

# Optimierungsverfahren für alternative statistische Verteilungen

MANFRED BÖRGENS, WÖLFERSHEIM

**Zusammenfassung:** Das im Folgenden beschriebene stochastische Optimierungsverfahren schlägt eine Brücke von den Grundlagen der Wahrscheinlichkeitsrechnung zu vielfältigen Anwendungen, z. B. in der Medizin. Das Verfahren beruht in seinem Kern auf dem Satz von Bayes und verfolgt das Ziel, bestimmte Entscheidungen unter Unsicherheit zu fällen. Eine Voraussetzung dafür ist das Vorliegen empirischer oder theoretischer statistischer Verteilungen. Für stetige Verteilungen wird eine Verbindung von Stochastik und Differentialrechnung hergestellt. Der Rechenweg für die Optimierung lässt sich z. B. in Excel® implementieren.

Das Optimierungsverfahren wird auch in der Statistischen Qualitätssicherung eingesetzt. Diese Anwendung wird in einem Folgeartikel vorgestellt.

## Einführung

Optimierungsverfahren werden im Schulunterricht außerhalb der Differentialrechnung nur wenig behandelt. Die Stochastik bietet aber verschiedene Ansätze für die Behandlung von Optimalitätsproblemen, die sich unter dem Begriff „Entscheidungen unter Unsicherheit“ zusammenfassen lassen und von praktischer Relevanz sind, siehe z. B. Rice (2007, ch. 9). Für die Umsetzung im Mathematikunterricht der Sekundarstufe 2 werden dafür nur stochastische Grundkenntnisse vorausgesetzt (bedingte Wahrscheinlichkeiten, Satz von Bayes, empirische Verteilungen und gängige Wahrscheinlichkeitsverteilungen, u. a. Normalverteilung).

Es soll zunächst nicht die Methodik im Vordergrund stehen, sondern es werden zwei Fallbeispiele vorgestellt, um in die Problemstellung einzuführen. Das erste Fallbeispiel basiert auf diskreten, das zweite auf stetigen Verteilungen. Sie sollen zeigen, dass es um optimale Entscheidungen geht, die auf der Grundlage von Risikobewertungen getroffen werden. Bei beiden Beispielen sind die verwendeten Wahrscheinlichkeitswerte fiktiv gewählt und stammen nicht aus einer empirischen Studie.

Wie aus den Beispielen deutlich werden wird, stehen am Anfang des Optimierungsverfahrens immer zwei Verteilungen, die sich auf dieselbe Zufallsvariable (ZV) beziehen. Die eine Verteilung entsteht, wenn die ZV bei Vorliegen eines bestimmten Ereignisses E erhoben wird, die andere unter der Bedingung  $\neg E$ .

Man kann dabei beispielsweise an die ZV „Blutdruck“ denken, die bei Vorliegen einer bestimmten Erkrankung eine andere Wahrscheinlichkeitsverteilung in der Bevölkerung hat als bei Menschen ohne diese Erkrankung.

Hinter den einführenden Beispielen wird dann eine Methode zur Vorhersage des Ereignisses E entwickelt. E ist für einen Beobachter in der Regel nicht leicht erkennbar, aber eine Prognose für das Vorliegen von E wird durch die Erhebung eines Wertes der ZV ermöglicht. Das Kriterium für die Optimalität der Entscheidung auf E oder  $\neg E$  beruht auf den Erträgen für richtige bzw. falsche Entscheidungen. Um das Verfahren rechnerisch umsetzbar zu machen, wird der Satz von Bayes eingesetzt (Rice 2007, ch. 1.5, ch. 9).

## 1 Fallbeispiel „Erdbebenwarnung“ – Problemstellung

In einem erdbebengefährdeten Gebiet gibt ein seismologisches Institut Vorwarnstufen 1 bis 6 bekannt. Solche Vorwarnungen beruhen auf der Erhebung von seismischen Messwerten, sind aber nur von begrenzter Zuverlässigkeit. Empirisch lässt sich aber feststellen: Je höher die Warnstufe, desto größer ist die Wahrscheinlichkeit, dass in den nächsten 24 Stunden ein Erdbeben auftritt. Dies lässt sich aus den relativen Häufigkeitsverteilungen in Tabelle 1 ablesen.

	kein Beben	Beben
1	0,35	0,05
2	0,2	0,1
3	0,15	0,15
4	0,15	0,2
5	0,1	0,2
6	0,05	0,3

Tabelle 1: Relative Häufigkeiten für Vorwarnstufen

Die relativen Häufigkeiten sind nicht zeilenweise, sondern spaltenweise zu lesen. Ein Beispiel: In den Fällen, in denen einer Vorwarnung kein Beben folgte, kam die Warnstufe 4 mit der relativen Häufigkeit 0,15 vor.

Abbildung 1 zeigt die Häufigkeitsverteilungen grafisch.

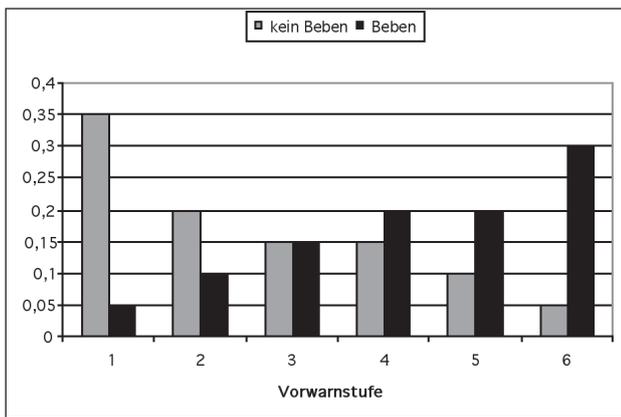


Abbildung 1: Empirische Häufigkeiten für Vorwarnstufen

Wie reagiert ein Unternehmer, dessen Lagerraum erdbebengefährdet ist, auf eine Vorwarnung? Es gibt die Möglichkeit, das Lager zu räumen und den wertvollen Inhalt zeitweise an einen sicheren Ort zu bringen. Wie lässt sich berechnen, ab welcher Vorwarnstufe sich das lohnt? Offenbar benötigt man dafür noch weitere Informationen, nämlich die Auftretenshäufigkeit von Erdbeben nach einer Vorwarnung und die Kosten für eine Lagerräumung bzw. für den Verlust des Lagerguts.

## 2 Fallbeispiel „Diagnose“ – Problemstellung

Eine physiologische Größe wird als möglicher Indikator für eine bestimmte Krankheit vermutet und deshalb in einer Testreihe bei Kranken und Gesunden gemessen. Die Messwerte erweisen sich in beiden Gruppen empirisch als (näherungsweise) normalverteilt mit gleicher Varianz, siehe Abb. 2. Will man diesen physiologischen Wert zur Diagnose heranziehen, so stellt sich die Frage, wie groß der Messwert ausfallen muss, um eine Therapie zu rechtfertigen.

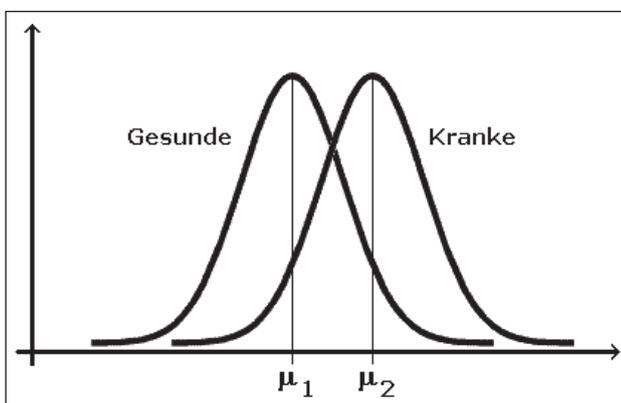


Abbildung 2: Wahrscheinlichkeitsdichten für eine physiologische Größe

Auch hier sind weitere Informationen erforderlich, u. a. die Auftretenshäufigkeit der Krankheit. Schwierig (aber erforderlich) ist die Bewertung der Kosten und Risiken, etwa bei der Frage, welche Schäden durch die Behandlung von Gesunden entstehen.

## 3 Vierfeldertafel

Für einen systematischen Ansatz, der u. a. die beiden Fallbeispiele abdeckt, müssen zunächst die Konsequenzen von richtigen und falschen Entscheidungen bewertet werden, dies wird auch als *Risikobewertung* bezeichnet. In einer Vierfeldertafel werden dafür *Erträge* festgelegt. Positive Erträge sind Gewinne, Einnahmen o.ä., negative Erträge sind Kosten, Verluste o.ä. Schaut man sich die Fallbeispiele an, so wird klar, dass die Festlegung der Erträge meistens nicht einfach sein wird, aber selbst Schätzwerte sind immer noch besser als der völlige Verzicht auf ein Optimierungsverfahren.

Bei den hier behandelten Problemen treten *dichotome* Ereignisse auf, also Paare von Ereignissen ( $\neg E$ ,  $E$ ). Das erste Ereignispaar liegt in der Zukunft oder ist zumindest zum Zeitpunkt des Optimierungsverfahrens unbekannt. Wir bezeichnen es mit  $(G, S)$ , was die Interpretation „Gut“ / „Schlecht“ nahelegt. Tabelle 2 zeigt, was dies für unsere Fallbeispiele bedeuten würde.

	$\neg E = G$	$E = S$
Erdbebenwarnung	kein Erdbeben	Erdbeben
Diagnose	gesund	krank

Tabelle 2: Dichotome Ereignisse bei den Fallbeispielen

Natürlich ist es willkürlich, ob man  $G$  als  $\neg E$  oder  $E$  wählt. In beiden Fallbeispielen ist  $S$  das mehr interessierende, also vorherzusagende Ereignis (Erdbeben, Krankheit), deshalb wird hier  $E = S$  gewählt.

Das zweite Ereignispaar betrifft die Entscheidung. Die Entscheidung wollen wir  $A$  nennen, wenn aufgrund des Messwertes vom Vorliegen von  $G$  ausgegangen wird, ansonsten lautet die Entscheidung  $Z$ . Das Ereignispaar  $(A, Z)$  soll die Interpretation „Annahme“ / „Zurückweisung“ nahelegen. Tabelle 3 zeigt dies für die beiden Fallbeispiele.

	Erdbebenwarnung	Diagnose
Messwert	Vorwarnstufe	physiol. Wert
$A$	keine Aktion	keine Aktion
$Z$	Lager räumen	Therapie

Tabelle 3: Entscheidungen bei den Fallbeispielen

Tabelle 4 zeigt die allgemeine Form der Vierfeldertafel und ein Beispiel für die Erdbebenwarnung in Abschnitt 1. Dabei sind  $h_{AG}$ ,  $h_{AS}$ ,  $h_{ZG}$  und  $h_{ZS}$  die Erträge, die für die vier kombinierten Ereignisse  $A \wedge G$ ,  $A \wedge S$ ,  $Z \wedge G$  und  $Z \wedge S$  (im Mittel) zu erwarten sind. Es wurde angenommen, dass die Schädigung des Lagerinhalts bei einem Erdbeben im Mittel 100 Geldeinheiten kostet, die Räumung des Lagers dagegen nur 3 Geldeinheiten.

	G	S
A	$h_{AG}$	$h_{AS}$
Z	$h_{ZG}$	$h_{ZS}$

	G	S
A	0	-100
Z	-3	-3

Tabelle 4: Vierfeldertafel (rechts für Fallbeispiel 1)

#### 4 Optimale Entscheidung

Es soll eine Entscheidung getroffen werden, ob man von einem (noch unbekanntem) Ereignis G ausgehen soll (Entscheidung A) oder von  $S = \neg G$  (Entscheidung  $Z = \neg A$ ).

Die Kombinationen  $A \wedge G$  und  $Z \wedge S$  bedeuten also richtige Entscheidungen,  $A \wedge S$  und  $Z \wedge G$  Fehlentscheidungen (außer in den seltenen Fällen, in denen A und Z die gleiche Wirkung haben).

Die Konsequenzen (in der Regel *Risiken*) der Entscheidung lassen sich in einer Vierfeldertafel (siehe Abschnitt 3) für die richtigen und falschen Entscheidungen kostenmäßig bewerten. Daraus ergibt sich die Baumstruktur in Abb. 3; dort stehen die Wahrscheinlichkeiten entlang der Kanten des Graphen und an den Endknoten. Die Wahrscheinlichkeiten werden ab Abschnitt 5 berechnet. Auf der unteren Ebene sind richtige Entscheidungen und Fehlentscheidungen symbolisch markiert.

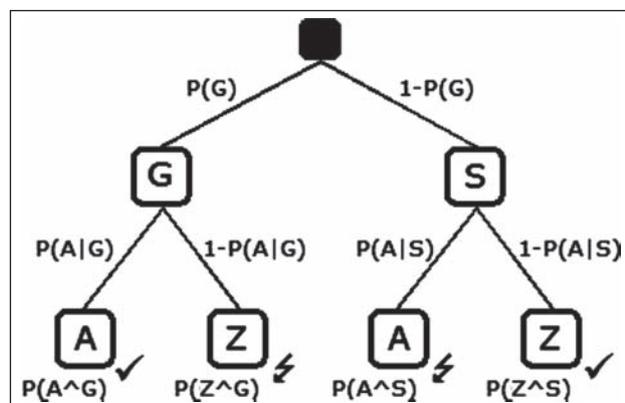


Abbildung 3: Allgemeine Baumstruktur für das Optimierungsproblem

Nun soll hergeleitet werden, wie die Entscheidungen A und Z zu treffen sind.

Bisher war in den Fallbeispielen immer von empirischen Messwerten die Rede. Nun sollen aber die erhobenen Werte als Realisationen einer Zufallsvariablen (ZV)  $X$  verstanden werden.  $X$  ist im Fallbeispiel 1 die Warnstufe und im Fallbeispiel 2 die physiologische Größe. Eine konkrete Realisation  $X = x$  soll die Entscheidung A / Z zwischen  $\neg E = G$  und  $E = S$  bestimmen. Der Schluss von  $X = x$  auf A bzw. Z soll ertragsoptimal erfolgen.

$P(C)$  soll im folgenden für die Wahrscheinlichkeit eines Ereignisses C stehen,  $P(C | B)$  für die bedingte Wahrscheinlichkeit von C unter der Bedingung B.

Es liege nun die konkrete Realisation  $X = x$  vor. Welchen Ertrag hat man im Mittel zu erwarten, wenn die Entscheidung A getroffen wird? Das ist ein einfacher Erwartungswert, der sich als das folgende gewichtete Mittel von  $h_{AG}$  und  $h_{AS}$  schreiben lässt:

$$P(G | X = x) \cdot h_{AG} + P(S | X = x) \cdot h_{AS}$$

Der Ertrag bei Entscheidung Z ist entsprechend

$$P(G | X = x) \cdot h_{ZG} + P(S | X = x) \cdot h_{ZS}$$

Die Entscheidung A ist im Vergleich zu Z günstiger oder gleich günstig, falls

$$(1) \quad P(G | X = x) \cdot h_{AG} + P(S | X = x) \cdot h_{AS} \geq P(G | X = x) \cdot h_{ZG} + P(S | X = x) \cdot h_{ZS}$$

Für die Entscheidung Z ist das Ungleichheitszeichen umzukehren.

Nun sind aber  $P(G | X = x)$  und  $P(S | X = x)$  unbekannt und lassen sich auch aus den vorhandenen Verteilungen nicht direkt ablesen. Hier wird ein Ansatz nach Bayes nötig werden (siehe nächste Abschnitte). Zunächst soll aber (1) rechnerisch vereinfacht werden:

(1) ist äquivalent zu

$$(2) \quad \beta \cdot P(G | X = x) \geq P(S | X = x)$$

$$\text{mit } \beta = (h_{AG} - h_{ZG}) / (h_{ZS} - h_{AS}).$$

Es wird  $h_{ZS} > h_{AS}$  vorausgesetzt. Das ist sinnvoll, denn die richtige Vorhersage von S hat einen höheren Wert als das „Verfehlen“ von S. In der Regel wird auch  $h_{AG} > h_{ZG}$  sein, aber die Formel ist ebenso richtig für  $h_{AG} = h_{ZG}$ , d. h.  $\beta = 0$ .

Was haben wir bis hierhin bewiesen? **Wir treffen die Entscheidung A genau dann, wenn für die konkrete Realisation  $X = x$  die Ungleichung (2) erfüllt ist.** Da  $P(G | X = x)$  noch unbekannt ist, sind weitere Überlegungen erforderlich, die für diskrete und stetige ZV  $X$  getrennt in den Abschnitten 5 und 7 geführt werden.

## 5 Diskrete Wahrscheinlichkeitsverteilungen

Für die Wahrscheinlichkeitsverteilungen der ZV  $X$  bei Vorliegen von  $E$  bzw.  $\neg E$  übernimmt man die empirischen Verteilungen (z. B. aus dem Fallbeispiel in Abschnitt 1). Diese Verteilungen benennen wir wie folgt:

$$p_G(x) = P(X = x | G)$$

$$p_S(x) = P(X = x | S)$$

Für das Fallbeispiel in Abschnitt 1 ergeben sich die gleichen Wertetabellen wie in Tabelle 1.

Nun folgt der Bayes'sche Ansatz für das vorliegende Problem.

$G/S$  sind konkurrierende Hypothesen im Bayes'schen Sinne. Wir betrachten zuerst  $G$ :

$P(G)$  a-priori-Wahrscheinlichkeit

$P(X = x | G)$  Likelihood

$P(G | X = x)$  a-posteriori-Wahrscheinlichkeit

Hier taucht die letzte erforderliche Komponente für die notwendige Vorinformation auf (nach Verteilungen und Vierfeldertafel), nämlich  $P(G)$ . Für die Lösung der Fallbeispiele wird das bedeuten, dass die Auftretenshäufigkeiten von Erdbeben bzw. Krankheit nachgeliefert werden müssen.

Nach dem Satz von Bayes gilt:

$$P(G | X = x) = P(X = x | G) \cdot P(G) / P(X = x)$$

Für  $S$  ergibt sich analog:

$$P(S | X = x) = P(X = x | S) \cdot P(S) / P(X = x)$$

Die Likelihoods wurden bereits mit Hilfe der Wahrscheinlichkeitsverteilungen  $p_G$  und  $p_S$  ausgedrückt. Dies setzen wir in die beiden Bayes'schen Formeln ein:

$$P(G | X = x) = p_G(x) \cdot P(G) / P(X = x)$$

$$P(S | X = x) = p_S(x) \cdot P(S) / P(X = x)$$

Also ist (2) äquivalent zu:

$$(3) \alpha \cdot p_G(x) \geq p_S(x) \text{ mit } \alpha = \beta \cdot P(G) / P(S).$$

Dies ist der entscheidende Schritt: Die unbekanntenen  $P(G | X = x)$  und  $P(S | X = x)$  in (2) konnten mit der Bayes'schen Formel durch bekannte Terme ausgedrückt werden, weil  $P(X = x)$  beim Einsetzen in (2) weggefallen ist.

*(1) – (3) sind also äquivalente Beschreibungen für die Entscheidung A. Durch Umkehrung der Ungleichheitszeichen erhält man die Entscheidungsregel für Z.*

## 6 Fallbeispiel „Erdbebenwarnung“ – Lösung

$\beta$  wurde in (2) definiert. Mit Tabelle 4 erhalten wir  $\beta = 3/97$ .

Einer Vorwarnung möge ein Erdbeben empirisch mit Wahrscheinlichkeit  $P(S) = 0,12$  folgen.

So erhalten wir mit Abschnitt 5  $\alpha = 22/97 \approx 0,227$ .

Nun liegen alle Informationen vor, um (3) anwenden zu können.

$x$	$\alpha p_G(x)$	$p_S(x)$	Entscheidung
1	0,0794	0,05	A
2	0,0454	0,10	Z
3	0,0340	0,15	Z
4	0,0340	0,20	Z
5	0,0227	0,20	Z
6	0,0113	0,30	Z

Tabelle 5: Entscheidungstafel für Fallbeispiel 1

**Optimale Entscheidung: Erdbebenwarnung der Stufe 1 ignorieren, ab Stufe 2 Lagerraum räumen.**

Die Optimalität der Entscheidung bezieht sich aufgrund der Herleitung in Abschnitt 4 auf die Minimierung der Kosten. Das wird in Abschnitt 9 noch vertieft.

Mit Hilfe der Entscheidungsregel (3) lässt sich der Baum aus Abb. 3 mit Inhalt füllen. Die Wahrscheinlichkeiten auf der oberen Ebene stehen für empirische Werte (hier fiktiv gewählt). Setzt man  $c$  als obere Grenze für  $x$  bei Entscheidung A (hier  $c = 1$ ), so ist im Baumdiagramm auf der zweiten Ebene  $P(A|G) = \sum_{x \leq c} p_G(x)$  und  $P(A|S) = \sum_{x \leq c} p_S(x)$ . Die ausmultiplizierten Wahrscheinlichkeiten in der untersten Zeile werden in **Abschnitt 9** zur Berechnung des Gesamtertrags verwendet. Das Ergebnis mit anschaulichen Interpretationen der Ereignisse zeigt Abb. 4.

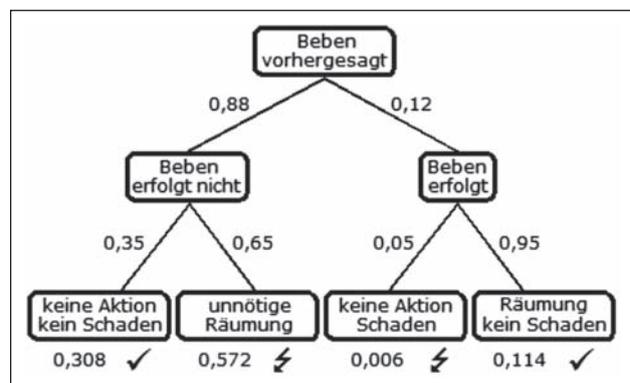


Abbildung 4: Wahrscheinlichkeitsbaum für Fallbeispiel 1

## 7 Stetige Wahrscheinlichkeitsverteilungen

Hat die ZV eine stetige Wahrscheinlichkeitsverteilung (wie z. B. in Abschnitt 2), so muss wegen  $P(X = x) = 0$  definiert werden, was unter  $P(G | X = x)$  zu verstehen ist:

$$P(G | X = x) = \lim_{\delta \rightarrow 0} P(G | E_\delta).$$

$E_\delta$  steht als Kurzform für das Ereignis

$$x - \delta < X < x + \delta.$$

Analog wird definiert:

$$P(S | X = x) = \lim_{\delta \rightarrow 0} P(S | E_\delta).$$

Natürlich muss noch nachgewiesen werden, dass diese Grenzwerte existieren; dies schlägt eine Brücke von der Stochastik zur Analysis.

Der Bayes'sche Ansatz ist hier weitgehend analog zum diskreten Fall:

$P(G)$  a-priori-Wahrscheinlichkeit

$P(E_\delta | G)$  Likelihood

$P(G | E_\delta)$  a-posteriori-Wahrscheinlichkeit

Analog zu Abschnitt 5 seien  $p_G$  und  $p_S$  die Wahrscheinlichkeitsdichten für  $X$ , je nachdem, ob  $G$  oder  $S$  vorliegt.  $f_G$  und  $f_S$  seien die zugehörigen stetigen Verteilungsfunktionen.

Dann ist nach dem Satz von Bayes und dem Satz von der totalen Wahrscheinlichkeit:

$$\begin{aligned} P(G | E_\delta) &= \frac{P(E_\delta | G) \cdot P(G)}{P(E_\delta)} = \frac{P(E_\delta | G) \cdot P(G)}{P(E_\delta | G) \cdot P(G) + P(E_\delta | S) \cdot P(S)} = \\ &= \left( 1 + \frac{P(E_\delta | S) \cdot P(S)}{P(E_\delta | G) \cdot P(G)} \right)^{-1} = \left( 1 + \frac{f_S(x + \delta) - f_S(x - \delta)}{f_G(x + \delta) - f_G(x - \delta)} \cdot \frac{P(S)}{P(G)} \right)^{-1}. \end{aligned}$$

Mit der Regel von l'Hospital folgt:

$$\begin{aligned} P(G | X = x) &= \lim_{\delta \rightarrow 0} P(G | E_\delta) = \lim_{\delta \rightarrow 0} \left( 1 + \frac{p_S(x + \delta) + p_S(x - \delta)}{p_G(x + \delta) + p_G(x - \delta)} \cdot \frac{P(S)}{P(G)} \right)^{-1} = \\ &= \left( 1 + \frac{2 \cdot p_S(x)}{2 \cdot p_G(x)} \cdot \frac{P(S)}{P(G)} \right)^{-1} = \frac{p_G(x) \cdot P(G)}{p_G(x) \cdot P(G) + p_S(x) \cdot P(S)}. \end{aligned}$$

Für  $p(S | X = x)$  ist im Zähler  $G$  durch  $S$  zu ersetzen. Mit (2) folgt dann wegen der Gleichheit der Nenner:

$$\beta \cdot p_G(x) \cdot P(G) \geq p_S(x) \cdot P(S) \Leftrightarrow \alpha \cdot p_G(x) \geq p_S(x)$$

Also gilt auch hier (3) als äquivalente Beschreibung der Entscheidung  $A$ .

## 8 Fallbeispiel „Diagnose“ – Lösung

In Abschnitt 2 möge die Krankheit (beispielsweise) bei 9% der Bevölkerung vorliegen. Tabelle 6 zeigt eine mögliche Vierfeldertafel.

	G	S
A	0	-100
Z	-10	100

Tabelle 6: Vierfeldertafel für Fallbeispiel 2

$h_{AG} = 0$  : Eine gesunde Person wird nicht behandelt.

$h_{AS} = -100$  : Eine kranke Person wird nicht behandelt. -100 steht für die Folgekosten.

$h_{ZG} = -10$  : Eine gesunde Person wird behandelt. -10 steht für die Kosten der Behandlung und der auftretenden Nebenwirkungen der (unnötigen) Behandlung.

$h_{ZS} = 100$  : Eine kranke Person wird behandelt. 100 steht für den Wert der Heilung (Arbeitsfähigkeit, Lebensqualität, Vermeidung oder Reduzierung von Folgekosten) abzüglich der Kosten der Behandlung.

Man berechnet  $\beta = \frac{1}{20}$ . Wegen  $P(S) = \frac{9}{100}$  folgt

$$\alpha = \frac{91}{180}.$$

Die beiden alternativen Normalverteilungen aus Abschnitt 2 sollen  $\mu_G = 120$ ,  $\mu_S = 160$  und  $\sigma = 30$  aufweisen. Dann ist

$$\frac{p_S(x)}{p_G(x)} = \exp\left(\frac{(x-120)^2 - (x-160)^2}{2 \cdot 30^2}\right) = \exp\left(\frac{2}{45} \cdot x - \frac{56}{9}\right).$$

Auf A wird nach (3) entschieden bei  $p_S(x)/p_G(x) \leq \alpha$ , also hier bei  $x \leq 124,653$ .

**Bei Messwerten unterhalb 124,653 wird auf eine Behandlung verzichtet, ansonsten wird die untersuchte Person therapiert.**

Mit Hilfe dieser Entscheidungsregel lässt sich der Baum aus Abb. 3 mit Inhalt füllen. Die Wahrscheinlichkeiten auf der oberen Ebene stehen für empirische Werte (hier fiktiv gewählt). Setzt man  $c$  als obere Grenze für  $x$  bei Entscheidung A (hier  $c = 124,653$ ), so ist im Baumdiagramm auf der zweiten Ebene:

$$P(A|G) = \int_{-\infty}^c p_G(x) dx = f_G(c)$$

$$P(A|S) = \int_{-\infty}^c p_S(x) dx = f_S(c)$$

Das Ergebnis mit anschaulichen Interpretationen der Ereignisse zeigt Abb. 5.

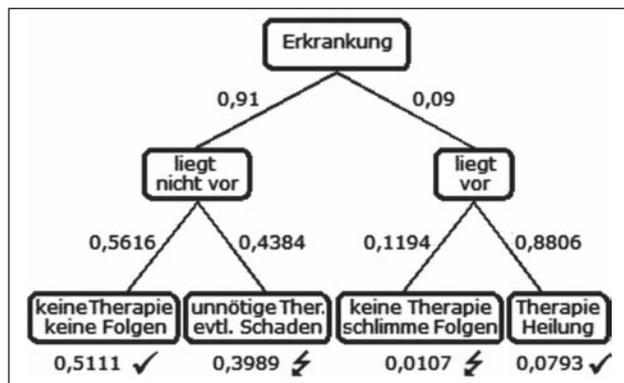


Abbildung 5: Wahrscheinlichkeitsbaum für Fallbeispiel 2

## 9 Erwarteter Ertrag

Den Erwartungswert für den Ertrag wollen wir mit  $H$  bezeichnen. Mit Hilfe von Tabelle 4 erhält man:

$$\begin{aligned} (4) \quad H &= h_{AG} \cdot P(A \wedge G) + h_{AS} \cdot P(A \wedge S) \\ &\quad + h_{ZG} \cdot P(Z \wedge G) + h_{ZS} \cdot P(Z \wedge S) \\ &= h_{AG} \cdot P(A | G) \cdot P(G) + h_{AS} \cdot P(A | S) \cdot p(S) \\ &\quad + h_{ZG} \cdot P(Z | G) \cdot p(G) + h_{ZS} \cdot P(Z | S) \cdot p(S) \end{aligned}$$

Wir wollen  $H$  für die Fallbeispiele berechnen.

### Fallbeispiel 1

Mit (4) und den Werten aus Tabelle 4 und Abb. 4 berechnet man  $H \approx -2,658$ , siehe Abb. 6.

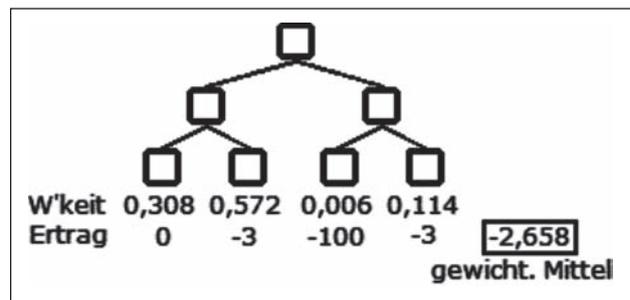


Abbildung 6: Erwarteter Ertrag bei Fallbeispiel 1

Als durchschnittliche **Kosten** pro Erdbebenwarnung sind 2,658 zu erwarten – unter der Voraussetzung, dass der Anweisung gefolgt wird (d. h. in Kurzform:  $c = 1$ ). Dies sind dann die minimal möglichen durchschnittlichen Kosten überhaupt.

Das Verständnis der Lernenden für das dargestellte Verfahren kann unterstützt werden durch die Berechnung des erwarteten Ertrags bei Variation der Grenze  $c$ . Hebt man diese Grenze beispielsweise von  $c = 1$  auf  $c = 2$  an, so bedeutet das inhaltlich einen weniger vorsichtigen Umgang mit der Erdbebenwarnung, denn dann würden auch Warnungen der Stufe 2 ignoriert. Aus Abschnitt 6 übernimmt man für  $c = 2$ :

$$P(A | G) = \sum_{x \leq c} p_G(x) = 0,55 \quad (\Leftrightarrow P(Z | G) = 0,45)$$

$$P(A | S) = \sum_{x \leq c} p_S(x) = 0,15 \quad (\Leftrightarrow P(Z | S) = 0,85)$$

Dies ergibt eingesetzt in (4):

$$H = -3,294$$

Das Ignorieren der Warnstufe 2 führt also zu deutlich höheren erwartenden Kosten.

### Fallbeispiel 2

Nach Abschnitt 8 lautet die Entscheidung:

A (keine Behandlung) bei  $x \leq 124,653 =: c$ .

$$P(A | \bullet) = \int_{-\infty}^c p_{\bullet}(x) dx = f_{\bullet}(c)$$

$$P(Z | \bullet) = \int_c^{+\infty} p_{\bullet}(x) dx = 1 - f_{\bullet}(c)$$

Diese Integrale lassen sich mit der Funktion NORMVERT in Excel® berechnen; so erhält man Tabelle 7. Die Wahrscheinlichkeiten wurden bereits in Abb. 5 verwendet. Auf eine Darstellung wie in Abb. 6 wird hier verzichtet, da für (4) bereits die bedingten Wahrscheinlichkeiten genügen.

$P(A   G) \approx 0,5616$	$h_{AG} = 0$
$P(A   S) \approx 0,1194$	$h_{AS} = -100$
$P(Z   G) \approx 0,4384$	$h_{ZG} = -10$
$P(Z   S) \approx 0,8806$	$h_{ZS} = 100$
$P(G) = 0,91$	$P(S) = 0,09$
<b><math>H \approx 2,86</math></b>	

Tabelle 7: Erwarteter Ertrag bei Fallbeispiel 2

Als durchschnittlicher **Gewinn** pro Patient ist ca. 2,86 zu erwarten – unter der Voraussetzung, dass der Anweisung gefolgt wird (d. h. in Kurzform:  $c = 124,653$ ). Dies ist dann der maximal mögliche durchschnittliche Gewinn überhaupt.

Auch hier kann die Variation der Grenze  $c$  das Verfahren für die Lernenden transparent machen. Setzt man etwa  $c = 130$ , so bedeutet das inhaltlich, dass die Therapie seltener verordnet wird. Die Auswertung der Integrale ergibt dann:

$$p(A | G) \approx 0,6306$$

$$p(A | S) \approx 0,1587$$

$$p(Z | G) \approx 0,3694$$

$$p(Z | S) \approx 0,8413$$

Dies ergibt eingesetzt in (4):

$$H \approx 2,78.$$

Die zurückhaltendere Anwendung der Therapie lässt somit den zu erwartenden Ertrag sinken.

Bei der Anwendung von (4) in den beiden Fallbeispielen wird die Berechnung des erwarteten Ertrages  $H$  dadurch erleichtert, dass sich die Messwerte durch  $c$  disjunkt in zwei Teilintervalle (für A und für Z) zerlegen lassen.

Es gibt aber Anwendungsbeispiele, in denen die Messwertbereiche für A und Z nicht diese einfache Struktur haben. Dadurch wird die Verwendung von (4) aber nicht beeinträchtigt. Ein gängiges Beispiel dafür sind zwei Normalverteilungen für  $p_G$  und  $p_S$  mit gleichem Erwartungswert, aber unterschiedlichen Varianzen. Dann zerfallen die Messwerte in der Regel in drei Intervalle  $(-\infty, c_1)$ ,  $(c_1, c_2)$ ,  $(c_2, +\infty)$ ; die beiden äußeren stehen für A und das innere für Z – oder umgekehrt.

## 10 Resümee

Es wurde ein Verfahren vorgestellt, das optimale Entscheidungen zwischen zwei Alternativen G und S ermöglicht. Das Optimalitätskriterium ist die Maximierung des erwarteten Ertrags. Als „Input“ werden zwei Wahrscheinlichkeitsverteilungen für dieselbe ZV benötigt, je eine unter der Bedingung G bzw. S. Für G und S müssen a-priori-Wahrscheinlichkeiten bekannt sein. Die Entscheidung kann A lauten („Annahme“ von G) oder Z („Zurückweisung“ von G). Die Erträge für die vier Kombinationen von G/S und A/Z gehören ebenfalls zum Input (4-Felder-Tafel). Damit erhält man eine Entscheidungsregel für A/Z (siehe Abb. 7). Diese lässt sich im Unterricht mit stochastischen Grundkenntnissen herleiten und bietet zusätzlich die Möglichkeit, Elemente der Analysis einzubinden.

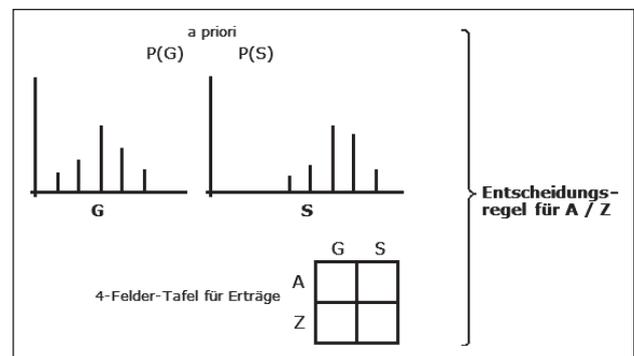


Abbildung 7: Übersicht über das Verfahren

## Literatur

Rice, J. A. (2007): *Mathematical Statistics and Data Analysis*, 3rd ed. Belmont: Duxbury.

## Anschrift des Verfassers

Manfred Börgens  
 Technische Hochschule Mittelhessen  
 Wilhelm-Leuschner-Straße 13  
 61169 Friedberg  
 manfred.boergens@mnd.thm.de