Urne liegen fünf Kugeln: davon drei rote und zwei weiße. Daraus werden ohne Zurücklegen zwei Kugeln gezogen

Wir betrachten folgende Ereignisse:

$$\begin{aligned} R_i &: \text{"Im } i\text{-ten Zug ist die Kugel rot"} \\ W_i &: \text{"Im } i\text{-ten Zug ist die Kugel weiß"} \end{aligned}$$

Damit ist insbesondere $W_i = \overline{R_i}$.
Im ersten Zug gilt

$$P(R_1) = 0{,}6\,, \quad P(W_1) = 0{,}4\,.$$

Bei der Frage nach den Wahrscheinlichkeiten im zweiten Zug ist das Ergebnis des ersten Zugs selbstverständlich entscheidend:

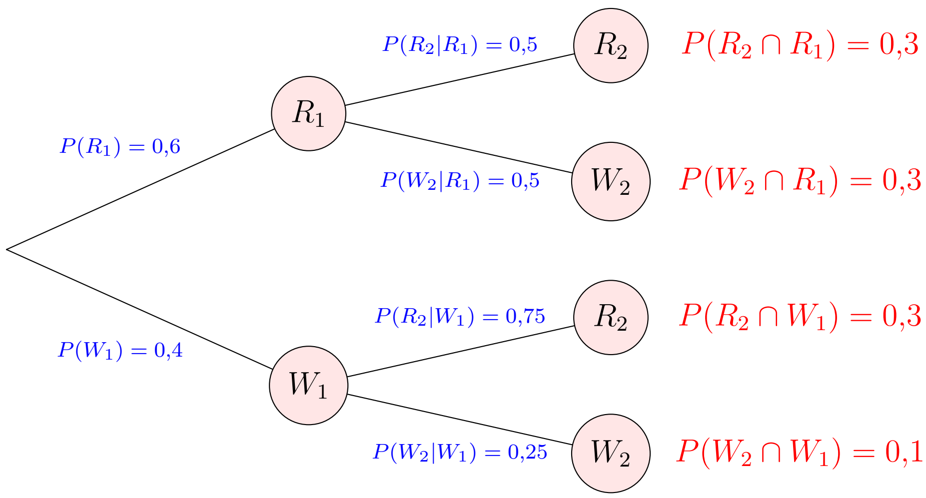
$$\begin{aligned} P(R_2|R_1) &= 0{,}5\,, & P(W_2|R_1) &= 0{,}5\,, \\ P(R_2|W_1) &= 0{,}75\,, & P(W_2|W_1) &= 0{,}25\,. \end{aligned}$$

Diese Wahrscheinlichkeiten erhält man etwa durch Abzählen einer "neuen" verkleinerte Urne.

- Im Fall R_1 hat man eine mit vier Kugeln: davon zwei rot und zwei weiß.
- Im Fall W_1 hat man eine mit vier Kugeln: davon drei rot und eine weiß.

In beiden Fällen startet man mit einem "neuen, anderen" Zufallsexperiment.

Das Beispiel der Urne sieht mit einem Wahrscheinlichkeitsbaum wie folgt aus:



Hier haben wir die "Schnittwahrscheinlichkeiten" mit Hilfe der Multiplikationsregel am Baum berechnet, z. B. $P(R_1 \cap R_2) = P(R_1) \cdot P(R_2|R_1)$.

Das (bzw. die umgeformte Gleichung) nehmen wir nun her, um unbekannte bedingte Wahrscheinlichkeiten für die allgemeine Situation zu definieren:

Bedingte Wahrscheinlichkeit

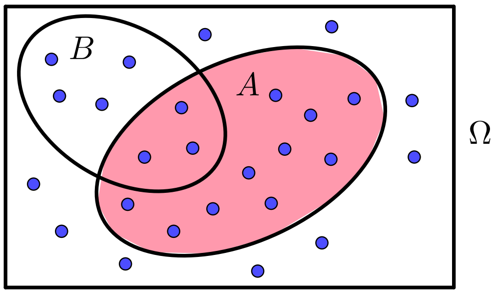
Ist A ein Ereignis und hängt das Ereignis B von A ab, dann <u>definiert</u> man

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Die vorige Definition ist zwar durch das Urnenexperiment motiviert, sie lässt sich aber auch ausgehend von einem Laplace-Experimente begründen.

Wir wählen ein Laplace-Experiment mit $\#\Omega = 26$ und weiter zwei Ereignisse A und B mit $\#A = 13$ und $\#B = 7$. Damit ist mit der üblichen Rechenregel "*Anzahl der Ergebnisse, für die das Ereignis eintritt, geteilt durch Anzahl aller möglichen Ergebnisse*" $P(A) = \frac{1}{2}$ und $P(B) = \frac{7}{26}$

Fragt man nun aber nach der Wahrscheinlichkeit $P(B|A)$, dann entspricht das einer geänderten Ausgangssituation: Man sucht hier nach Ergebnissen in B, die gleichzeitig auch Ergebnisse in A sind, **aber** dieses Mal nur innerhalb den Ergebnissen, die bereits in A lagen! Das heißt, man darf als Ergebnisse aus B nur die in $A \cap B$ zählen und die Ergebnismenge ist auf die Ergebnisse in A reduziert, siehe Skizze.



Damit ergibt sich

$$P(B|A) = \frac{\#(A \cap B)}{\#A} = \frac{3}{13}$$

und entspricht damit

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{3}{26}}{\frac{1}{2}}$$

Wir machen dieses Beispiel konkreter:

Beispiel 4. Wir nehmen eine Urne mit 26 Kugeln die von 1 bis 26 nummeriert sind und betrachten die Ereignisse A :"gezogene Kugel hat ungeraden Wert" sowie B :"gezogene Kugel hat Wert größer als 19".

Dann ist $\#A = 13$ und $\#B = 7$, also $P(A) = \frac{13}{26} = 0{,}5$ und $P(B) = \frac{7}{26} \approx 0{,}27$. Weiter ist $\#(A \cap B) = 3$, sodass $P(A \cap B) = \frac{3}{26} \approx 0{,}12$.

Sucht man die Wahrscheinlichkeit dafür, dass

"eine Kugel mit ungeradem Wert, einen Wert hat, der größer als 19 ist",

dann entspricht das $P(B|A)$. Durch Abzählen oder mit Hilfe der Definition ergibt sich

$$P(B|A) = \frac{3}{13} = \frac{\frac{3}{26}}{\frac{13}{26}} \approx 0{,}23\,.$$

Fragt man nach der Wahrscheinlichkeit, dass

"der Wert einer Kugel ungerade ist, wenn der Wert größer als 19 ist", dann ist das $P(A|B)$. Auch hier ergibt sich die Wahrscheinlichkeit durch Abzählen oder mit Hilfe der Definition:

$$P(A|B) = \frac{3}{7} = \frac{\frac{3}{26}}{\frac{7}{26}} \approx 0{,}43\,.$$

Achtung!

Es ist manchmal schwierig, bei einer sprachlichen Formulierung eines Ereignisses, die beteiligten Teilereignisse herauszufinden. Manchmal ist es aber auch schwierig herauszufinden, welches der in der Formulierung vorkommenden Ereignisse die Bedingung beschreibt – hier hilft einem oft das Schlüsselwort **wenn**.[2]

Bemerkung 5.

1. Die bedingte Wahrscheinlichkeit $P(\,\cdot\,|A)$ kann als Wahrscheinlichkeitsverteilung eines "neuen" Zufallsexperiments interpretiert werden.

   In diesem Sinne ist in dem obigen Beispiel W_2 wieder das Gegenereignis zu R_2 aber eben abhängig davon, was vorher passiert ist!

2. Aus dieser Eigenschaft – aber auch aus der ersten Vollständigkeitsregel für Bäume – ergibt sich folgende Formel, die für zwei beliebige Ereignisse A, B gültig ist:

$$\boxed{P(B|A) = 1 - P(\bar{B}|A)}$$

1.2 Die Multiplikationsformel für bedingte Wahrscheinlichkeiten

Stellt man die Formel aus der Definition um, dann erhält man den

Multiplikationsformel für bedingte Wahrscheinlichkeiten

Sind A und B Ereignisse, wobei B von A abhängt, dann gilt

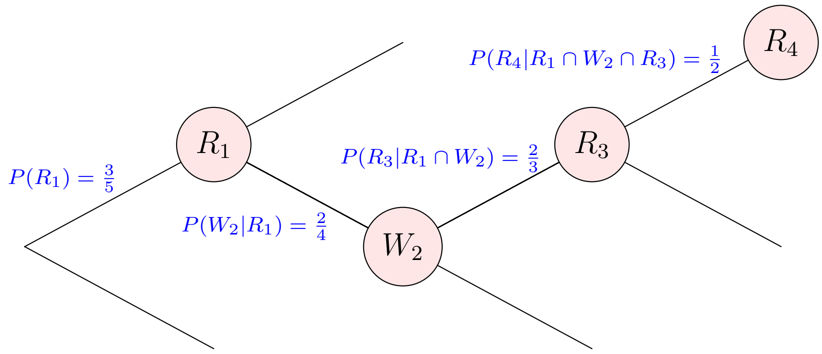
$$P(A \cap B) = P(A) \cdot P(B|A)\,.$$

Das war genau das Resultat der Pfadregel im Urnenexperiment, dass uns zur Definition führte.

Kehren wir also nochmal zu diesem zurück und sehen uns die Ereignisse R_i und W_i nach z. B. vier Ziehungen an.

[2] Dazu kann es hilfreich sein, das Ereignis umzuformulieren. Im ersten Teil von Beispiel 4 ist das etwa *[...]"eine Kugel einen Wert größer als 18 hat, wenn der Wert ungerade ist"*.

Ein Ausschnitt des Wahrscheinlichkeitsbaums sieht dann wie folgt aus:



Aus diesem Baum ergibt sich z. B.

$$P(R_1 \cap W_2 \cap R_3 \cap R_4) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} = 0{,}1$$

oder ohne Zahlen

$$P(R_1 \cap W_2 \cap R_3 \cap R_4) = P(R_1) \cdot P(W_2|R_1) \cdot P(R_3|R_1 \cap W_2) \cdot P(R_4|R_1 \cap W_2 \cap R_3)\,.$$

Auf diese Weise lässt sich eine allgemeine Multiplikationsformel für bedingte Wahrscheinlichkeiten formulieren (obwohl man in der Praxis die gesuchte Wahrscheinlichkeit eher sukzessive berechnen würde):

Bemerkung 6. Sind $A_1, A_2, A_3, \ldots, A_n$ Ereignisse, dann ist

$$\begin{aligned} &P(A_1 \cap \ldots \cap A_n) \\ &\qquad = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot \ldots \cdot P(A_n|A_1 \cap \ldots \cap A_{n-1}) \end{aligned}$$

2 Die totale Wahrscheinlichkeit und der Satz von Bayes

Beide Begriffe aus der Überschrift, ergeben sich als direkte Folgerung aus der Definition der bedingten Wahrscheinlichkeit bzw. dem Multiplikationssatz als dessen Umformulierung.

Dennoch sind beide sehr nützlich bei der Bearbeitung spezieller Fragestellungen.

2.1 Die totale Wahrscheinlichkeit

Wir betrachten ein Zufallsexperiment mit Ergebnismenge Ω. Ist $A \subset \Omega$ ein Ereignis, dann ist $\Omega = A \cup \bar{A}$ eine disjunkte Zerlegung von Ω.

Ist jetzt $B \subset \Omega$ ein beliebiges Ereignis, dann ist

$$B = (B \cap A) \cup (B \cap \bar{A})$$

eine disjunkte Zerlegung von B. Damit gilt

$$P(B) = P(B \cap A) + P(B \cap \bar{A})\,.$$

Weiter haben wir wegen des Multiplikationssatzes $P(B \cap A) = P(A) \cdot P(B|A)$ und $P(B \cap \bar{A}) = P(\bar{A}) \cdot P(B|\bar{A})$.

Zusammen gibt das den Satz von der totalen Wahrscheinlichkeit:

Satz von der totalen Wahrscheinlichkeit

Sind $A, B \subset \Omega$ Ereignisse eines Zufallsexperiments mit Ergebnismenge Ω, dann gilt

$$P(B) = P(A) \cdot P(B|A) + P(\bar{A}) \cdot P(B|\bar{A})$$

Bemerkung 7. Die totale Wahrscheinlichkeit hat eine Verallgemeinerung.

Ist $\Omega = A_1 \cup A_2 \cup \ldots \cup A_n$ eine disjunkte Zerlegung der Ergebnismenge Ω eines Zufallsexperiments, dann gilt für ein beliebiges Ereignis $B \subset \Omega$:

$$P(B) = P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + \ldots + P(A_n) \cdot P(B|A_n)\,.$$

Die Begründung für diese Formel ergibt sich aus einer einfachen Erweiterung der obigen Argumente.

2.2 Der Satz von Bayes

Sind $A, B \subset \Omega$ zwei Ereignisse, dann gilt wegen des Multiplikationssatzes $P(A \cap B) = P(A) \cdot P(B|A)$ und ebenso $P(A \cap B) = P(B) \cdot P(A|B)$. Damit bekommt man eine Verbindung zwischen den bedingten Wahrscheinlichkeiten $P(A|B)$ und $P(B|A)$:

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{P(B)}\,.$$

Zusammen mit der totalen Wahrscheinlichkeit gibt das den Satz von Bayes:

Satz von Bayes

Sind $A, B \subset \Omega$ Ereignisse eines Zufallsexperiments mit Ergebnismenge Ω, dann gilt

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{P(A) \cdot P(B|A) + P(\bar{A}) \cdot P(B|\bar{A})}$$

Bemerkung 8. Auch der Satz von Bayes hat eine Variante, die auf der Verallgemeinerung der totalen Wahrscheinlichkeit beruht:

Ist $\Omega = A_1 \cup A_2 \cup \ldots \cup A_n$ eine disjunkte Zerlegung der Ergebnismenge Ω eines Zufallsexperiments, dann gilt für ein beliebiges Ereignis $B \subset \Omega$:

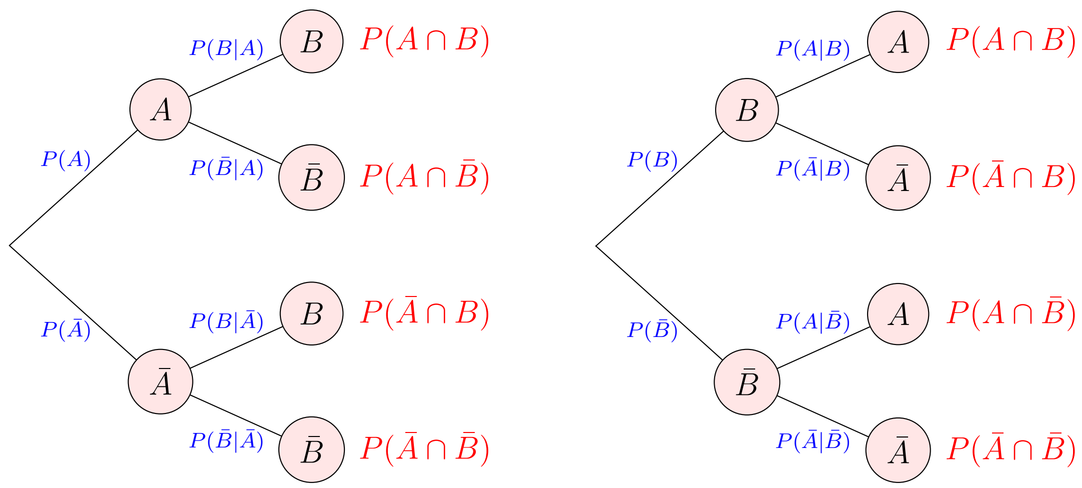
$$P(A_i|B) = \frac{P(A_i) \cdot P(B|A_i)}{P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + \ldots + P(A_n) \cdot P(B|A_n)}$$

3 Wahrscheinlichkeitsbaum und Vierfeldertafel

Wir betrachten zwei Ereignisse A und B eines Zufallsexperiments.

Wie wir wissen, lassen sich diese mit einem **Wahrscheinlichkeitsbaum** darstellen. Dabei gibt es zwei Varianten, je nachdem ob man mit A oder B startet.

Wahrscheinlichkeitsbäume für zwei Ereignisse



Stellt man die Grundversion der Bäume (schwarz/blau) auf, dann sieht man, dass zunächst Informationen fehlen.

Im ersten Baum[3] fehlen $P(B), P(\bar{B}), P(A|B), P(A|\bar{B}), P(\bar{A}|B), P(\bar{A}|\bar{B})$ sowie die rot notierten Wahrscheinlichkeiten der Schnitte.

Die roten kann man mit der Pfadregel berechnen, und die weiter fehlenden berechnet man mit dem Satz von Bayes und der totalen Wahrscheinlichkeit.

Z. B. $P(B) = P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})$ und damit $P(A|B) = \dfrac{P(A \cap B)}{B}$.

Eine alternative Art der Darstellung ist die **Vierfeldertafel**. Diese ist symmetrischer als der Wahrscheinlichkeitsbaum, nämlich insofern, als sie die Ereignisse A und B gleichwertig behandelt.

Vierfeldertafel für zwei Ereignisse

	B	$\bar{B}$	
A	$P(A \cap B)$	$P(A \cap \bar{B})$	$P(A)$
$\bar{A}$	$P(\bar{A} \cap B)$	$P(\bar{A} \cap \bar{B})$	$P(\bar{A})$
	$P(B)$	$P(\bar{B})$	1

[3] Zur Diskussion des zweiten Baumes vertauscht man die Rollen von A und B.

In der Grundversion der Tafel (schwarz/blau) fehlen ebenfalls Informationen, nämlich alle Wahrscheinlichkeiten außer die der Schnitte.

Insbesondere erhält man die roten Wahrscheinlichkeiten zeilen- bzw. spaltenweise als Summe der davor stehenden Wahrscheinlichkeiten (wegen der Eigenschaften disjunkter Mengen).

Die bedingten Wahrscheinlichkeiten berechnet man dann mit dem Satz von Bayes.

Z. B. Ist $P(A) = P(A \cap B) + P(A \cap \bar{B})$ und damit $P(B|A) = \dfrac{P(A \cap B)}{P(A)}$ und $P(\bar{B}|A) = \dfrac{P(A \cap \bar{B})}{P(A)}$.

Hinweis 9.

- Ob man einen Wk-Baum oder die VF-Tafel verwendet ist ein wenig Geschmachssache, kann aber auch von der Art der Fragestellung abhängen.
- Oft stellt man im Rahmen eines Problems oder einer Aufgabe einen Wk-Baum (oder eine VF-Tafel) auf ohne ihn (oder sie) vollständig ausfüllen zu können, da Informationen fehlen. Dann sind in der Regel aber weitere Informationen vorgegeben, die im Baum (oder der Tafel) nicht stehen. Das sind solche, die die wir oben berechnet haben.
  Dann füllt man die Lücken im Baum (oder in der Tafel), indem man die Rechnungen von oben "rückwärts" durchführt.

Hinweis 10. Die Vierfeldertafel ist auch geeignet, um Aufgaben oder Probleme im Zusammenhang mit Häufigkeiten (absolut oder relativ) zu bearbeiten:

	B	$\bar{B}$	
A	$H(A \cap B)$	$H(A \cap \bar{B})$	$H(A)$
$\bar{A}$	$H(\bar{A} \cap B)$	$H(\bar{A} \cap \bar{B})$	$H(\bar{A})$
	$H(B)$	$H(\bar{B})$	Gesamtzahl

Die Tafel zu den relativen Häufigkeiten erhält man dann, indem man alle Inhalte durch die Gesamtzahl teilt – das ist dann i. W. die alte Vierfeldertafel mit $h(\cdot)$ statt $P(\cdot)$.

Bedingte relative Häufigkeiten lassen sich aus der Tafel sinnvoll bestimmen, z. B. $h(B|A) = \dfrac{H(A \cap B)}{H(A)}$ (siehe dazu auch die Diskussion vor Beispiel 4 auf Seite 3).

4 Anwendung: Beurteilung von (medizinischen) Tests

- Lein I et al.: SARS-CoV-2: Testergebnisse richtig einordnen. *Deutsches Ärzteblatt* Heft 47 (2020) A2304-A2305. https://www.aerzteblatt.de/archiv/titel/dae/2020/47/sars-cov-2-testergebnisse- richtig-einordnen-90bd0d91-5482-48c7-8485-7b1ecc2d0737
- Faller H: Sensitivität, Spezifität, positiver und negativer Vorhersagewert. *Rehabilitation* **44** (2005) 44-49, https:/www.doi.org/10.1055/s-2004-834624
- Seifried J, et al.: Was ist bei Antigentests zur Eigenanwendung (Selbsttests) zum Nachweis von SARS-CoV-2 zu beachten? *Epidemiologisches Bulletin* **8** (2021) 3-9. https://www.rki.de/DE/Aktuelles/Publikationen/Epidemiologisches-Bulletin/2021/08_21
- Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Sensitivität und Spezifität. Vers. 26.08.2022. https://www.gesundheitsinformation.de/sensitivitaet-und-spezifitaet.html

Bei der Beurteilung medizinischer Test spielen die Begriffe **Sensitivität**, **Spezifität** und **Prävalenz** eine besondere Rolle. Darüber hinaus werden auch noch die Begriffe **positiver Vorhersagewert** (ppv), **negativer Vorhersagewert** (npv) sowie die Begriff **falsch-positiv**, **richtig-positiv**, **falsch-negativ** und **richtig-negativ**.

Bei der Lektüre der drei obigen Texte kann man sich nun folgende Fragen stellen:

1. Wie hängen die Fachbegriffe mit unsere Beschreibung von Wahrscheinlichkeiten zusammen? Wie kann man diese Begriff in unsere Schreibweise "übersetzen"?
2. Wieso reichen Sensitivität, Spezifität und Prävalenz aus, um die zusätzlichen Begriffe zu erklären? Welche weiteren Wahrscheinlichkeiten kann man in diesem Zusammenhang noch berechnen?
3. Welche Probleme können in der Beurteilung eines Test auftreten, wenn man nur die Werte der Sensitivität, Spezifität und, Prävalenz kennt? Hierbei muss man gegebenenfalls beachten, dass es sich nicht wirklich um Wahrscheinlichkeiten handelt, sondern dass die Werte sich als statistisch erlangte Werte handelt (etwa um relative Häufigkeiten).

Zur Bearbeitung von Frage 3 können die folgenden Beispiele hilfreich sein:

- Werte (eines realistischen Tests): Sensitivität 80%, Spezifität 95%, Prävalenz 1%.
- Ein Schwangerschaftstest liefert folgende absolute Häufigkeiten:

	Schwanger	Nicht schwanger	Insges.
Positiver Test	13	367	380
Negativer Test	2	4538	4540
Insges.	15	4905	4920

5 Stochastische Unabhängigkeit

Wie wir gesehen haben, ist die Eigenschaft, dass die Wahrscheinlichkeit von Ereignissen von einander abhängen, sehr natürlich. Daher ist es auch interessant, wann das denn nicht der Fall ist.

Die naive Interpretation, zwei Ereignisse, die sich in ihrer Formulierung nicht beeinflussen, als unabhängig zu bezeichnen greift da jedoch nicht.

Dazu sehen wir uns das Beispiel des Würfelwurfs nochmal an. Wir betrachten die Ereignisse $G:$"Der Wurf ist gerade" und $U:$ "Der Wurf ist ungerade".
In diesem Fall sind $P(G) = P(U) = \frac{1}{2}$ und $P(G|U) = P(U|G) = 0$ – da ja $U \cap G = \varnothing$: die Ereignisse beeinflussen sich gegenseitig nicht.
Dennoch hängt die Wahrscheinlichkeit, dass U eintritt sehr wohl davon ab, ob G eingetreten ist oder nicht: die Wahrscheinlichkeit für U ist 1 oder 0, je nachdem, ob G eingetreten ist oder nicht! Sie sind also hohem Maße von einander abhängig.

Die **stochastische Abhängigkeit** und **stochastische Unabhängigkeit** von Ereignissen ist also vielmehr eine Eigenschaft der Wahrscheinlichkeiten. Sie lässt sich in der Regel nicht so einfach aus den Formulierungen der Ereignisse herauslesen.

Wir starten mit der Definition:

Zwei Ereignisse $A, B \subset \Omega$ mit $P(A) \neq 0, P(B) \neq 0$ heißen **stochastisch unabhängig**, wenn

$$P(A \cap B) = P(A) \cdot P(B)$$

Fakten 11. Die folgenden Eigenschaften sind äquivalent:[4]

1. A und B sind stochastisch unabhängig
2. $P(A|B) = P(A)$
3. $P(B|A) = P(B)$
4. A und $\bar{B}$ sind stochastisch unabhängig
5. B und $\bar{A}$ sind stochastisch unabhängig

Wir sehen uns hier die Begründung von zwei dieser Äquivalenzen an (die fehlenden Begründungen kann man dann ganz analog durchführen):

Aus Punkt 1 folgt der Punkt 4:

$$\begin{aligned} P(A)P(\bar{B}) &= P(A)\big(1 - P(B)\big) \\ &= P(A) - P(A)P(B) \\ &\overset{1.}{=} P(A) - P(A \cap B) \\ &\overset{\substack{\text{disj.}\\ \text{Zerl.}}}{=} P(A \setminus B) \\ &\overset{\substack{\text{Mengen}\\ \text{gleich}}}{=} P(A \cap \bar{B})\,. \end{aligned}$$

[4]Wegen Punkt 2 nutzt man in der Literatur teilweise $P(A) = P(A|B)$ als Definition der stochastischen Unabhängigkeit der Ereignisse A und B.

Vertauscht man hierin die Rollen von B und $\bar{B}$, dann folgt aus Punkt 4 der Punkt 1. Beides zusammen gibt dann die Äquivalenz der Punkte 1 und 4.

Wegen des Multiplikationssatzes $P(A \cap B) = P(A) \cdot P(B|A)$ ist

$$\begin{aligned} P(A \cap B) = P(A) \cdot P(B) &\iff P(A) \cdot P(B|A) = P(A) \cdot P(B) \\ &\iff P(B|A) = P(B)\,. \end{aligned}$$

Also ist Punkt 1 äquivalent zu Punkt 3.

Beispiel 12. Sprachlich schließen sich die Ereignisse A und $\bar{A}$ aus. Sie sind jedoch stochastisch abhängig: $0 = P(A|\bar{A})$ aber $P(A) \neq 0$ (siehe etwa das Eingangsbeispiel des Würfelwurfs mit G und $U = \bar{G}$).

Betrachten wir eine leichten Änderung des Beispiels 4 (Dort waren A und B stochastisch abhängig!)

Beispiel 13. Wir nehmen wieder die Urne mit 26 Kugeln die von 1 bis 26 nummeriert sind. Jetzt betrachten wir die Ereignisse A :"gezogene Kugel hat ungeraden Wert" und (leicht verändert) B :"gezogene Kugel hat Wert größer als 18".

Dann ist $\#A = 13$ und $\#B = 8$, also $P(A) = \frac{13}{26} = 0{,}5$ und $P(B) = \frac{8}{26} \approx 0{,}31$. Weiter ist $\#(A \cap B) = 4$, sodass $P(A \cap B) = \frac{4}{26} \approx 0{,}15$.

Wir suchen auch hier die Wahrscheinlichkeit dafür, dass

"eine Kugel mit ungeradem Wert, einen Wert hat, der größer als 18 ist",

also $P(B|A)$. Ähnlich wie zuvor ergibt sich hier

$$P(B|A) = \frac{4}{13} = \frac{\frac{4}{26}}{\frac{13}{26}} \approx 0{,}31\,.$$

Die Wahrscheinlichkeit für

"der Wert einer Kugel ungerade ist, wenn der Wert größer als 18 ist"

bekommt man auch wie in Beispiel 4:

$$P(A|B) = \frac{4}{8} = \frac{\frac{4}{26}}{\frac{8}{26}} = 0{,}5\,.$$

Man sieht: Auch wenn die Ereignisse sprachlich nicht unvereinbar sind, so sind sie durch die kleine Änderung nun stochastisch unabhängig.