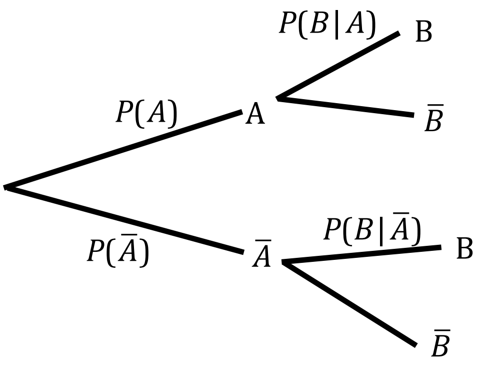
4. Bedingte Wahrscheinlichkeit

Bedingte Wahrscheinlichkeit ist ein neuer Fachbegriff, der uns inhaltlich bereits in den vorhergehenden Kapiteln begegnet ist. Er taucht in mehrstufigen Zufallsexperimenten auf. Am einfachsten und deutlichsten lässt sich die bedingte Wahrscheinlichkeit an einem einfachen Baumdiagramm erläutern. Wir haben ein zweistufiges Experiment. Auf der ersten Stufe kann das Ereignis A passieren oder nicht, auf der zweiten Stufe das Ereignis B.

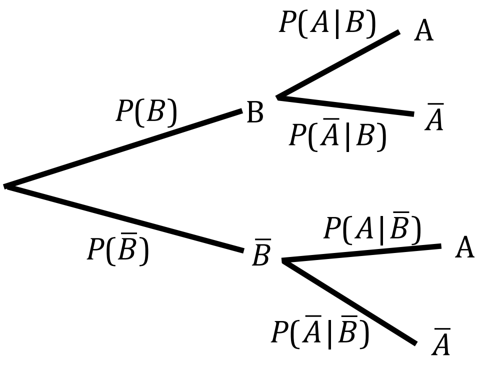


Die Wahrscheinlichkeit von B hängt von A bzw. $\overline{A}$ ab. Dazu schreiben wir $P(B|A)$ bzw. $P(B|\overline{A})$. $P(B|A)$ wird gesprochen als: „ Die Wahrscheinlichkeit für B unter der Bedingung A.“

Inhaltlich-anschaulich fragt diese Wahrscheinlichkeit also: „Wie groß ist die Wahrscheinlichkeit für B, wenn A bzw. $\overline{A}$ bereits passiert ist?“

Bleibt die Frage, warum wir dem Begriff der bedingten Wahrscheinlichkeit ein extra Kapitel widmen, wo wir ihn doch längst schon angewendet haben. Das liegt daran, dass in der bedingten Wahrscheinlichkeit einige überraschende Ergebnisse stecken.

Eine Überraschung können wir erhalten, wenn wir das oben dargestellte, einfache Baumdiagramm **umdrehen** (umgekehrtes Baumdiagramm). Das kommt in der Praxis durchaus vor. Zum Beispiel kann A das Bestehen eines Tests sein (mit dem etwa ein Schüler oder Bewerber getestet wird) und B ist das Ereignis, dass der Bewerber tatsächlich gut ist (auf Grund anderer Gütebeurteilungen). Das obere Baumdiagramm ist dann die Situation, dass man den Test anwendet und dann wissen will, mit welcher Wahrscheinlichkeit man einen guten Bewerber erhält. $P(B|A)$ ist dann die Wahrscheinlichkeit, dass ein Bewerber gut ist, wenn er den Test bestanden hat.



Das umgekehrte Baumdiagramm beschreibt die Situation, dass man (vor der Anwendung) den Test testet. Man kennt die Güte von Probanden (auf Grund anderer Beurteilungen), lässt diese den Test machen und schaut nun, mit welcher Wahrscheinlichkeit sie den Test bestehen. $P(A|B)$ ist also die Wahrscheinlichkeit, dass ein Proband den Test besteht, wenn bekannt ist, dass er gut ist. Natürlich gilt auch für die bedingten Wahrscheinlichkeit, also denen für die zweite Stufe, die einfache Regel:

$$\underbrace{P(A|B)}_{\text{Ereignis}}+\underbrace{P(\overline{A}|B)}_{\text{Gegenereignis}}=1$$

Rein formal muss man hier darauf achten, dass die Bedingung in beiden Wahrscheinlichkeiten die gleiche ist.

Man darf aber nicht Ereignis und Gegenereignis in der Bedingung kombinieren. Die Summe $P(A \mid B) + P(A \mid \overline{B})$ hat keinen regelhaften, festen Wert. (Unterschiedliche Stränge eines anderen Hauptzweiges hängen nicht zusammen!)

Eine weitere, in der Praxis wichtige Anwendung ist die Diagnose von Krankheiten. Ein Test kann ergeben, dass man eine Krankheit hat. Dabei muss man aber nicht krank sein. Die umgekehrte Situation ist ebenso möglich. Der Test von Diagnoseverfahren ist aufwendig und erfordert sorgfältige, wissenschaftliche Arbeit. An der Universität Bremen gibt es dafür zwei Institute (BIBS, KKS), die über Deutschland hinaus anerkannte Institute sind.

Zusammenhänge zwischen beiden Baumdiagrammen

Hängen die Wahrscheinlichkeiten in beiden Baumdiagrammen zusammen oder sind diese völlig unabhängig voneinander? Ein Zusammenhang ergibt sich zwischen beiden Bäumen über die Enden. Läuft man im ersten Baum entlang des obersten Pfades, so erhält man erst das Ereignis A, dann das Ereignis B. Am Ende steht also das Ereignis $A \cap B$ und man berechnet die Wahrscheinlichkeit $P\left(A \cap B\right)$ mit der Pfadregel $P(A) \cdot P(B \mid A) = P(A \cap B)$

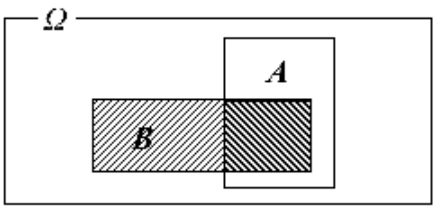
Bei dem umgekehrten Baum hat man ebenfalls einen Pfad, der über B und dann A läuft. Auch hier berechnet man die Wahrscheinlichkeit $P\left(A \cap B\right)$ über die Pfadregel:

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

Folglich ergibt sich die Gleichung: $P(A) \cdot P(B \mid A) = P(A \cap B) = P(B) \cdot P(A \mid B)$. Diese Gleichung verknüpft also eine bedingte Wahrscheinlichkeit $P(B \mid A)$ mit der, in der Ereignis und Bedingung gerade vertauscht sind: $P(A \mid B)$.

Im ersten Baum taucht die Wahrscheinlichkeit für B nur unter den Bedingungen auf. Im umgekehrten Diagramm braucht man aber die Wahrscheinlichkeit für B ohne Bedingung. Diese Wahrscheinlichkeit nennt man in diesem Zusammenhang auch die totale Wahrscheinlichkeit. Die zweite wichtige Regel ist die zur Berechnung der totalen Wahrscheinlichkeit. Auch diese stellt eine Verbindung zwischen den Werten beider Baumdiagramme her.

Zur Herleitung der Formel für die totale Wahrscheinlichkeit beginnen wir mit reiner Mengenlehre:



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

Hier sind $B \cap A$ und $B \cap \overline{A}$ immer disjunkt. Dann gilt nach den Axiomen für die Wahrscheinlichkeit:

$P(B) = P(B \cap A) + P(B \cap \overline{A})$

Diese Formel können wir aber auch direkt im Baumdiagramm ablesen.

Die Wahrscheinlichkeiten für die Schnittmengen lassen sich mit der Pfadregel berechnen:

$P(A \cap B) = P(A) \cdot P(B \mid A)$ und $P(\overline{A} \cap B) = P(\overline{A}) \cdot P(B \mid \overline{A})$

Also ergibt sich als **Formel für die totale Wahrscheinlichkeit**

$$\boxed{P(B) = P(A) \cdot P(B \mid A) + P(\overline{A}) \cdot P(B \mid \overline{A})}$$

Diese theoretischen Betrachtungen wollen wir konkret an einem Beispiel betrachten. Wir wählen dazu das Einführungsbeispiel aus Kapitel 0, das dort unter der Überschrift „Qualitätstest“ aufgeführt ist. Es geht dabei um ein Büro, in dem zwei Sekretärinnen arbeiten und sich die Arbeiten aufteilen. Beide arbeiten mit unterschiedlicher Sicherheit. Die in der Arbeit gegebenen Daten formulieren wir gleich in Ereignissen und deren Wahrscheinlichkeiten:

A: „Arbeit wurde von Sekretärin 1 erledigt.“ P(A) = 0,6

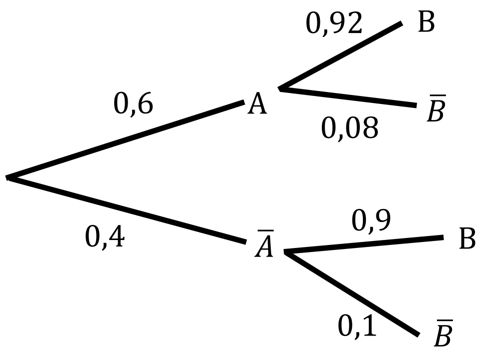
$\overline{A}$: „Arbeit wurde von Sekretärin 2 erledigt.“ P($\overline{A}$) = 1 – P(A) = 0,4

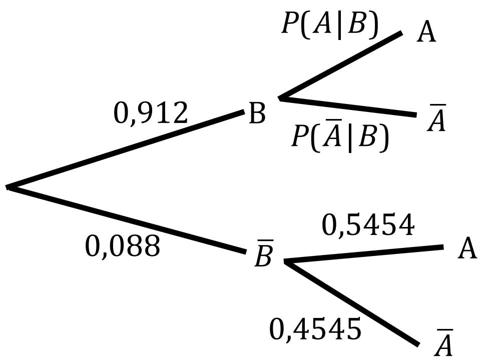
B: „Arbeit war fehlerfrei.“ Für Sekreträrin 1 galt das zu 92%. In unserer Notation für bedingte Wahrscheinlichkeiten bedeutet das: P(B|A) = 0,92

Bei Sekretärin 2 waren es nur 90%: P(B| $\overline{A}$) = 0,9

$\overline{B}$: „Arbeit war fehlerhaft.“ P($\overline{B}$ |A) = 0,08

P($\overline{B}$ | $\overline{A}$) = 0,1





Wir tragen alle Angaben übersichtlich in ein Baumdiagramm ein. Diesen Werten kann man entnehmen, dass Sekretärin 1 (A) mehr arbeitet und sicherer ist.

Der in der Aufgabe beschriebene Test durch den Chef geht in den Bedingungen genau umgekehrt vor. Er stellt einen Fehler fest, d.h. Ereignis $\overline{B}$ liegt vor, und forscht dann nach der Verursacherin des Fehlers. Seine kleine Statistik misst also die Wahrscheinlichkeiten P(A| $\overline{B}$) bzw. P($\overline{A}$ | $\overline{B}$). Alle diese Angaben tauchen im umgekehrten Baumdiagramm auf. Für die Berechnung dieser Werte bestimmt man zunächst die totale Wahrscheinlichkeit:

$P(B) = 0,6 \cdot 0,92 + 0,4 \cdot 0,9 = 0,912$

Womit auch P($\overline{B}$) = 1 – P(B) = 0,088 bekannt ist.

Wir wollen nun die Wahrscheinlichkeiten P(A| $\overline{B}$) und P($\overline{A}$ | $\overline{B}$) aus den Werten des Ausgangsdiagramms bestimmen.

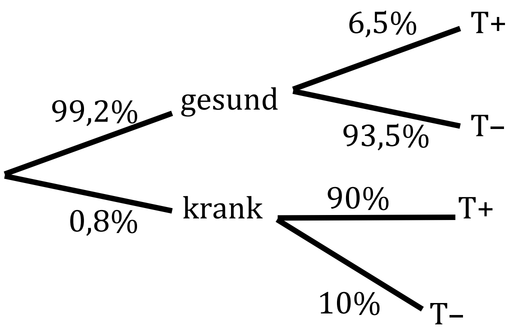
Es existiert eine Verknüpfung über $P(A \cap \overline{B})$:

$$P(A \mid \overline{B}) \cdot P(\overline{B}) = P(A) \cdot P(\overline{B} \mid A)$$

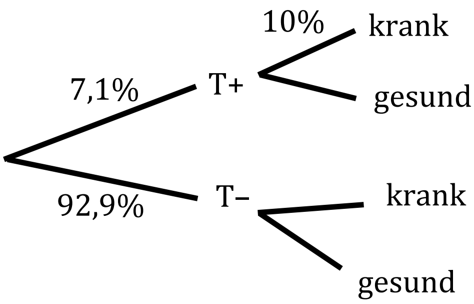
$$P(A \mid \overline{B}) = \frac{P(A) \cdot P(\overline{B} \mid A)}{P(\overline{B})} = \frac{0,6 \cdot 0,08}{0,088} = \frac{6}{11} \approx 0,5454$$

Die (verlässlicheren) Daten aus dem ersten Diagramm legen also zwangsläufig fest, dass in dem unzureichenden Test durch den Chef die leistungsfähigere Sekretärin schlechter abschneiden muss.

Das 2. Beispiel beruht auf statistischen Daten, die im Rahmen einer umfangreichen Studie zur Mammographie erhoben wurden. Dabei ist dieses Beispiel die Konkretisierung eines allgemeinen Problems in der Medizin. Man möchte eine Krankheit diagnostizieren (hier Brustkrebs bei älteren Frauen) und verwendet dazu ein Testverfahren (hier die Mammographie). Solche Testverfahren müssen nun selbst auch getestet werden, bevor sie offiziell zugelassen werden. Dazu werden Menschen, deren Gesundheitszustand aus anderen Untersuchungen bekannt ist, dem zu testenden Test unterzogen und es wird festgestellt, wie weit der Test den Gesundheitszustand korrekt anzeigt.



Das hier dargestellte Baumdiagramm stellt diese Messdaten dar: T+ besagt, dass der Test die Krankheit anzeigt, T- steht dafür, dass die Krankheit nicht festgestellt wurde.
Bei gesunden Frauen zeigt die Mammographie in 93,5% der Fälle dieses richtig an, bei Frauen mit Brustkrebs kann man das zu 90% der Mammographie entnehmen. Die totalen Wahrscheinlichkeiten für gesund/krank besagen, dass 0,8% der Frauen Brustkrebs haben.



Wird nun das Testverfahren, hier also die Mammographie, angewendet, um Brustkrebs festzustellen, kehren sich die Bedingungen um: Wie groß ist die Wahrscheinlichkeiten, Brustkrebs zu haben, wenn die Mammographie Indizien dafür anzeigt? Überraschender Weise ist diese Wahrscheinlichkeiten nur etwa 10%. Das Verfahren, das in der Testsituation eine Sicherheit von über 90% hatte, erweist sich in der Praxis als unsicher.

Die Erklärung findet man wieder in den entsprechenden bedingten Wahrscheinlichkeiten, die man genau so ausrechnen könnte wie im ersten Beispiel.
Es existiert jedoch noch ein zweites Lösungsverfahren. Nämlich die **Vier-Felder-Tafel**. Ausgehend von der Gesamtzahl der Probanden (26.057 unten rechts) kann man die Werte entsprechend der Wahrscheinlichkeiten im oberen Baumdiagramm nach und nach ausfüllen (schwarze Pfeile). (Tatsächlich sind die Zahlen in der Tabelle die Rohdaten, nach denen die Wahrscheinlichkeiten als relative Häufigkeiten berechnet wurden.)

	krank	gesund	
T+	179	1.671	1.850
T−	20	24.187	24.207
	199	25.858	26.057

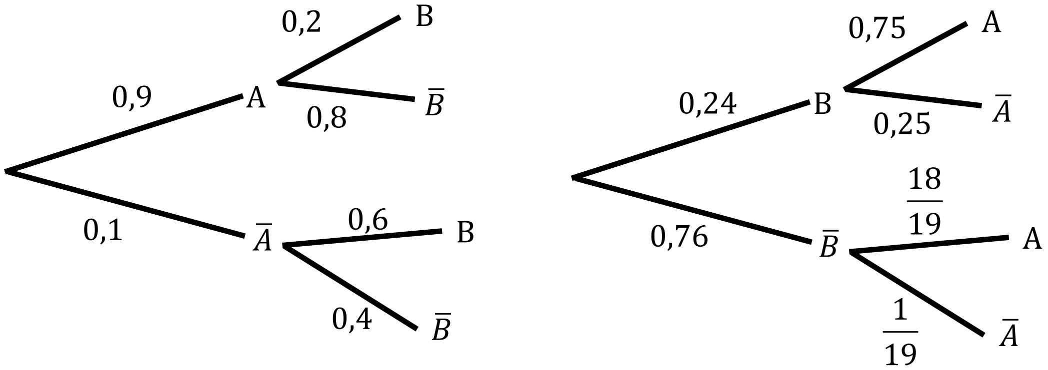
Die Wahrscheinlichkeiten für den umgekehrten Baum erhält man, indem man zunächst entlang der grauen Pfeile die Summen am rechten Rand berechnet und dann die entsprechenden Wahrscheinlichkeiten. Hier kann man sehr übersichtlich sehen, dass die Mammographie bei 1850 Frauen Hinweise auf eine Erkrankung geliefert hat, dass davon aber nur 179 tatsächlich Brustkrebs hatten. Das sind $\frac{179}{1850} \approx 9{,}7\%$.

<u>Eine kurze Übung zur Vier-Felder-Tafel</u>
Gegeben ist die nachfolgende Vier-Felder-Tafel.

	B	$\overline{B}$	$\sum$
A	18	72	90
$\overline{A}$	6	4	10
$\sum$	24	76	100

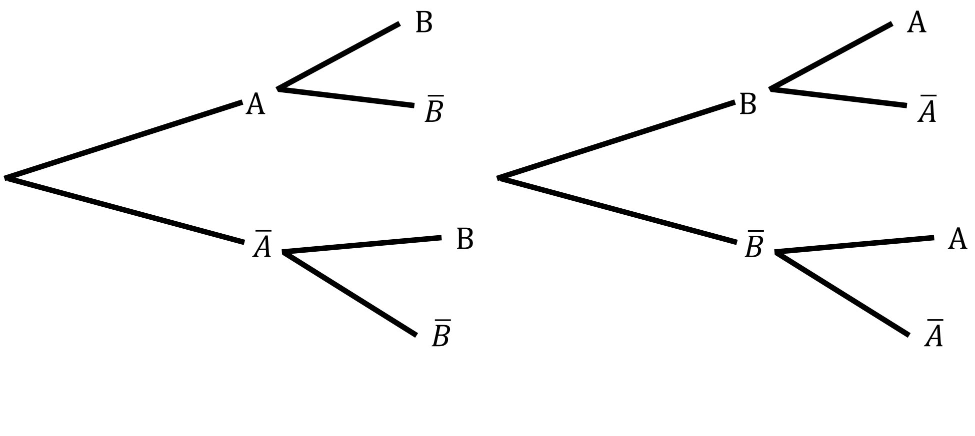
Aus dieser Tafel sollen die beiden Baumdiagramme mit den zugehörigen Wahrscheinlichkeiten (relativen Häufigkeiten) abgeleitet werden.

Lösung:



<u>Die **Formel von Bayes** (Satz von Bayes) als Zusammenfassung</u>

Für diese abstrakte Betrachtung gehen wir wie so oft von einem simplen Baumdiagramm eines zweistufigen Versuchs und dessen Umkehrung aus.



Die Wahrscheinlichkeiten in einem Baumdiagramm legen die Wahrscheinlichkeiten im anderen Baumdiagramm eindeutig fest. Das heißt, wenn man in dem einen Baumdiagramm alle Wahrscheinlichkeiten kennt, kann man die Wahrscheinlichkeiten beim umgekehrten Baumdiagramm ausrechnen. Dafür kann man die Formel von Bayes benutzen.
Zur Herleitung der Formel gehen wir wieder über das gemeinsame Eintreten von zwei Ereignissen. Dieses stellt die Verbindung zwischen beiden Baumdiagrammen her und geschieht immer über das Ende der Pfade des Baumdiagramms.

$P(A \cap B) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B)$. Dies kann nach der bedingten Wahrscheinlichkeit im umgekehrten Baumdiagramm auflösen: $P(A|B) = \frac{P(A) \cdot P(B|A)}{P(B)}$

Um die Wahrscheinlichkeit von P(B) zu ermitteln, benutzen wir die Formel für die totale Wahrscheinlichkeit:

$$P(B) = P(B \cap A) + P(B \cap \overline{A})$$
$$= P(A) \cdot P(B \mid A) + P(\overline{A}) \cdot P(B \mid \overline{A})$$

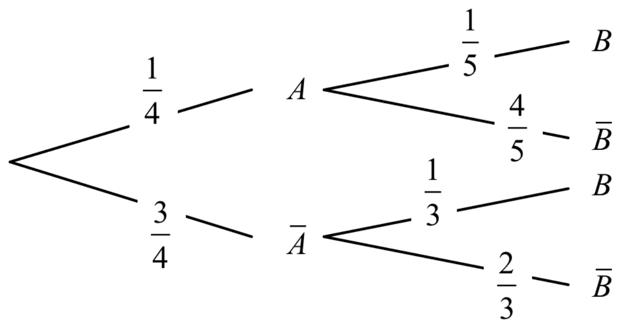
Die Wahrscheinlichkeit von P(B) ist nun ermittelt und wird in die obige Formel eingesetzt.

Satz von Bayes:

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

Übungsaufgaben zum Kapitel 4

Ü1 Gegeben ist das folgende Baumdiagramm



a. Schreiben Sie das Diagramm auf mit absoluten Zahlen bei einer Gesamtversuchszahl von 600.
b. Erstellen Sie aus den Zahlen in a. eine Vierfeldertafel.
c. Erzeugen Sie aus der Vierfeldertafel das umgekehrte Baumdiagramm
   i. mit den absoluten Zahlen
   ii. mit den Wahrscheinlichkeiten
d. Lesen Sie aus den entsprechenden Baumdiagrammen ab:
   $P(A \mid B)$ $P(A \mid \overline{B})$ $P(\overline{A} \mid \overline{B})$ $P(B \mid \overline{A})$ $P(\overline{B} \mid A)$
e. Wir interpretieren die gesamten Daten in dem Zusammenhang:
   A: Person ist Wähler der Grünen
   B: Person ist Befürworter einer Müllverbrennungsanlage
   Wie viel Prozent der Grünenwähler befürworten die Müllverbrennungsanlage?
   Wie viel Prozent der Befürworter der Müllverbrennungsanlage sind Grünenwähler?

Ü2 Eine Firma beschäftigt drei Mitarbeiter, die telefonische Anfragen von Kunden beantworten sollen. Herr Alleskönner kann 90% aller Frage zur Zufriedenheit der Kunden beantworten, Frau Besserwisser 80% und Herr Chancenlos noch gerade 70%. Berechnen Sie unter der Annahme, dass alle drei Mitarbeiter gleich viele Telefonate beantworten, die Wahrscheinlichkeiten, dass

a. ein Kunde an Herrn Alleskönner gerät und eine zufrieden stellende Antwort bekommt.
b. ein Kunde an Herrn Chancenlos gerät und eine nicht zufrieden stellende Antwort bekommt.
c. ein Kunde mit der Antwort, die er erhält, nicht zufrieden ist.
d. ein unzufriedener Kunde an Frau Besserwisser geraten ist.
e. eine Antwort, die zur Zufriedenheit des Kunden ausfiel, von Herrn Chancenlos gegeben wurde.
f. eine Antwort, die den Kunden nicht zufrieden stellt, von Herrn Alleskönner gegeben wurde.

Ü3 Die formale Definition der bedingten W' ist: $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$.

a. Begründen Sie diese Gleichung aus dem Rechnen im Baumdiagramm.
b. Erläutern Sie die Gleichung $P(A \mid \Omega) = P(A)$, wobei Ω wie üblich die Ergebnismenge ist.
c. Beweisen Sie formal die Gleichung $P(A \cup B \mid B) = 1$.
d. Berechnen Sie analog zu c die W'en $P(A \mid A)$ und $P(A \mid \overline{A})$.
e. Verwenden Sie die inhaltliche Erklärung für die bedingte W': „*P*(A|B) ist die W' für A, wenn bereits B eingetreten ist.“ Erläutern Sie damit inhaltlich - anschaulich die Gleichung unter b, c und d.

Ü4 In einem bayrischen Touristenort sind zur Hochsaison drei Mal so viele Touristen wie Einheimische. Touristen tragen zu 70% einen Tirolerhut, Einheimische nur zu 25%.

a. Sie fragen einen Menschen mit Tirolerhut nach dem Weg. Wie groß ist die W', dass der Mensch ein Einheimischer ist?
b. Sie fragen einen Menschen ohne Tirolerhut nach dem Weg. Wie groß ist die W', dass der Mensch ein Einheimischer ist?

Was ist also günstiger, wenn Sie möglichst schnell eine verlässliche Wegauskunft haben möchten?

Ü5 Eine Krankheit kommt in der Bevölkerung mit einer W' von 2% vor. Der Test für die Krankheit ist zurzeit noch sehr unsicher, er zeigt nämlich sowohl für die Kranken wie auch die Gesunden ihren Zustand mit einer Sicherheit von 75% an.

a. Der Test zeigt für eine Person das Vorhandensein der Krankheit an. Wie groß ist die W', tatsächlich Krank zu sein?
b. Die Firma möchte nun durch weitere Investitionen den Test genauer machen. Sie sollen entscheiden, ob die Investitionen eher in die genauere Indikation bei den Kranken oder bei den Gesunden fließen soll.

Rechnen Sie dazu die Situation unter a durch, wenn die Genauigkeit

i. für die Kranken (die Sensitivität) von 75% auf 90% gesteigert wurde.
ii. für die Gesunden (die Spezifität) von 75% auf 90% gesteigert wurde.

Wohin sollte also die Investition fließen?