Kapitel 10

Entscheidungen bei Risiko

1. Einordnung der Entscheidungen bei Risiko

Wie in der Betriebswirtschaftslehre schlechthin findet sich auch in der Entscheidungstheorie das Nebeneinander deskriptiver und normativer Theorien. Gegenstand deskriptiver Entscheidungsforschung ist die Identifizierung von Situationsmerkmalen, die bestimmte tatsächliche Entscheidungen erklären. Auch der ***deskriptiven Entscheidungstheorie*** liegen Modelle zugrunde: Modelle, welche die Suche nach einschlägigen Situationsmerkmalen erleichtern, und Modelle, welche eine Hypothesenbildung erlauben, die einer empirischen Überprüfung zugänglich sind. Die (praktisch-) ***normative Entscheidungstheorie*** dient dazu, Entscheidungsempfehlungen abzuleiten. Demnach ist zu erklären, wie ausgehend von einer fixierten Zielsetzung aus einer Menge von Handlungsalternativen diejenige mit der besten Zielerreichung ausgewählt werden kann. Die praktisch-normative Entscheidungstheorie steht im Mittelpunkt der folgenden Überlegungen. Die vorgegebenen Zielsetzungen sind die Einkommenserzielung und die Verringerung von Einkommensunsicherheiten. Die geeignete Umsetzung dieser generellen Zielorientierung, zum Beispiel die angemessene Einbeziehung von Risikopräferenzen, ist Gegenstand dieses Kapitels.

Die Elemente des ***Grundmodells der Entscheidungstheorie*** sind bereits bekannt:[1] Handlungsalternativen, Umweltzustände (gleichbedeutend war auch von Zukunftsentwicklungen die Rede) und Ergebnisse. Je nachdem, wie viele Umweltzustände eintreten können und welche Informationen über Eintrittswahrscheinlichkeiten vorliegen, lassen sich Entscheidungen bei Sicherheit, bei Risiko und bei Ungewissheit unterscheiden. Im Weiteren werden ausschließlich Entscheidungen bei Risiko betrachtet. Es wird also vorausgesetzt, dass der Entscheider in der Lage ist, mindestens eine grobe subjektive Wahrscheinlichkeitsverteilung für die Zustände anzugeben. Der Fall sicherer Erwartungen ist als Grenzfall der degenerierten Wahrscheinlichkeitsverteilung darin enthalten.

Da Entscheidungen bei Risiko auf Basis von Wahrscheinlichkeitsurteilen untersucht werden, ist es erforderlich, mit einigen Grundlagen der Wahrscheinlichkeitsrechnung zu beginnen. Kalküle für Risikosituationen beruhen darauf.

---

[1] Siehe Kapitel 2, Abschnitt 1.5.

Anschließend werden besonders wichtige Entscheidungsprinzipien für Risikosituationen vorgestellt, nämlich das *Bernoulli*-Prinzip, das (μ, σ)-Prinzip und die stochastische Dominanz.

2. Wahrscheinlichkeitsrechnung

2.1 Zufallsvariablen und Wahrscheinlichkeitsverteilungen

> Eine ***Zufallsvariable*** ist eine Funktion, die jedem Elementarereignis aus einer Ereignismenge eine reelle Zahl zuordnet.

Die Ereignismenge ist der Definitionsbereich der Zufallsvariablen, die Menge der reellen Zahlen oder einer Teilmenge davon ist deren Wertebereich. In der Sprache des oben vorgestellten Grundmodells der Entscheidungstheorie[2] entspricht ein Elementarereignis einem Zustand. Die Ereignismenge kann zum Beispiel darin bestehen, dass die Marktentwicklung bei einem bestimmten Produkt normal, gut oder schlecht verlaufen kann. Das Elementarereignis „normale Marktentwicklung" entspricht einem Zustand in der Terminologie des Grundmodells der Entscheidungstheorie. Ein Ereignis kann aber auch mehrere Zustände enthalten, zum Beispiel das Ereignis: „Die Marktentwicklung ist schlecht oder normal." Eine Zufallsvariable ordnet einer jeden Ausprägung der Marktentwicklung eine bestimmte Absatzmenge zu (beispielsweise 300, 400 oder 120 Produkteinheiten je Tag).

Man spricht von einem ***zufälligen Vorgang***, wenn auch bei einem fixierten Bedingungsrahmen das Resultat nicht eindeutig vorhergesagt werden kann. Der Bedingungsrahmen besteht beispielsweise in den eingesetzten absatzpolitischen Maßnahmen, entscheidungstheoretisch also in den Aktionen. Im Mangel an Vorhersagbarkeit schlägt sich der Einfluss des Zufalls nieder. Für einen Statistiker zu ungenau, bei betriebswirtschaftlichen Fragestellungen aber akzeptabel ist es, den Zufall als die Wirkung all dessen zu beschreiben, über das keine präzisen Aussagen gemacht werden können. Gerade mit Blick auf die Anwendung in der Entscheidungstheorie ist schließlich hervorzuheben, dass eine Zufallsvariable nicht nur den Zufall bestimmt wird, sondern auch durch Entscheidungen beeinflusst werden kann, wenn auch eben nicht in einer sicheren Art und Weise.

Wesentlich für die quantitative Erfassung des Zufalls sind Wahrscheinlichkeiten. Als ***Wahrscheinlichkeitsverteilung*** bezeichnet man eine Zuordnung von Zahlen zu Ereignissen, die bestimmten Axiomen genügt:

2 Siehe dazu Kapitel 2, Abschnitt 1.5.

1. Die Wahrscheinlichkeit für den Eintritt irgendeines Zustands (oder allgemein: eines Ereignisses) ist nichtnegativ und nicht größer als Eins:

$$0 \leq w(z_j) \leq 1 \quad (j = 1, \dots, n),$$

wobei
$w(z_j)$ Wahrscheinlichkeit für den Eintritt des Zustands j.

2. Aufgrund der Vollständigkeit der Einteilung aller denkbaren Umweltentwicklungen muss irgendeiner der Umweltzustände eintreten. Das Ereignis „Eintritt irgendeines Zustands" ist demnach das sichere Ereignis. Die Wahrscheinlichkeit für den Eintritt des sicheren Ereignisses beträgt Eins:

$$w(s) = 1,$$

wobei
s sicheres Ereignis (Vereinigungsmenge aller Zustände: $s = \cup_j z_j$).

3. Per Konstruktion des Grundmodells der Entscheidungstheorie sind die einzelnen Zustände überschneidungsfrei abgegrenzt. Der Eintritt eines Zustands ist also mit dem Eintritt anderer Zustände unvereinbar. Die Wahrscheinlichkeit, dass von zwei Zuständen (oder allgemein: von zwei unvereinbaren Ereignissen) der eine oder der andere eintritt, ergibt sich deshalb aus der Summe der einzelnen Wahrscheinlichkeiten:

$$w(z_j \cup z_k) = w(z_j) + w(z_k) \quad (j, k = 1, \dots, n; k \neq j),$$

wobei
$z_j \cup z_k$ Ereignis, dass Zustand j oder Zustand k eintritt.

Von besonderer Bedeutung ist die Beachtung dieser Axiome vor allem, wenn subjektive Wahrscheinlichkeiten verwendet werden, was bei betriebswirtschaftlichen Fragestellungen häufig der Fall ist.

In der angegebenen Form gelten die Axiome für ***diskrete Zufallsvariablen***, also für solche Zufallsvariablen, die nur abzählbar viele Werte annehmen können. Daneben gibt es auch ***stetige Zufallsvariablen***; diese werden im Weiteren allerdings nicht explizit einbezogen.[3]

2.2 Wahrscheinlichkeits- und Verteilungsfunktion

Die Wahrscheinlichkeiten für den Eintritt bestimmter Ereignisse sind ein wichtiger Baustein der quantitativen Erfassung der Unsicherheit von Erwartungen. Letztlich geht es aber um die Erwartung bezüglich der Ergebnisse. Sie lassen sich durch die Wahrscheinlichkeitsfunktion und die Verteilungsfunktion beschreiben.

3 Vgl. dazu ausführlich *Bamberg/Baur/Krapp* (2012), Abschnitte 8.5 und 8.6.

> Die ***Wahrscheinlichkeitsfunktion*** $f(x)$ gibt an, mit welcher Wahrscheinlichkeit einzelne Ergebnisse x realisiert werden.

Im Beispiel der Marktentwicklung kann die Wahrscheinlichkeitsfunktion das folgende Aussehen haben:[4]

$$f(x) = w(\tilde{x} = x) = \begin{cases} 0{,}2 & \text{für} \quad x = 120 \\ 0{,}7 & \text{für} \quad x = 300, \\ 0{,}1 & \text{für} \quad x = 400 \end{cases}$$

wobei

x Marktentwicklung, gemessen in täglich absetzbaren Produkteinheiten.

Grundsätzlich ist zu fragen, welche Wahrscheinlichkeitsfunktion in einer bestimmten Situation zu unterstellen ist. Im vorgetragenen Beispiel kann sie das Ergebnis der Expertise eines Gutachters sein, die neben den Wahrscheinlichkeiten für die einzelnen Szenarien auch die Quantifizierung von „normaler", „schlechter" und „guter" Marktentwicklung umfassen muss. Realistisch gestellte Entscheidungsprobleme werden häufig nur durch derartig ***grobe Informationen*** beschreibbar sein. Möglich ist es aber auch, mangels anderer Anhaltspunkte alle denkbaren Umweltentwicklungen als gleich wahrscheinlich einzustufen oder, weil man es für eine ***akzeptable Annäherung*** an die Realität hält, von bestimmten, analytisch gut handhabbaren Wahrscheinlichkeitsfunktionen, beispielsweise der Normalverteilung, auszugehen.

Sofern es sich um eine diskrete Zufallsvariable handelt, ist auch die Wahrscheinlichkeitsfunktion eine diskrete Funktion. Ein Beispiel für eine operationale Wahrscheinlichkeitsverteilung ist die ***Gleichverteilung,*** die jedem denkbaren Ergebnis die gleiche Eintrittswahrscheinlichkeit zuordnet:

$$f(x_j) = \frac{1}{n} \quad (j = 1, \dots, n).$$

Die grafische Darstellung einer Wahrscheinlichkeitsfunktion entspricht einem Stabdiagramm (Abbildung 10.1).

Neben der Wahrscheinlichkeitsfunktion dient auch die Verteilungsfunktion der Beschreibung einer Wahrscheinlichkeitsverteilung.

> Die ***Verteilungsfunktion*** $F(x)$ gibt an, mit welcher Wahrscheinlichkeit ein bestimmter Ergebniswert x nicht überschritten wird.

[4] In diesem Abschnitt über die Wahrscheinlichkeitsrechnung werden zur Verdeutlichung Zufallsvariable $\tilde{x}$ mit Tilden gekennzeichnet, die entsprechenden Größen ohne Tilden stehen für bestimmte Ausprägungen (Realisationen) der Zufallsgrößen.

Für die Verteilungsfunktion an der Stelle x gilt

$$F(x) = w(\tilde{x} \leq x) = \sum_{e \leq x} w(e).$$

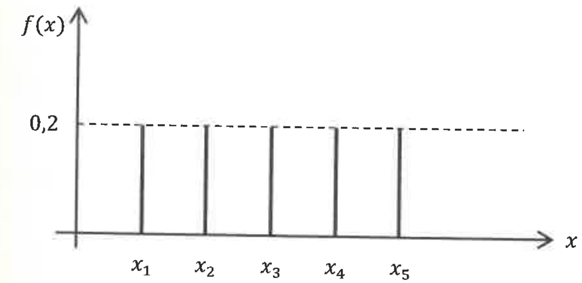

Abbildung 10.1: Wahrscheinlichkeitsfunktion der Gleichverteilung.

Aufgrund dieser Definitionen und der Axiome der Wahrscheinlichkeitsrechnung hat die Verteilungsfunktion einen Wertebereich von $[0; 1]$, außerdem ist sie (zumindest schwach) monoton steigend. Angewendet auf das Beispiel der Marktentwicklung erhält man

$$F(x) = \begin{cases} 0 & \text{für} \quad x < 120 \\ 0{,}2 & \text{für} \quad 120 \leq x < 300 \\ 0{,}9 & \text{für} \quad 300 \leq x < 400 \\ 1 & \text{für} \quad 400 \leq x \end{cases}$$

Die Verteilungsfunktion hat also die Gestalt einer Treppenfunktion. Die Abbildung 10.2 gibt als Beispiel die Verteilungsfunktion der Gleichverteilung wieder.

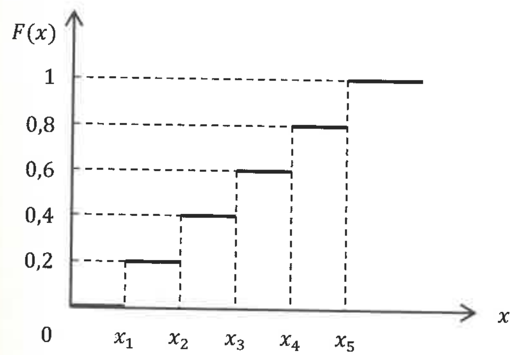

Abbildung 10.2: Verteilungsfunktion der Gleichverteilung.

Die Umkehrfunktion der Verteilungsfunktion ordnet den Wahrscheinlichkeitsniveaus die zugehörigen Quantile zu.

Das ***w-Quantil*** ist definiert als der kleinste Wert des Wertebereichs einer Zufallsvariablen, für den die Verteilungsfunktion mindestens den Wert w annimmt.

Das heißt, das w-Quantil ist der kleinste Ergebniswert x_w, der mit einer Wahrscheinlichkeit von mindestens w nicht überschritten wird. Es gilt also

$$x_w = \min\{x | F(x) \geq w\}.$$

Abbildung 10.3 verdeutlicht dies wieder anhand der Gleichverteilung:

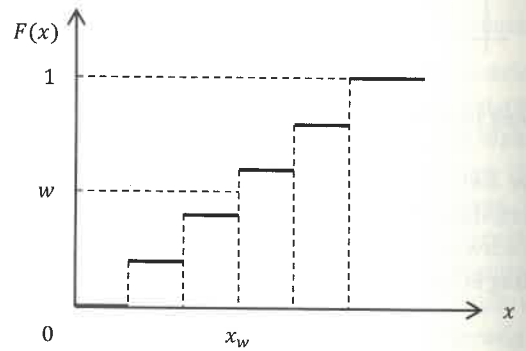

Abbildung 10.3: w-Quantil der Gleichverteilung.

Unter den Quantilen besonders hervorzuheben ist das 50%-Quantil, das als ***Median*** oder Zentralwert bezeichnet wird. Der Median hat bei Verteilungen mit einer stetigen Verteilungsfunktion die Eigenschaft, dass mit gleicher Wahrscheinlichkeit größere und kleinere Ergebnisse realisiert werden. Bei diskreten Verteilungen gilt dies nicht allgemein.

2.3 Parameter von Wahrscheinlichkeitsverteilungen

Mit Hilfe von Wahrscheinlichkeits- und Verteilungsfunktion werden *alle* Informationen über eine Wahrscheinlichkeitsverteilung erfasst. Allerdings kann das Arbeiten mit diesen Funktionen unhandlich sein. Problematisch ist zudem der hohe für die Schätzung einer kompletten Wahrscheinlichkeitsverteilung erforderliche ***Informationsbedarf***. Daher beschränkt man sich häufig auf bestimmte Parameter von Wahrscheinlichkeitsverteilungen, welche die wesentlichen Charakteristika der Verteilung wiedergeben. Einerseits ist damit, von wenigen Ausnahmen abgesehen, eine ***Einschränkung des Informationsgehalts*** gegeben, andererseits sinkt aber auch der Schätzaufwand. Dabei ist zu beachten, dass die Parameter direkt, ohne Rückgriff auf die Wahrscheinlichkeitsverteilung, geschätzt werden können.

Für die Eigenschaften einer Wahrscheinlichkeitsverteilung charakterisierend sind vor allem die Lage der Verteilung, deren Streuung und deren Symmetrie. Im Weiteren werden insbesondere die beiden erstgenannten Merkmale hervorgehoben.

Der ***Erwartungswert*** ist definiert als die mit den Eintrittswahrscheinlichkeiten gewichtete Summe der Ergebnisse.

$$\mathrm{E}\{\tilde{x}\} = \mu_x = \sum_{j=1}^{n} x_j f(x_j),$$

wobei

$\mathrm{E}\{\cdot\}$ Erwartungswertoperator[5]

μ_x Kurzschreibweise für den Erwartungswert.

Der Erwartungswert ist ein Lageparameter und steht für das mittlere Niveau von Ergebnissen, mit dem angesichts der Zufallseinflüsse gerechnet werden kann. Aus der Definition des Erwartungswertes ergibt sich, dass gleichermaßen bessere oder schlechtere Ergebnisse realisiert werden können. Deshalb ist die Streuung der Ergebnisse um den Erwartungswert herum eine wesentliche Eigenschaft einer Zufallsvariablen. Das wichtigste Streuungsmaß ist die Varianz.

Die ***Varianz*** ist definiert als Erwartungswert der quadrierten Abweichung einer Zufallsvariablen von ihrem Erwartungswert.

$$\mathrm{Var}\{\tilde{x}\} = \sigma_x^2 = \mathrm{E}\{(\tilde{x} - \mu_x)^2\},$$

wobei

$\mathrm{Var}\{\cdot\}$ Varianzoperator

σ_x^2 Kurzschreibweise für die Varianz.

Den gleichen Informationsgehalt wie die Varianz hat die ***Standardabweichung***. Sie ist definiert als positive Quadratwurzel der Varianz und wird mit σ_x symbolisiert. Die Standardabweichung ist häufig anschaulicher als die Varianz, weil sie dieselbe Dimension wie die Ergebnisse hat. Dagegen ist die Dimension der Varianz (bei Geldbeträgen zum Beispiel €^{2}) häufig unanschaulich. Die Varianz ist allerdings analytisch leichter zu handhaben. Aufgrund der Definition des Risikos[6] können Varianz und Standardabweichung unmittelbar als Risikomaße interpretiert werden.

---

5 Ein ***Operator*** ist eine bestimmte Rechenvorschrift, die auf das nachfolgende Argument auszuführen ist, im vorliegenden Fall die Erwartungswertbildung.

6 Siehe Kapitel 2, Abschnitt 1.5

Mit der Varianz eng verbunden ist die ***Kovarianz***. Sie misst den linearen Zusammenhang zwischen zwei Zufallsvariablen. Die Kovarianz steht damit für den Risikozusammenhang, also für die Gleichartigkeit der Auswirkungen des Zufalls auf zwei Zufallsvariablen.

> Die ***Kovarianz*** ist definiert als Erwartungswert des Produkts der Abweichungen zweier Zufallsvariablen von ihrem Erwartungswert.

$$\mathrm{Cov}\{\tilde{x};\tilde{y}\} = \sigma_{xy} = \mathrm{E}\{(\tilde{x}-\mu_x)(\tilde{y}-\mu_y)\},$$

wobei
$\mathrm{Cov}\{\cdot;\cdot\}$ Kovarianzoperator
σ_{xy} Kurzschreibweise für die Kovarianz.

Anders als die Varianz kann die Kovarianz auch negative Werte annehmen. Dies ist der Fall, wenn die Ausprägungen der Zufallsvariablen x tendenziell dann hoch sind, wenn die Ausprägungen von y niedrig sind, und umgekehrt. Nimmt die Kovarianz den Wert Null an, besteht kein linearer Zusammenhang zwischen den Zufallsvariablen.

Neben der Kovarianz kann man den ***Korrelationskoeffizienten*** als Maß für den Risikozusammenhang verwenden. Er ist definiert als

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x\sigma_y},$$

wobei
ρ_{xy} Korrelationskoeffizient zwischen x und y
σ_j Kurzschreibweise für die Standardabweichung von j $(j = x, y)$.

Der Korrelationskoeffizient ist auf den Wertebereich $\rho_{xy} \in [-1;+1]$ beschränkt und gibt deshalb einen Einblick nicht nur in die Richtung, sondern auch in die Stärke des linearen Zusammenhangs zwischen x und y.

Ein genauerer Blick auf Kovarianz und Varianz zeigt, dass die Varianz als Spezialfall der Kovarianz angesehen werden kann, nämlich als die Kovarianz einer Zufallsgröße mit sich selber. Dies erkennt man unmittelbar, wenn in der Definition der Kovarianz $\tilde{y}$ durch $\tilde{x}$ ersetzt wird. Dementsprechend gilt:

$$\mathrm{Cov}\{\tilde{x};\tilde{x}\} = \sigma_{xx} = \sigma_{\tilde{x}}^2 = \mathrm{Var}\{\tilde{x}\}.$$

Die Ermittlung von Parametern einer Wahrscheinlichkeitsverteilung und Möglichkeiten zu deren Interpretation seien an einem einfachen Beispiel verdeutlicht, das im Kern der ***Szenariotechnik***[7] entspricht. Es geht um die Absatzmengen verschiedener Produkte in Abhängigkeit von der Entwicklung eines bestimmten, innovativen Marktsegmentes. Die Produkte 1 und 2 sind Produktinnovationen. Das zum Vergleich herangezogene Produkt 3 stellt eine traditionelle Produktvariante dar. Insgesamt lassen in unserem Beispiel die vorhandenen Informationen die folgende Abschätzung zu:

	Entwicklung des innovativen Marktsegments		
Absatzmenge	schlecht (0,2)	normal (0,6)	gut (0,2)
Produkt 1	100	300	400
Produkt 2	240	260	380
Produkt 3	360	240	120

Tabelle 10.1: Absatzmengen und Marktentwicklung.

Für die Verteilungsparameter erhält man (die Berechnung wird an je einem Beispiel verdeutlicht):

$$\mu_1 = 0{,}2\cdot 100 + 0{,}6\cdot 300 + 0{,}2\cdot 400 = 280;$$

$$\mu_2 = 280;\ \ \mu_3 = 240;$$

$$\sigma_1^2 = 0{,}2(100-280)^2 + 0{,}6(300-280)^2 + 0{,}2(400-280)^2 = 9.600;$$

$$\sigma_2^2 = 2.560;\ \ \sigma_3^2 = 5.760;$$

$$\begin{aligned}\sigma_{12} = {} & 0{,}2(100-280)(240-280) \\ & +0{,}6(300-280)(260-280) + 0{,}2(400-280)(380-280) = 3.600;\end{aligned}$$

$$\sigma_{13} = -7.200;\ \ \sigma_{23} = -3.360;$$

$$\rho_{12} = \frac{3.600}{\sqrt{9.600}\sqrt{2.560}} = 0{,}565;$$

$$\rho_{13} = -0{,}968;\ \ \rho_{12} = -0{,}875.$$

Die Parameter zeigen, dass die neuen Produkte einen höheren Absatz erwarten lassen als die traditionelle Variante (μ_3 ist kleiner als μ_1 und μ_2). Bei Produkt 2 streuen die Absatzmengen deutlich weniger stark um ihren Erwartungswert (σ_2^2 ist relativ klein) als bei den anderen Produkten. Zudem erkennt man, dass sich die Absatzmengen der Innovationen 1 und 2 tendenziell ähnlich entwickeln (ρ_{12} ist deutlich größer als 0), das eingeführte Produkt hingegen dann höhere Absatzmengen verspricht, wenn der Markt für Innovationen sich nicht gut entwickelt ($\rho_{13}, \rho_{23} < 0$).

2.4 Rechenregeln für Parameter von Wahrscheinlichkeitsverteilungen

Bei der Lösung von Modellen ist es häufig erforderlich, Rechenregeln für die Verteilungsparameter anzuwenden. Es gilt

$$\mathrm{E}\{a + b\tilde{x}\} = a + b\mu_x;$$

7 Siehe dazu *Brockhoff* (2005), S. 760.

$$E\{a\tilde{x} + b\tilde{y}\} = a\mu_x + b\mu_y;$$

$$E\left\{\sum_{i=1}^{n} a_i\tilde{x}_i\right\} = \sum_{i=1}^{n} a_i\mu_i\,;$$

$$\text{Var}\{\tilde{x}\} = E\{\tilde{x}^2\} - (E\{\tilde{x}\})^2;$$

$$\text{Var}\{a + b\tilde{x}\} = b^2\sigma_x^2;$$

$$\text{Var}\{a\tilde{x} + b\tilde{y}\} = a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\sigma_{xy};$$

$$\text{Var}\left\{\sum_{i=1}^{n} a_i\tilde{x}_i\right\} = \sum_{i=1}^{n} a_i^2\sigma_i^2 + \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} a_i a_j\sigma_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j\sigma_{ij}\,;$$

$$\text{Cov}\{\tilde{x}; a\} = 0;$$

$$\text{Cov}\{\tilde{x}; \tilde{y}\} = E\{\tilde{x}\tilde{y}\} - E\{\tilde{x}\}\,E\{\tilde{y}\};$$

$$\text{Cov}\{(a + b\tilde{x}); (c + d\tilde{y})\} = bd\sigma_{xy};$$

$$\text{Cov}\{(\tilde{x} + \tilde{z}); (\tilde{y} + \tilde{z})\} = \text{Cov}\{(\tilde{x} + \tilde{y}); \tilde{z}\} = \text{Cov}\{\tilde{x}; \tilde{z}\} + \text{Cov}\{\tilde{y}; \tilde{z}\} = \sigma_{xz} + \sigma_{yz};$$

$$\rho\{\tilde{x}; a\} = 0;$$

$$\rho\{(a + b\tilde{x}); (c + d\tilde{y})\} = \rho_{xy}.$$

Neben diesen elementaren Rechenregeln ist die ***Jensensche Ungleichung*** von Interesse. Es gilt

$$E\{f(\tilde{x})\} < f(E\{\tilde{x}\}),$$

wenn $f(E\{\tilde{x}\}) < \infty$ und $f(\cdot)$ eine konkave Funktion ist $(f''(\cdot) < 0)$. Ist $f(\cdot)$ eine konvexe Funktion, kehrt sich das Vorzeichen um.

2.5 Mehrdimensionale Zufallsvariablen

2.5.1 Gemeinsame, bedingte und Randverteilungen

Häufig wirken zwei oder mehrere Zufallswirkungen auf ein unsicheres Ergebnis. Zum Beispiel vereinbart ein Exporteur, dass der Kaufpreis bis spätestens drei Monate nach Lieferung in der Währung des Abnehmers zu begleichen ist (vereinbarter Abnahmepreis: 1.100 \$). Dann hängt der in Euro bewertete Verkaufserlös einerseits davon ab, ob der ausländische Abnehmer seiner Forderung nachkommen kann (denkbar ist, dass aufgrund einer Insolvenz nur ein geringerer Betrag gezahlt wird oder die Zahlung ganz ausfällt), andererseits bestimmt auch die unsichere Wechselkursentwicklung den Erlös in Heimatwährung. Solche Zusammenhänge lassen sich durch mehrdimensionale Zufallsvariablen erfassen.

Die ***gemeinsame Wahrscheinlichkeitsfunktion***

$$w(x_j \cap y_k),$$

wobei

$x_j \cap y_k$ Ereignis, dass zugleich $x = x_j$ und $y = y_k$ eintreten,

gibt die Wahrscheinlichkeit an, dass die Zufallsvariable x die Ausprägung x_j und die Zufallsvariable y die Ausprägung y_k annehmen. Aus der gemeinsamen Wahrscheinlichkeitsfunktion resultiert die ***Randverteilung***, also die isolierte Wahrscheinlichkeitsfunktion für eine einzelne Zufallsvariable. Es gilt

$$w(x_j) = \sum_{k=1}^{n_y} w(x_j \cap y_k),$$

wobei

n_ℓ Anzahl der Ausprägungen der Zufallsvariablen ℓ $(\ell = x, y)$.

Sowohl für die Zahlung in Auslandswährung (hier: US-Dollar) als auch für den Dollar-Wechselkurs gebe es jeweils zwei mögliche Ausprägungen: $x_1 =$ 1.100[\$], $x_2 =$ 500[\$] sowie $y_1 =$ 0,75[€/\$], $y_2 =$ 0,78[€/\$]. Die gemeinsame Wahrscheinlichkeitsfunktion und die Randwahrscheinlichkeitsfunktionen stellt man üblicherweise in einer Matrix dar, woraus sich auch die Bezeichnung Randverteilung ergibt.

	$y = 0{,}78$	$y = 0{,}75$	Summe
$x = 1.100$	0,5	0,45	0,95
$x = 500$	0	0,05	0,05
Summe	0,5	0,5	1

Tabelle 10.2: Gemeinsame Wahrscheinlichkeitsverteilung und Randverteilungen.

Weitere Informationen erhält man, wenn aus der gemeinsamen Wahrscheinlichkeitsverteilung bedingte Wahrscheinlichkeiten ermittelt werden. Die bedingte Wahrscheinlichkeit $w(y_k|x_j)$ gibt die Wahrscheinlichkeit für das Ereignis $\tilde{y} = y_k$ an, sofern der Eintritt des Ereignisses $\tilde{x} = x_j$ als gegeben betrachtet wird. Man erhält bedingte Wahrscheinlichkeiten aus

$$w(y_k|x_j) = \frac{w(y_k \cap x_j)}{w(x_j)}.$$

Konkret gilt im Beispiel unter anderem:

$$w(\tilde{x} = 500|\tilde{y} = 0{,}78) = \frac{0}{0{,}5} = 0$$

oder

$$w(\tilde{y} = 0{,}75|\tilde{x} = 1.100) = \frac{0{,}45}{0{,}95} = 0{,}474.$$

Interpretiert man einen hohen Dollar-Wechselkurs als Zeichen einer guten Wirtschaftsentwicklung in den USA, erkennt man aus der ersten Gleichung, dass bei einer guten Konjunktur der Abnehmer niemals zahlungsunfähig wird.

Die unbedingte Wahrscheinlichkeit für ein Ereignis lässt sich aus den bedingten Wahrscheinlichkeiten rekonstruieren, indem man den Erwartungswert der bedingten Wahrscheinlichkeiten über die Bedingungen bildet (***Satz über die totale Wahrscheinlichkeit***):

$$w(y_k) = \sum_{k=1}^{n_y} w(y_k \cap x_j) = \sum_{j=1}^{n_x} w(x_j)w(y_k|x_j).$$

Dies lässt sich unschwer am Beispiel der Wahrscheinlichkeit für den Wechselkurs von 0,75[€/$] belegen. Es gilt

$$w(\tilde{y} = 0{,}75) = w(\tilde{y} = 0{,}75 \cap \tilde{x} = 1.100) + w(\tilde{y} = 0{,}75 \cap \tilde{x} = 500)$$
$$= 0{,}45 + 0{,}05 = 0{,}5$$

oder, bei der Verwendung der zweiten Gleichung,

$$w(\tilde{y} = 0{,}75) = w(\tilde{x} = 1.100)w(\tilde{y} = 0{,}75|\tilde{x} = 1.100)$$
$$+w(\tilde{x} = 500)w(\tilde{y} = 0{,}75|\tilde{x} = 500) = 0{,}95 \cdot \frac{0{,}45}{0{,}95} + 0{,}05 \cdot 1 = 0{,}5.$$

2.5.2 *Bayes*ianisches Lernen

Aus der Definition von bedingten Wahrscheinlichkeiten und dem Satz über die totale Wahrscheinlichkeit folgt der Satz von *Bayes*, eine besonders wichtige Anwendung der Erkenntnisse über gemeinsame und bedingte Wahrscheinlichkeiten:

$$w(x_j|y_k) = \frac{w(x_j)w(y_k|x_j)}{\sum_{h=1}^{n_x} w(x_h)w(y_k|x_h)}.$$

Der ***Satz von Bayes*** ermöglicht den Schluss von beobachtbaren Variablen (Signalen) auf Ausprägungen einer nicht beobachtbaren Zufallsvariablen. Erforderlich dafür sind Kenntnisse über systematische stochastische Zusammenhänge. Da Unsicherheitssituationen generell durch einen Mangel an Informationen gekennzeichnet sind, ist es wichtig, verfügbare Informationen möglichst gut auszunutzen. Der damit verbundene Vorteil sei an einem konkreten Beispiel verdeutlicht:

Der Manager einer Venture-Capital-Gesellschaft[8] erwägt die Investition in ein innovatives Projekt. Aus langjähriger Erfahrung ist ihm bekannt, dass in der großen Menge der ihm vorgelegten Projektanträge nur ein sehr kleiner Anteil tatsächlich lohnender Projekte enthalten ist. Für das nachfolgende numerische Beispiel möge dieser Anteil 10% betragen.

Dieser Erfahrungswert stellt eine erste Schätzung für die Wahrscheinlichkeit eines Projekterfolgs dar. Es gilt also

$$w(g) = 0{,}1; \;\; w(\bar{g}) = 0{,}9,$$

wobei

g Ereignis erfolgreiches Projekt
$\bar{g}$ Ereignis erfolgloses Projekt (Ereignis „nicht g“).

Die Wahrscheinlichkeiten $w(g)$ und $w(\bar{g})$ bezeichnet man als ***A-priori-Wahrscheinlichkeiten***, weil sie vor Gewinnung zusätzlicher Informationen geschätzt werden.

Ein kennzeichnendes Merkmal des Venture-Capital-Geschäfts besteht in der ***Due-Diligence-Prüfung*** der Projekte. Eine solche Prüfung führt zu einer detaillierten Abschätzung der Erfolgschancen eines konkreten Projekts. Eine solche Prüfung ist natürlich nur sinnvoll, wenn sie einen Einfluss auf die Entscheidung hat. Das bedeutet, nur wenn die Prüfung ein positives Signal ergibt, wird das Projekt tatsächlich finanziert, anderenfalls nicht.

Allerdings ist selbst eine sorgfältige Prüfung nicht perfekt, sondern geht mit kaum vermeidbaren Fehleinschätzungen einher. Zu unterscheiden sind ***Fehler erster Art*** und ***Fehler zweiter Art***. Der Fehler erster Art (auch: „α-Fehler“) besteht darin, ein tatsächlich lohnendes Projekt als nicht lohnend zu klassifizieren. Der Fehler zweiter Art (auch: „β-Fehler“) tritt auf, wenn ein tatsächlich nicht lohnendes Projekt als lohnend klassifiziert wird. Grundsätzlich möchte man gerne beide Fehler vermeiden. Die Gefahr eines Fehlers erster Art lässt sich vermindern, wenn die Abschätzung des Projekterfolgs sehr vorsichtig vorgenommen wird. Genau dies würde aber den Fehler zweiter Art erhöhen. In aller Regel besteht also ein Trade-off zwischen den beiden Fehlerarten. Für die Abwägung im Einzelnen ist von Bedeutung, welcher Fehler die schwerer wiegenden Nachteile mit sich bringt. In unserem Beispiel gehen wir davon aus, dass einer Venture-Capital-Gesellschaft sehr viele Projekte vorgelegt werden, so dass nicht alle tatsächlich guten Projekte finanziert werden könnten. In diesem Fall ist es besonders wichtig, die knappen Mittel nur in gute Projekte zu investieren; dagegen ist es weniger schädlich, eines der guten Projekte als schlecht zu klassifizieren. Daher sollte die Prüftechnik so eingestellt werden, dass der

[8] Vgl. zum Venture Capital Kapitel 7, Abschnitt 4.2.

Fehler erster Art kleiner ist als der Fehler zweiter Art. Für das numerische Beispiel möge gelten $\alpha = 0{,}06$ sowie $\beta = 0{,}12$. Man erhält für die bedingten Wahrscheinlichkeiten der Prüfergebnisse:

$$w(p|g) = 0{,}94;\ w(\bar{p}|g) = \alpha = 0{,}06;$$

$$w(p|\bar{g}) = \beta = 0{,}12;\ w(\bar{p}|\bar{g}) = 0{,}88,$$

wobei

p Ereignis, dass die Prüfung ein positives Signal ergibt

$\bar{p}$ Ereignis des negativen Signals (Ereignis „nicht p").

In Kenntnis dieser Eigenschaften der Due-Diligence-Prüfung können die A-priori-Wahrscheinlichkeiten nunmehr angepasst werden; dies bezeichnet man als ***Bayesianisches Lernen***. Die Anwendung des Satzes von *Bayes* führt im Falle eines positiven Signals zu

$$w(g|p) = \frac{w(g)w(p|g)}{w(g)w(p|g) + w(\bar{g})w(p|\bar{g})} = \frac{0{,}1 \cdot 0{,}94}{0{,}1 \cdot 0{,}94 + 0{,}9 \cdot 0{,}12} = 0{,}4653$$

sowie im Falle eines negativen Signals zu

$$w(g|\bar{p}) = \frac{w(g)w(\bar{p}|g)}{w(g)w(\bar{p}|g) + w(\bar{g})w(\bar{p}|\bar{g})} = \frac{0{,}1 \cdot 0{,}06}{0{,}1 \cdot 0{,}06 + 0{,}9 \cdot 0{,}88} = 0{,}0075.$$

Diese modifizierten Schätzungen für den Projekterfolg bezeichnet man als ***A-posteriori-Wahrscheinlichkeiten***, weil sie nach Heranziehung der zusätzlichen Informationen bestimmt werden.

Im vorliegenden Beispiel zeigt der Kalkül an, dass trotz des a priori sehr hohen Anteils schlechter Projekte nach einem positiven Signal mit einer beträchtlichen Wahrscheinlichkeit davon ausgegangen werden kann, dass es sich um ein lohnendes Projekt handelt. Dagegen ist die Wahrscheinlichkeit für ein gutes Projekt nach einem schlechten Signal nur noch verschwindend gering.

In Fortführung dieses Beispiels könnte man weiter überlegen, ob es sich lohnt, eine zusätzliche, unabhängige Expertise durch einen zweiten Kapitalgeber einzuholen.[9] Erneut wird dies nur dann geschehen, wenn das zusätzliche Signal die Entscheidung beeinflusst. Im modifizierten Szenario wird das Projekt also nur finanziert, wenn beide Kapitalgeber im Rahmen ihrer jeweiligen Due Diligence zu einem positiven Urteil kommen. Wendet der zweite Kapitalgeber eine Prüftechnologie mit denselben Parametern α und β an, führt die Bedingung zweier positiver Signale zu einer A-posteriori-Wahrscheinlichkeit für den Projekterfolg von

[9] Siehe dazu im Einzelnen *Neus/Sturm* (2010). Damit die zusätzliche Prüfung tatsächlich zu einem unabhängigen Ergebnis führt, sollte der zweite Kapitalgeber über ein anderes Spezialwissen verfügen als der erste.

$$w(g|p \cap p) = \frac{w(g)w(p \cap p|g)}{w(g)w(p \cap p|g) + w(\bar{g})\big(1 - w(p \cap p|\bar{g})\big)}$$

$$= \frac{0{,}1 \cdot 0{,}94^2}{0{,}1 \cdot 0{,}94^2 + 0{,}9 \cdot 0{,}12^2} = 0{,}8721.$$[10]

Die nochmals sehr stark angestiegene A-posteriori-Wahrscheinlichkeit bedeutet, dass die Gefahr eines fehlgeschlagenen Projekts deutlich verringert werden kann, wenn es zwei unabhängige Prüfungen gibt. Zugleich erklärt dieses Ergebnis, warum es im Bereich des Venture Capital sehr häufig zur ***Syndizierung***, also zur gemeinsamen Finanzierung von Projekten durch zwei Venture-Capital-Gesellschaften kommt.

3. *Bernoulli*-Prinzip

3.1 Die Konzeption

Das *Bernoulli*-Prinzip geht zurück auf einen Aufsatz aus dem 18. Jahrhundert.[11] Ziel dieses Beitrags war die Entwicklung einer Theorie zur Bewertung unsicherer Einkommensaussichten („Lotterien"). Als eine erste Idee für eine solche Bewertung kann man den erwarteten Gewinn[12] der Lotterie ansehen. Als Repräsentant der mittleren Gewinnchancen scheint der Erwartungswert ein guter Anhaltspunkt für den Wert zu sein. Jedoch zeigt *Bernoulli* unter Heranziehung des „***St. Petersburger Spiel****s*", dass der Erwartungswert ein höchst unplausibler Bewertungsansatz für Gewinnchancen sein kann.

Die Spielregeln sind wie folgt: Die Gewinne werden durch Münzwürfe bestimmt. Fällt bei dem ersten Wurf „Kopf", erfolgt eine Gewinnauszahlung von $g = 1$ und das Spiel ist beendet; bei „Zahl" gibt es keine Auszahlung und das Spiel geht weiter. Fällt beim zweiten Wurf „Kopf", erfolgt eine Gewinnauszahlung von $g = 2$ und das Spiel ist beendet; bei „Zahl" gibt es keine Auszahlung und das Spiel geht weiter, usw. Zusammengefasst gilt: Das Spiel endet beim ersten Eintritt des Ergebnisses „Kopf"; tritt dieses Ereignis im n-ten Wurf ein, beträgt die zugehörige Gewinnauszahlung $g = 2^{n-1}$. Die unsicheren Gewinnaussichten haben insgesamt die Eigenschaft, dass mit relativ hohen Wahrscheinlichkeiten relativ geringe Gewinne erzielt werden, mit sehr kleinen Wahrscheinlichkeiten aber sehr hohe Gewinne. Der Gewinnerwartungswert ist

[10] Man kommt zum selben Ergebnis, wenn die erste A-posteriori-Wahrscheinlichkeit $w(g|p)$ bei nochmaliger Anwendung des Satzes von *Bayes* als neue A-priori-Wahrscheinlichkeit verwendet wird.

[11] *Bernoulli* (1738).

[12] Als Gewinn wird hier die Auszahlung an den Spieler bezeichnet. Damit verbundene Aufwendungen wie beispielsweise der Kaufpreis eines Lotterieloses werden nicht einbezogen.

grundsätzlich in der Lage, dies zu einer Wertziffer zusammenzufassen. Für das St. Petersburger Spiel gilt

$$E\{g\} = \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 4 + \frac{1}{16} \cdot 8 + \dots = \sum_{k=1}^{\infty} \frac{1}{2^k} \cdot 2^{k-1} = \sum_{k=1}^{\infty} \frac{1}{2} = \infty.$$

Im Ergebnis beträgt der Gewinnerwartungswert im St. Petersburger Spiel also ∞. Nun stellt sich die Frage, ob dieses Ergebnis Teil eines sinnvollen Bewertungsansatzes sein kann. Dies wäre dann der Fall, wenn die Zahlungsbereitschaft für die Teilnahme an diesem Spiel dem ermittelten „Wert" entspricht. Tatsächlich ist aber kaum vorstellbar, dass irgendein Spieler einen unendlich hohen Preis für die Teilnahme an dem Spiel bezahlen würde. Dies lässt ernste Zweifel an der generellen Eignung des Erwartungswertes für die Bewertung unsicherer Einkommensaussichten aufkommen.

Laut *Bernoulli* ist dies darauf zurückzuführen, dass die unterschiedlich hohen Gewinne ungleichmäßig wahrgenommen werden. Eine Gewinnsteigerung ausgehend von dem Gewinn 1 im ersten Wurf wird anders bewertet als eine Gewinnsteigerung ausgehend von dem Gewinn 67.108.864 im 27-sten Wurf. Angesichts des riesigen Ausgangsgewinns von über 67 Mio. wäre ein geringerer Nutzenzuwachs zu erwarten. Mit anderen Worten, in die Bewertung von Lotterien sollte der abnehmende Grenznutzen[13] eingehen. *Bernoulli* selbst schlägt die logarithmische Nutzenfunktion $u(g) = \ln g$ vor. Das maßgebliche Bewertungskriterium ist dann nicht der Gewinnerwartungswert sondern der ***Nutzenerwartungswert***. Auf Basis der logarithmischen Nutzenfunktion erhält man einen Nutzenerwartungswert von

$$E\{u(g)\} = \sum_{k=1}^{\infty} \frac{1}{2^k} \cdot \ln 2^{k-1} = \ln 2 \left(\underbrace{\sum_{k=1}^{\infty} \frac{k}{2^k}}_{=2} - \underbrace{\sum_{k=1}^{\infty} \frac{1}{2^k}}_{=1} \right) = \ln 2.$$

Der Nutzenerwartungswert in Höhe von $E\{\ln g\} = \ln 2$ entspricht einer Zahlungsbereitschaft (einem ***„Sicherheitsäquivalent"***[14]) von $s = 2$. Dieser Wert für das skizzierte St. Petersburger Spiel wirkt nicht unplausibel. Vom Ergebnis her vermag der von *Bernoulli* vorgeschlagene Bewertungsansatz demnach eher zu überzeugen als die Heranziehung des Gewinnerwartungswertes. Unbefriedigend bleibt jedoch die alleine aus einer gewissen Plausibilität heraus vorgenommene Wahl der Nutzenfunktion.

Diese Schwäche ist zugleich der Ausgangspunkt für eine axiomatische Herangehensweise an das Problem der Bewertung unsicherer Einkommen. Die

[13] Siehe dazu Kapitel 2, Abschnitt 1.1.3.

[14] Siehe dazu Abschnitt 3.6 dieses Kapitels.

Grundidee lässt sich so beschreiben:[15] Aus bestimmten Annahmen über rationales Handeln in Risikosituationen ergibt sich zwingend eine Nutzenfunktion über die Ergebnisse. Diese Nutzenfunktion wird im englischen Sprachraum als *v. Neumann-Morgenstern*-Nutzenfunktion bezeichnet; im deutschen Sprachraum ist auch der Begriff *Bernoulli*-Nutzenfunktion geläufig. Entscheidungskriterium für einen rationalen Entscheider ist der Erwartungswert des Nutzens, der mit einer Aktion verbunden ist.

Das *Bernoulli*-Prinzip ist das am weitesten akzeptierte Entscheidungsprinzip für Risikosituationen, teilweise wird es sogar synonym mit Rationalverhalten in Risikosituationen verwendet.[16] Inwieweit diese starke Einschätzung zu teilen ist, hängt von der Beurteilung der Verhaltensannahmen ab, die dem *Bernoulli*-Prinzip zugrunde liegen.

Für die Darstellung der Annahmen über rationales Handeln ist es sinnvoll, eine bestimmte Schreibweise für Wahrscheinlichkeitsverteilungen zu verwenden. Die Bezeichnung ***Lotterie***

$$L = \{(x_1, w_1); \dots; (x_n, w_n)\} \text{ mit } \sum_{j=1}^{n} w_j = 1$$

steht für ein Szenario, in dem die Ergebnisse x_j $(j = 1, \dots, n)$ mit den Wahrscheinlichkeiten w_j realisiert werden.

Man spricht von einer ***zusammengesetzten Lotterie***, wenn mindestens eines der Ergebnisse seinerseits wieder eine Lotterie ist. Bei der zusammengesetzten Lotterie

$$L = \{(x_1, w_1); \dots; (x_{n-1}, w_{n-1}); ([x_{n1}, \hat{w}, x_{n2}], w_n)\}$$

werden mit den Wahrscheinlichkeiten w_j die Ergebnisse x_j $(j = 1, \dots, n-1)$ realisiert, mit der Restwahrscheinlichkeit w_n wird eine weitere Lotterie ausgespielt, die mit $\hat{w}$ zum Ergebnis x_{n1} und mit $(1 - \hat{w})$ zu x_{n2} führt.

Lotterien mit nur zwei Ergebnissen nennt man ***Basislotterien***. Diese lassen sich etwas vereinfacht darstellen als

$$L = \{x_1; w_1; x_2\},$$

weil das zweite Ergebnis mit der Wahrscheinlichkeit $1 - w_1$ eintritt.

3.2 Annahmen über rationales Handeln

Nachstehend wird nicht das ursprüngliche Annahmensystem nach *v. Neumann/Morgenstern* vorgestellt, sondern jenes von *Luce/Raiffa*.[17] Es ist zwar etwas

[15] Siehe *v. Neumann/Morgenstern* (1943).

[16] *Bitz* (1981), S. 192.

[17] *Luce/Raiffa* (1957), S. 23 ff.

umfangreicher als das nach *v. Neumann/Morgenstern*, zeichnet sich aber durch eine leichtere Zugänglichkeit aus und verdeutlicht überdies, durch welchen gedanklichen Prozess die Nutzenfunktionen bestimmt werden können. Dabei werden für eine bessere Nachvollziehbarkeit bei der Vorstellung der Annahmen kleine formale Ungenauigkeiten in Kauf genommen. Zum Beispiel wird die Möglichkeit unendlicher Mengen von Ergebnissen nicht einbezogen.

Ordnung der Ergebnisse

Der Entscheider kann alle Ergebnisse in eine Präferenzordnung bringen, die transitiv ist. Für jedes Paar von Ergebnissen x_1 und x_2 gilt

$$x_1 \succ x_2, \quad x_1 \sim x_2 \text{ oder } x_1 \prec x_2,$$

das heißt, entweder wird x_1 dem Ergebnis x_2 vorgezogen ($\succ$), es herrscht Indifferenz zwischen x_1 und x_2 ($\sim$) oder x_1 wird gegenüber x_2 als minderwertig angesehen ($\prec$). Zudem ist die Präferenzordnung transitiv, das heißt, für drei Ergebnisse x_1, x_2 und x_3 gilt

$$x_1 \begin{Bmatrix} \succ \\ \sim \\ \prec \end{Bmatrix} x_2 \text{ und } x_2 \begin{Bmatrix} \succ \\ \sim \\ \prec \end{Bmatrix} x_3 \quad \Rightarrow \quad x_1 \begin{Bmatrix} \succ \\ \sim \\ \prec \end{Bmatrix} x_3.$$

Wird also x_1 dem Ergebnis x_2 vorgezogen und x_2 dem Ergebnis x_3, so wird auch x_1 dem Ergebnis x_3 vorgezogen. Die analogen Aussagen gelten für die schwache Präferenz[18] und für die Indifferenzrelation. Eine vollständige Präferenzordnung über die Ergebnisse enthält folglich ein bestes Ergebnis x_{max} und ein schlechtestes Ergebnis x_{min}.

Stetigkeit

Zu jedem Ergebnis x_j gibt es eine Basislotterie

$$L_j = \{x_{max}; \pi(x_j); x_{min}\}$$

mit der Eigenschaft $x_j \sim L_j$. Der Entscheider ist also in der Lage, zu jedem Ergebnis eine individuell äquivalente Basislotterie mit dem besten bzw. schlechtesten Ergebnis zu benennen. Die Äquivalenz wird durch Wahl der Indifferenzwahrscheinlichkeiten $\pi(x_j)$ oder kurz π_j bewirkt. Aufgrund der Definition von x_{max} und x_{min} gilt für die zugehörigen Indifferenzwahrscheinlichkeiten zwingend $\pi_{max} = 1$ und $\pi_{min} = 0$.

18 Die schwache Präferenz schließt die Indifferenz ein, ist also entsprechend dem mathematischen „$\geq$" konstruiert.

Substituierbarkeit

In einer Lotterie L kann jedes Ergebnis x_j durch die äquivalente Lotterie L_j ersetzt werden. Gibt es eine einfache Lotterie

$$L = \{(x_1; w_1); \ldots; (x_n; w_n)\}$$

und eine zusammengesetzte Lotterie

$$\begin{aligned} L^* &= \{(L_1; w_1); \ldots; (L_n; w_n)\} \\ &= \{([x_{max}; \pi_1; x_{min}]; w_1); \ldots; ([x_{max}; \pi_n; x_{min}]; w_n)\}, \end{aligned}$$

dann gilt $L \sim L^*$. Die Lotterie L^* ist eine zusammengesetzte Lotterie, in der letztlich nur die Ergebnisse x_{min} oder x_{max} eintreten können.

Reduktion zusammengesetzter Lotterien

Jede zusammengesetzte Lotterie ist gleichwertig zu einer Lotterie mit denselben Einzelergebnissen und Wahrscheinlichkeiten, die nach den Regeln der Wahrscheinlichkeitsrechnung ermittelt werden. Gibt es eine zusammengesetzte Lotterie

$$L^* = \{([x_{max}; \pi_1; x_{min}]; w_1); \ldots; ([x_{max}; \pi_n; x_{min}]; w_n)\}$$

und eine ***reduzierte Lotterie***

$$L^{**} = \left\{ x_{max}; \sum_{j=1}^{n} \pi_j w_j; x_{min} \right\},$$

dann gilt $L^* \sim L^{**}$.

Monotonie

Eine Basislotterie wird einer anderen Basislotterie mit denselben Ergebnissen genau dann vorgezogen, wenn das präferierte Ergebnis mit der höheren Wahrscheinlichkeit erzielt wird. Bei

$$x_1 \succ x_2$$

gilt

$$\{x_1; w; x_2\} \succ \{x_1; \hat{w}; x_2\}$$

genau dann, wenn $w > \hat{w}$. Die entsprechende Aussage gilt für die schwache Präferenz.

Transitivität der Präferenz zwischen Lotterien

Die Präferenz zwischen Lotterien ist transitiv. Es gilt also

$$L \begin{Bmatrix} \succ \\ \sim \\ \prec \end{Bmatrix} L^* \text{ und } L^* \begin{Bmatrix} \succ \\ \sim \\ \prec \end{Bmatrix} L^{**} \quad \Rightarrow \quad L \begin{Bmatrix} \succ \\ \sim \\ \prec \end{Bmatrix} L^{**}.$$

Insbesondere gilt weiter

$$L_1^{**} \succ L_2^{**} \quad \Rightarrow \quad L_1 \succ L_2.$$

Ableitung der Entscheidungsvorschrift

Gibt es also zwei Lotterien L_1 und L_2, dann sind diese wegen der Stetigkeit und der Substituierbarkeit äquivalent zu L_1^* bzw. L_2^* und weiter wegen der Reduzierbarkeit äquivalent zu L_1^{**} bzw. L_2^{**}. Infolge der Monotonie gilt

$$L_1^{**} \succ L_2^{**} \quad \Leftrightarrow \quad \sum_{j=1}^{n} \pi_{1j} w_j > \sum_{j=1}^{n} \pi_{2j} w_j.$$

Schließlich folgt wegen der Transitivität der Präferenz zwischen Lotterien

$$L_1 \succ L_2 \quad \Leftrightarrow \quad \sum_{j=1}^{n} \pi_{1j} w_j > \sum_{j=1}^{n} \pi_{2j} w_j.$$

Aus den Verhaltensannahmen des *Bernoulli*-Prinzips folgt demnach zwingend die Entscheidungsvorschrift „Maximiere den Erwartungswert der Indifferenzwahrscheinlichkeiten". Da die individuelle Bewertung genau durch die Zuweisung der Indifferenzwahrscheinlichkeiten π_j vorgenommen wird, lassen sich diese Wahrscheinlichkeiten als Nutzen u der Ergebnisse interpretieren:

$$u(x_j) = \pi_j.$$

3.3 *Bernoulli*-Befragung

3.3.1 Der Entscheidungsprozess

Ausgangspunkt einer Entscheidung nach dem *Bernoulli*-Prinzip ist das Grundmodell der Entscheidungstheorie: Die Auswahl einer bestimmten Aktion a_i führt zu einer Lotterie, bei der je nach eintretendem Zustand eines der Ergebnisse x_{ij} realisiert wird. Durch die „*Bernoulli*-Befragung"[19] kann man den Entscheidungsprozess klären:

1. Zunächst werden nach der Ordnungsannahme alle Ergebnisse x_{ij} in eine Präferenz-Reihenfolge gebracht.

[19] Vgl. bspw. *Laux/Gillenkirch/Schenk-Mathes* (2014), S. 117 ff.

2. Anschließend werden nach der Stetigkeitsannahme für alle Ergebnisse die zugehörigen Indifferenzwahrscheinlichkeiten $\pi(x_{ij}) = \pi_{ij}$ ermittelt.

3. Nach der Substituierbarkeitsannahme können alle Ergebnisse x_{ij} durch die äquivalenten Basislotterien $L_{ij} = \{x_{max}; \pi_{ij}; x_{min}\}$ ersetzt werden, ohne dass sich die Präferenz über die ursprünglichen, komplexen Lotterien verändert. Die modifizierten, zusammengesetzten Lotterien führen nur noch zu den Ergebnissen x_{max} oder x_{min}.

4. Nach der Reduktionsannahme können die zusammengesetzten Lotterien reduziert werden; dies führt bei Alternative a_i zu

$$L_i = \left\{ x_{max}; \sum_{j=1}^{n} \pi_{ij} w_j ; x_{min} \right\},$$

da die Indifferenzwahrscheinlichkeiten π_{ij} unabhängig von den Umweltzuständen sind.

5. Nach der Monotonieannahme wird diejenige Basislotterie vorgezogen, bei der das bessere Ergebnis mit der größeren Wahrscheinlichkeit erzielt wird.

6. Da infolge der Transitivitätsannahme die Präferenzordnung über Lotterien transitiv ist, wird die Präferenz über Alternativen a_i durch den Ausdruck

$$\sum_{j=1}^{n} \pi_{ij} w_j = \mathrm{E}\{\pi_{ij}\}$$

vollständig erfasst. Es ist diejenige Alternative auszuwählen, die den Erwartungswert der Indifferenzwahrscheinlichkeiten, oder kürzer: den Nutzenerwartungswert maximiert. Die Entscheidungsvorschrift des *Bernoulli*-Prinzips lautet also:

> Wähle diejenige Handlungsalternative, die den maximalen ***Erwartungswert des Nutzens*** der Ergebnisse herbeiführt.

3.3.2 Ein Beispiel

Ein Beispiel mag den Sachverhalt zusätzlich verdeutlichen. Ausgangspunkt sei die folgende Ergebnismatrix:

	z_1 (0,25)	z_2 (0,5)	z_3 (0,25)
a_1	120	120	160
a_2	100	160	100

Tabelle 10.3: Ergebnismatrix.

Der erste Schritt besteht darin, die Präferenzordnung über die Ergebnisse festzustellen. In diesem Fall ist das leicht, weil es sich um eindimensionale Größen (zum Beispiel um Vermögen) handelt. Es gilt $160 \succ 120 \succ 100$. Im zweiten Schritt werden Lotterien betrachtet, die entweder zum besten ($x_{max} = 160$) oder zum schlechtesten Ergebnis ($x_{min} = 100$) führen. Bei dem schlechtesten bzw. besten Ergebnis müssen die Indifferenzwahrscheinlichkeiten per Konstruktion und somit unabhängig von individuellen Präferenzen $\pi(x_{min}) = 0$ bzw. $\pi(x_{max}) = 1$ betragen. Für das dritte Ergebnis $x = 120$ muss der Entscheider die Indifferenzwahrscheinlichkeit $\pi(120)$ bestimmen mit der Eigenschaft $120 \sim \{160; \pi(120); 100\}$. Einzig an dieser Stelle gehen individuelle Präferenzen in den Entscheidungsprozess ein. Die Befragung eines Entscheiders ergebe die folgende Indifferenzwahrscheinlichkeit:

x	100	120	160
$\pi(x)$	0	0,5	1

Tabelle 10.4: Indifferenzwahrscheinlichkeiten.

Aufgrund des Substitutionsprinzips können die einzelnen Ergebnisse in der Ergebnismatrix durch die zu ihnen äquivalenten Lotterien ersetzt werden. Demnach sind alle Alternativen äquivalent zu zusammengesetzten Lotterien, bei denen mit unterschiedlichen Wahrscheinlichkeiten nur die Ergebnisse 160 oder 100 erzielt werden. Die Alternativen unterscheiden sich durch die Gesamtwahrscheinlichkeit für das beste Ergebnis oder, anders ausgedrückt, durch die Erwartungswerte der Indifferenzwahrscheinlichkeiten. Diese lassen sich wie folgt ermitteln:

$$\mathrm{E}\{\pi|a_1\} = 0{,}75\pi(120) + 0{,}25\pi(160) = 0{,}625,$$

$$\mathrm{E}\{\pi|a_2\} = 0{,}5\pi(100) + 0{,}5\pi(160) = 0{,}5.$$

Die Aktion 1 hat also den höheren Erwartungswert der Indifferenzwahrscheinlichkeiten (des Nutzens) und wird demnach ihrer Alternative vorgezogen. Um den Bezug zu *Bernoullis* Ausgangspunkt herzustellen, sei darauf hingewiesen, dass die Aktionen a_1 und a_2 gleichermaßen zu einem Ergebniserwartungswert von $\mathrm{E}\{x|a_1\} = \mathrm{E}\{x|a_2\} = 130$ führen. Bei Bewertung alleine auf Basis des Erwartungswerts wären die Aktionen also gleichwertig. Die Maximierung des Nutzenerwartungswertes führt offenbar im Allgemeinen zu einer anderen Entscheidung als die Maximierung des Ergebniserwartungswertes.

3.4 Normierung der Nutzenwerte und positive Lineartransformation

Die *Bernoulli*-Nutzenfunktionen sind zunächst auf Nutzenwerte im Intervall zwischen 0 (bei x_{min}) und 1 (bei x_{max}) normiert. Da es sich beim *Bernoulli*-

Prinzip jedoch nicht um eine Methode zur Nutzenmessung handelt, sondern zur ***Auswahl aus Handlungsalternativen***, spielt das überhaupt keine Rolle.

Der Erwartungswert ist ein linearer Operator $(\mathrm{E}\{a + bx\} = a + b\mathrm{E}\{x\})$. Deshalb führt eine positive Lineartransformation der Nutzengrößen nicht zu einer veränderten Reihenfolge der Alternativenbewertung. Die Maximierung von $\mathrm{E}\{a + b\pi_{ij}\}$ führt bei $b > 0$ (daher die Einschränkung auf *positive* Lineartransformationen) stets zu derselben Entscheidung wie die Maximierung von $\mathrm{E}\{\pi_{ij}\}$. Nutzenfunktionen, die durch eine positive Lineartransformation auseinander hervorgehen, sind im Sinne des *Bernoulli*-Prinzips identisch.

Betrachtet man zum Beispiel die allgemeine quadratische Nutzenfunktion

$$u(x) = a + bx + cx^2,$$

so lässt diese sich im Falle von $b > 0$ transformieren zu

$$v(x) = \frac{u(x) - a}{b} = x + \hat{c}x^2,$$

mit $\hat{c} = c/b$. Offenbar ist die Nutzenfunktion $v(x)$ leichter handhabbar als $u(x)$, führt aber stets zu derselben Entscheidung wie die Funktion $u(x)$. Daher bietet es sich an, aus einer Klasse äquivalenter Nutzenfunktionen stets die einfachste auszuwählen.

3.5 Kritik an den Verhaltensannahmen

Ernstzunehmende Kritik am *Bernoulli*-Prinzip muss an den Rationalitätsannahmen ansetzen. Empirische Untersuchungen kommen häufig zu dem Ergebnis, dass es Regelmäßigkeiten im Entscheidungsverhalten von Individuen gibt, die mit bestimmten Annahmen, insbesondere der ***Substitutionsannahme***, nicht vereinbar sind. Experimente, die ein solches Verhalten ausweisen, sind allerdings häufig in einer besonders unübersichtlichen Art und Weise konstruiert. Dies wird hier anhand eines bekannten Beispiels, dem ***Allais-Paradoxon***, verdeutlicht:

Allais ließ in den 50er Jahren Teilnehmer an Experimenten aus zwei Lotteriepaaren (a_1, a_2) und (b_1, b_2) jeweils eine Lotterie auswählen. Konkret gilt[20]

$$a_1 = \{100\ \text{Mio.}; 1; 0\}, \qquad a_2 = \{500\ \text{Mio.}; 0{,}98; 0\},$$

$$b_1 = \{500\ \text{Mio.}; 0{,}0098; 0\}, \qquad b_2 = \{100\ \text{Mio.}; 0{,}01; 0\}.$$

Die Geldbeträge lauten bei *Allais* auf alte französische Francs und sind damit ohne Inflationsanpassung ungefähr 650-mal höher als aktuelle Euro-Beträge. Die Entscheider äußerten überwiegend die Präferenzen $a_1 \succ a_2$ und $b_1 \succ b_2$.

[20] *Allais* (1979), S. 91.

Leicht lässt sich zeigen, dass dieses Entscheidungsverhalten mit dem *Bernoulli*-Prinzip nicht vereinbar ist. Denn $a_1 \succ a_2$ impliziert

$$1 \cdot u(100) + 0 \cdot u(0) > 0{,}98 \cdot u(500) + 0{,}02 \cdot u(0) \Leftrightarrow$$
$$u(100) > 0{,}98 \cdot u(500) + 0{,}02 \cdot u(0).$$

Die zweite Auswahl, $b_1 \succ b_2$, erfordert

$$0{,}01 \cdot u(100) + 0{,}99 \cdot u(0) < 0{,}0098 \cdot u(500) + 0{,}9902 \cdot u(0) \Leftrightarrow$$
$$0{,}01 \cdot u(100) < 0{,}0098 \cdot u(500) + 0{,}0002 \cdot u(0) \Leftrightarrow$$
$$u(100) < 0{,}98 \cdot u(500) + 0{,}02 \cdot u(0).$$

Offensichtlich ist es also im Sinne des *Bernoulli*-Prinzips widersprüchlich, zugleich a_1 und b_1 auszuwählen, und dies völlig ***unabhängig von der Nutzenfunktion*** des Entscheiders. Genau dies macht das *Allais*-Paradoxon aus. Die Intuition hinter den Entscheidungen ist dabei durchaus nachvollziehbar. Aus den a-Lotterien wird a_1 offenbar deshalb vorgezogen, weil der Betrag von 100 Mio. sicher realisiert wird. Angesichts des ohne Verlustmöglichkeit erzielbaren, sehr hohen Betrages von 100 Mio. fällt die Steigerung auf 500 Mio. weniger ins Gewicht als die zweiprozentige Möglichkeit eines Totalverlustes. Bei den b-Lotterien sieht das anders aus. Offensichtlich werden hier die 100 Mio. und die 500 Mio. als fühlbar andere Beträge erkannt, die Gewinnwahrscheinlichkeiten von 0,98% und 1% werden jedoch als fast identisch angesehen. Auch in anderen Experimenten lässt sich ein Entscheidungsverhalten nachweisen, das einerseits die Bedeutung geringer Wahrscheinlichkeiten oder geringer Wahrscheinlichkeitsunterschiede unterschätzt und das andererseits bei extrem hohen Gewinnen und Verlusten nicht hinreichend differenziert, jeweils verglichen mit den Anforderungen des *Bernoulli*-Prinzips.

Die ***Kritiker*** des *Bernoulli*-Prinzips ziehen daraus die Folgerung, dass Individuen häufig in einer mit dem *Bernoulli*-Prinzip nicht vereinbaren Weise entscheiden. Das *Bernoulli*-Prinzip wäre demnach als deskriptive Theorie überhaupt nicht geeignet und sollte auch nicht als Basis für eine normative Theorie verwendet werden. Wenn die Verhaltensannahmen als plausibel bezeichnet und akzeptiert werden, liegt das den Kritikern zufolge nur daran, dass die Implikationen der Annahmen nicht erkannt werden. Die einschlägigen Beispiele dienen dazu, kritische Implikationen aufzudecken.

Die ***Befürworter*** des *Bernoulli*-Prinzips argumentieren natürlich anders: Sie werfen den Designern der Experimente vor, Beispiele so zu entwickeln, dass Fehlentscheidungen wahrscheinlich sind. In „normalen" Entscheidungssituationen kommt es nicht zu derartigen Fehlern. Überdies erweist sich die Stärke des *Bernoulli*-Prinzips genau bei solchen Entscheidungen, die zunächst schwer zu durchschauen sind. Das *Bernoulli*-Prinzip führt zur Vermeidung von Fehlern, indem die Entscheidungssituation besser strukturiert wird. Die Fehler in verwickelten Situationen wären demnach kein Argument gegen das *Bernoulli*-Prinzip, sondern belegen gerade dessen Stärke. Die Argumente lassen sich anhand einer recht bekannten Grafik verdeutlichen:[21]

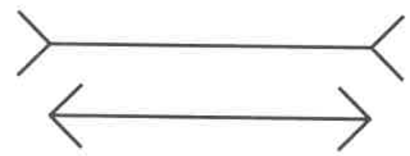

Abbildung 10.4: Optische Täuschung als Beispiel für falsche Entscheidungen.

Gefragt ist, welche der beiden horizontalen Linien länger ist. Die spontane Antwort eines jeden Betrachters wird lauten: die obere Linie. Tatsächlich sind die beiden Linien exakt gleich lang, die optische Fehleinschätzung wird durch die umgekehrte Anordnung der Pfeilspitzen hervorgerufen. Der erste Eindruck trügt also, das Messinstrument eines Zentimetermaßes legt die richtige Antwort offen.

Spontane, intuitiv geprägte Entscheidungen stehen, wie etwa das *Allais*-Paradoxon belegt, häufig im Widerspruch zu den Implikationen des *Bernoulli*-Prinzips. Muss man deshalb aber folgern, das *Bernoulli*-Prinzip sei empirisch falsch? Oder handelt es sich – wie analog das Zentimetermaß – bei dem *Bernoulli*-Prinzip um ein Analyseinstrument, mit dessen Hilfe normativ richtige Entscheidungen in komplizierten Risikosituationen getroffen werden können? Zumindest für das Gebiet normativer Entscheidungstheorien ist zu konstatieren, dass es trotz der genannten Befunde keine weithin akzeptierte Alternative zum *Bernoulli*-Prinzip gibt.[22]

Mit Blick auf deskriptive Entscheidungstheorien kommt man zu einer anderen Antwort. Denn zum einen gibt es neben dem dargestellten *Allais*-Paradoxon zahlreiche weitere empirische Regelmäßigkeiten, welche der Gültigkeit des *Bernoulli*-Prinzips im Besonderen und dem Rationalverhalten im Allgemeinen widersprechen.[23] Zum anderen gibt es deskriptive Entscheidungstheorien wie beispielsweise die ***Prospect-Theorie***[24], welche einige der Anomalien besser zu erfassen erlauben. Und schließlich gilt diese Beobachtung nicht nur für Ansätze mit nur einem Entscheider, sondern auch für spieltheoretische Modelle. Hier

21 Vgl. bspw. *Kahneman* (2012), S. 41.
22 Ähnlich *Bamberg/Trost* (1996), S. 659 f.
23 Eine umfassende Abhandlung zu den Regelmäßigkeiten begrenzt rationalen Entscheidungsverhaltens präsentiert *Kahneman* (2012).
24 Oder auch: Neue Erwartungstheorie; *ebenda*, S. 342 ff.

erfährt beispielsweise die Verwendung von Nutzenfunktionen mit ***Ungleichheitsaversion*** eine zunehmende Verbreitung.[25] Auch dabei dient die Loslösung vom *Bernoulli*-Prinzip dazu, reichhaltigere und offenbar empirisch besser zutreffende Ergebnisse herbeizuführen.

Kahnemans Theorie vom schnellen (intuitiven) und langsamen (rationalen) Denken erklärt in beeindruckender Eleganz und Schlüssigkeit, wie sich das Nebeneinander von spontanen Handlungen und komplexen Abwägungen bei der menschlichen Entwicklung bewährt hat. Dennoch sind die vielen nachgewiesenen „Effekte" nicht geeignet, die Eignung der *Bernoulli*-Prinzips als normative Theorie, gewissermaßen als Instrument des langsamen Denkens, zu widerlegen.

3.6 Nutzenfunktionen und Risikoeinstellungen

Im Weiteren werden nur noch *Bernoulli*-Nutzenfunktionen betrachtet. Die Auswahl zwischen Alternativen hängt offenbar vom Verlauf der Nutzenfunktion ab. Es wird unterstellt, dass ein größeres Ergebnis einem geringeren Ergebnis vorgezogen wird. Bei der überwiegend unterstellten Ergebnisdimension des Vermögens bedarf das keiner näheren Erläuterung. Bei anderen Ergebnisdimensionen (zum Beispiel Arbeitszeit) behilft man sich in der Regel damit, dass man sie in Form ihres monetären Äquivalents einbezieht.

Die positive Steigung allein charakterisiert die Nutzenfunktion jedoch nur unzureichend. Insbesondere lassen sich damit keine Aussagen über Risikopräferenzen treffen. Bisher wurde häufig betont, dass Entscheider risikoavers sind. Weniger riskante Alternativen werden also ceteris paribus vorgezogen. Das *Bernoulli*-Prinzip erlaubt eine genaue Definition der Risikoaversion.

> Bei ***Risikoaversion*** wird ein sicherer Betrag in Höhe des Ergebniserwartungswertes der unsicheren Ergebnisverteilung vorgezogen. Es gilt also $u(\mathrm{E}\{x\}) > \mathrm{E}\{u(x)\}$.

Die Interpretation, dass ein risikoaverser Entscheider jedes Risiko vermeidet, wäre ***völlig verfehlt***. Die generelle Risikovermeidung trifft nur dann zu, wenn alle anderen Merkmale der Ergebnisverteilung (beispielsweise der Erwartungswert) übereinstimmen. Zwar ist Risiko ein Ungut, für dessen Übernahme ein Preis, nämlich eine Risikoprämie, gefordert wird. Bei Abgeltung durch eine hinreichend große Risikoprämie wird jedoch grundsätzlich jedes Risiko übernommen. Umgekehrt gibt es zu jeder positiven Risikoprämie eine adäquate Risikomenge, die ein Entscheider zu übernehmen bereit ist. Ein risikoaverser Entscheider „always takes some part of a favorable gamble"[26]. Dabei ist eine Lotterie „favorable", wenn sie einen positiven Gewinnerwartungswert hat.

[25] Vgl. bspw. *Fehr/Schmidt* (1999); im Falle der Ungleichheitsaversion wird eine ungleiche Ergebnisverteilung als nutzenmindernd empfunden.

Analog zur Risikoaversion (oder Risikoscheu) spricht man von ***Risikofreude***, wenn in obiger Ungleichung das umgekehrte Vorzeichen gilt. Stimmen der Nutzen des Erwartungswerts und der Erwartungswert des Nutzens überein, spricht man von ***Risikoindifferenz*** oder Risikoneutralität. Im Weiteren wird durchweg die Risikoaversion untersucht, wobei als Grenzfall manchmal die Risikoindifferenz einbezogen wird.

Risikoaversion impliziert, dass positive Abweichungen von einem bestimmten Ergebnis weniger stark ins Gewicht fallen als gleich hohe negative Abweichungen. Der Verlauf der Nutzenfunktion ist also ***oberhalb eines bestimmten Ergebnisses flacher*** als unterhalb. Eine hierzu passende Redensart lautet: „Geld zu haben ist weniger gut, als keines zu haben schlecht ist." Zu derselben Aussage kommt man mit Hilfe der oben eingeführten *Jensen*schen Ungleichung[27]: Danach ist die Definitionsgleichung für Risikoaversion genau im Falle ***konkaver Funktionen*** $u(\cdot)$ erfüllt, also bei Funktionen mit abnehmender Steigung. Eine solche Nutzenfunktion ist demnach Ausdruck der Risikoaversion.

Mit der Definition der Risikoaversion eng verbunden ist das sogenannte Sicherheitsäquivalent einer Lotterie.

> Das ***Sicherheitsäquivalent*** einer Lotterie ist der sichere Ergebniswert, welcher den Entscheider indifferent lässt zwischen ebendiesem Ergebniswert und der unsicheren Lotterie.

Es gilt also

$$u(s) = \mathrm{E}\{u(x)\} \quad \Leftrightarrow \quad s = u^{-1}(\mathrm{E}\{u(x)\}),$$

wobei
s Sicherheitsäquivalent
u^{-1} Umkehrfunktion der Nutzenfunktion.

Bei Risikoaversion ist das Sicherheitsäquivalent stets kleiner als der Erwartungswert des unsicheren Ergebniswertes. Dies ergibt sich zum einen intuitiv daraus, dass ein risikoaverser Entscheider ein risikoloses Ergebnis auch dann einer unsicheren Lotterie gleich schätzt, wenn das mittlere Niveau des unsicheren Ergebnisses etwas höher ist. Die erwartete Ergebnissteigerung ausgehend vom Sicherheitsäquivalent muss größer sein als die erwartete Ergebnisminderung, damit beides einander ausgleicht. Zum anderen folgt dies wiederum aus

[26] *Arrow* (1970), S. 100.
[27] Siehe Abschnitt 2.4 dieses Kapitels.

den formalen Eigenschaften einer konkaven Nutzenfunktion und der *Jensen*schen Ungleichung. Den gleichen Informationsgehalt wie das Sicherheitsäquivalent hat die Risikoprämie.

> Die ***Risikoprämie*** λ ist definiert als Differenz zwischen Erwartungswert und Sicherheitsäquivalent.

Im Falle der Risikoaversion ist die Risikoprämie positiv:

$$\lambda = \mathrm{E}\{x\} - s = \mathrm{E}\{x\} - u^{-1}(\mathrm{E}\{u(x)\}) > 0,$$

wobei
λ Risikoprämie.

Anhand der konkaven Funktion $u(x) = \sqrt{x}$ lässt sich das leicht verdeutlichen. Zu beurteilen ist eine Basislotterie $L = \{1; 0{,}5; 49\}$. Der Erwartungswert des Ergebnisses beträgt $\mathrm{E}\{x\} = 25$, der Nutzen eines sicheren Ergebnisses in Höhe dieses Erwartungswerts $u(\mathrm{E}\{x\}) = 5$. Für den Nutzenerwartungswert gilt $\mathrm{E}\{u(x)\} = 0{,}5 \cdot 1 + 0{,}5 \cdot 7 = 4$. Das Sicherheitsäquivalent ergibt sich aus $\sqrt{s} = 4$ mit $s = 16$, die Risikoprämie als $\lambda = \mathrm{E}\{x\} - s = 9$. Grafisch dargestellt erhält man:

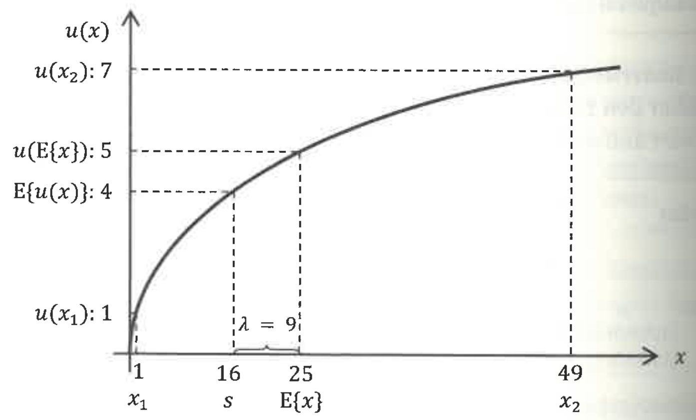

Abbildung 10.5:
Erwartungswert, Sicherheitsäquivalent und Risikoprämie bei Risikoaversion.

3.7 Maßgrößen für die Risikoaversion

Da sich Risikoaversion in der Rechts-Krümmung (Konkavität) der Nutzenfunktion äußert, könnte es angemessen erscheinen, den Grad der Krümmung, ausgedrückt durch den Betrag der zweiten Ableitung der Nutzenfunktion, als Maßstab für die Risikoaversion zu verwenden. Dies erweist sich jedoch als nicht

zweckmäßig, wie man angesichts der Eigenschaft der Gleichwertigkeit positiv lineartransformierter Nutzenfunktionen leicht erkennt. Eine derartige Transformation kann nämlich die zweite Ableitung der Funktion verändern, obwohl stets identische Entscheidungen getroffen werden. Diesen Effekt kann man durch Einbeziehung der ersten Ableitung der Funktion eliminieren.

Die ***absolute Risikoaversion*** ist definiert als

$$r_a(x) = -\frac{u''(x)}{u'(x)},$$

wobei
$r_a(x)$ absolute Risikoaversion.

Man bezeichnet sie auch als ***Pratt-Arrow-Maß***.[28] Neben der absoluten Risikoaversion verwendet man auch die Kennzahlen ***relative Risikoaversion*** und (absolute) ***Risikotoleranz***. Sie sind definiert als

$$r_r(x) = x r_a(x) = -x\frac{u''(x)}{u'(x)},$$

$$r_t(x) = \frac{1}{r_a(x)} = -\frac{u'(x)}{u''(x)},$$

wobei
$r_r(x)$ relative Risikoaversion
$r_t(x)$ Risikotoleranz.

Anhand der Eigenschaften dieser Kennzahlen lassen sich Klassen von Nutzenfunktionen bilden, die typische Implikationen haben. Die relative Risikoaversion lässt sich von der absoluten Risikoaversion anhand des folgenden Entscheidungsproblems abgrenzen: Ein Kapitalanleger hat die Möglichkeit, einen Geldbetrag v auf eine riskante und eine risikolose Anlagemöglichkeit aufzuteilen. Für sein unsicheres Endvermögen gilt dann

$$\begin{aligned}\tilde{w} &= (1+\tilde{r})x + (1+i)(v-x) \\ &= (1+i)v + (\tilde{r}-i)x,\end{aligned}$$

wobei
$\tilde{w}$ unsicheres Endvermögen
$\tilde{r}$ unsichere Rendite der riskanten Investition
i Verzinsung der risikolosen Anlagemöglichkeit
v Anfangsvermögen
x unsicher investierter Geldbetrag.

Die Entscheidung über x hängt neben der Risikoeinstellung ab von der Rendite der risikolosen Anlage, von der erwarteten Rendite der riskanten Anlage

[28] Nach *Pratt* (1964) und *Arrow* (1970).

und vom damit verbundenen Risiko. Hat der Entscheider eine konstante absolute Risikoaversion, verändert sich der optimale riskant angelegte Betrag x^* bei einer Veränderung des Anfangsvermögens v nicht. Nimmt hingegen die absolute Risikoaversion ab (bzw. zu), steigt (bzw. verringert sich) der riskant angelegte ***Betrag*** und damit das absolut übernommene Risiko mit wachsendem Vermögen. Es gilt also

$$\begin{Bmatrix}\text{abnehmende}\\ \text{konstante}\\ \text{zunehmende}\end{Bmatrix} \text{ absolute Risikoaversion } \Leftrightarrow \frac{\partial x^*}{\partial v} \begin{Bmatrix}>\\ =\\ <\end{Bmatrix} 0.$$

Für die relative Risikoaversion gelten entsprechende Aussagen, nun jedoch bezogen auf den riskant angelegten ***Anteil am Anfangsvermögen***, definiert als $y = x/v$, also das relativ übernommene Risiko. Man erhält

$$\begin{Bmatrix}\text{abnehmende}\\ \text{konstante}\\ \text{zunehmende}\end{Bmatrix} \text{ relative Risikoaversion } \Leftrightarrow \frac{\partial y^*}{\partial v} \begin{Bmatrix}>\\ =\\ <\end{Bmatrix} 0,$$

wobei
y unsicher investierter Anteil des Anfangsvermögens v.

Die Abhängigkeit von absoluter und relativer Risikoaversion vom Anfangsvermögen ermöglicht Aussagen über die ***Vermögenswirkungen auf Entscheidungen***.

Die Risikotoleranz erweist sich im Zusammenhang mit Fragen der Risikoaufteilung als besonders anschaulich, denn eine ***optimale Risikoaufteilung*** zwischen verschiedenen Individuen ist dann gegeben, wenn jeder den Risikoanteil übernimmt, der dem Anteil seiner Risikotoleranz an der Summe aller Risikotoleranzen entspricht.[29]

Häufig verwendete Nutzenfunktionen lassen sich wie folgt einordnen:

Nutzenfunktion	absolute Risikoaversion	relative Risikoaversion
Lineare Funktion $u(x) = x$	konstant (null)	konstant (null)
Logarithmische Funktion $u(x) = \ln x$	abnehmend	konstant
Potenzfunktion $u(x) = x^\gamma$ $(0 < \gamma < 1)$	abnehmend	konstant
Exponentielle Funktion $u(x) = 1 - \exp(-\gamma x)$	konstant	zunehmend
Quadratische Funktion $u(x) = x - \gamma x^2$	zunehmend	zunehmend

Tabelle 10.5: Nutzenfunktionen und Risikoaversionsmaße.

[29] Vgl. Kapitel 5, Abschnitt 1.4.5.

Empirische Untersuchungen kommen überwiegend zu dem Ergebnis, dass die absolute Risikoaversion abnehmend ist und die relative Risikoaversion einen konstanten Wert aufweist.[30] Die logarithmische und die Potenzfunktion weisen demnach besonders plausible Eigenschaften auf. Dennoch werden in Entscheidungsmodellen häufig andere Nutzenfunktionen unterstellt; der damit verbundene Vorteil der analytischen Vereinfachung überwiegt den Nachteil problematischer empirischer Implikationen. Dies gilt insbesondere bei solchen Fragestellungen, bei denen es auf den Grad der Risikoaversion ankommt, nicht aber auf die Vermögenswirkungen auf Entscheidungen.

4. (μ, σ)-Prinzip

4.1 Idee der Vereinfachung

Das *Bernoulli*-Prinzip hat offenbar spezifische Vorzüge, insbesondere die ausgeprägte Orientierung an rationalem Verhalten und das Vermeiden von Verstößen gegen das Rationalitätspostulat. Insbesondere in komplizierten Entscheidungssituationen kann jedoch die analytische Handhabung schwierig sein. Das Vermögen als das Argument der Nutzenfunktion kann in einer mehr oder minder komplizierten Weise von den Entscheidungsvariablen abhängen. Es ist leicht einzusehen, dass allein die Explikation der notwendigen Bedingung für das Optimum erhebliche Probleme aufwerfen kann. Und selbst wenn die Lösung eines einzelnen Entscheidungsproblems möglich ist, wird der Vergleich verschiedener Entscheidungssituationen erschwert, wenn nur implizite Lösungen ermittelt werden können. Gerade der Vergleich von Lösungen ist jedoch bei allgemeinen Untersuchungen von einiger Bedeutung.

Naheliegend ist daher der Gedanke, ähnlich wie bei der Beschreibung von Wahrscheinlichkeitsverteilungen durch deren Parameter Entscheidungen lediglich auf Basis der ***Verteilungsparameter*** des unsicheren Vermögens zu treffen. Angesichts der Erfassung von Risiko und Risikoeinstellungen bietet es sich an, Erwartungswert und Varianz des Vermögens heranzuziehen. Dies kennzeichnet das (μ, σ)-Prinzip. Unschwer sind dabei Varianz oder Standardabweichung als Risikomaße zu identifizieren. Betrachtet man als Grenzfall das risikoindifferente Verhalten, reduziert sich das (μ, σ)-Prinzip auf das μ-Prinzip, weil das Risiko für die Beurteilung keine Rolle spielt.

[30] *Friend* (1977), S. 66.

4.2 (μ, σ)-Prinzip und Risikoeinstellung

Eine allgemeine Nutzenfunktion nach dem (μ, σ)-Prinzip lässt sich also beschreiben als $u = u(\mu, \sigma^2)$. Jede (μ, σ^2)-Nutzenfunktion lässt sich in eine äquivalente (μ, σ)-Nutzenfunktion überführen.

Die für das *Bernoulli*-Prinzip abgeleiteten Aussagen über Risikoeinstellungen lassen sich bei geringfügiger Modifikation auf das (μ, σ)-Prinzip übertragen. Ein Entscheider ist risikoavers, wenn gilt

$$\frac{\partial u(\mu, \sigma^2)}{\partial \sigma^2} < 0.$$

Als Maß für die absolute Risikoaversion dient die Grenzrate der Substitution zwischen μ und σ^2 (und der Kehrwert als Maß für die Risikotoleranz):

$$r_{a(\mu,\sigma)} = \left.\frac{d\mu}{d\sigma^2}\right|_{u=konst.} = -\frac{\partial u/\partial \sigma^2}{\partial u/\partial \mu}.$$

Das Sicherheitsäquivalent s muss die Bestimmungsgleichung

$$s = u(s, 0) = u(\mu, \sigma^2)$$

erfüllen. Die Risikoprämie ergibt sich wieder als Differenz zwischen dem Erwartungswert des Ergebnisses und seinem Sicherheitsäquivalent:

$$\lambda = \mu - u(s, 0).$$

Eine besonders einfache, häufig verwendete (μ, σ)-Nutzenfunktion lautet

$$u(\mu, \sigma^2) = \mu - \frac{1}{2}\gamma\sigma^2.$$

In diesem Fall stimmt der Nutzen mit dem Sicherheitsäquivalent überein. Weiter erhält man

$$r_{a(\mu,\sigma)} = \frac{1}{2}\gamma \text{ sowie } \lambda = \frac{1}{2}\gamma\sigma^2.$$

für den Risikoaversionskoeffizienten bzw. für die Risikoprämie.

In grafischen Darstellungen wie in Abbildung 10.6 verwendet man zumeist die Standardabweichung als Risikomaß. Das Sicherheitsäquivalent s zu einer beliebigen (μ, σ)-Kombination x ergibt sich als Schnittpunkt der zugehörigen (μ, σ)-Indifferenzkurve mit der Ordinate. Die Abbildung zeigt auch, dass eine flacher verlaufende Indifferenzkurve eine geringere Risikoaversion impliziert, da derselben riskanten Position x ein höheres Sicherheitsäquivalent und eine geringere Risikoprämie zugewiesen werden. Der Grenzfall der Risikoindifferenz ist durch eine horizontale Indifferenzkurve gekennzeichnet.

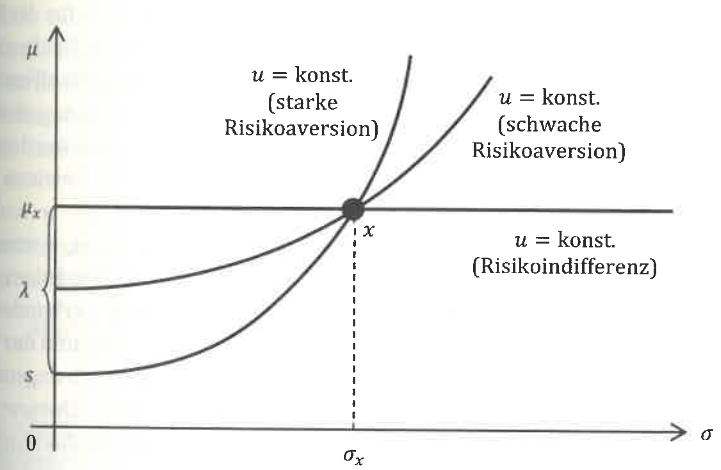

Abbildung 10.6:
Indifferenzkurve, Sicherheitsäquivalent und Risikoprämie beim (μ, σ)-Prinzip.

4.3 Schwächen des (μ, σ)-Prinzips

Zwar ist die analytische Handhabung des (μ, σ)-Prinzips einfacher als die des *Bernoulli*-Prinzips, auch können die wesentlichen Erkenntnisse in Bezug auf Risikoeinstellungen entsprechend angepasst werden. Jedoch sind deutliche Nachteile des (μ, σ)-Prinzips zu konstatieren.

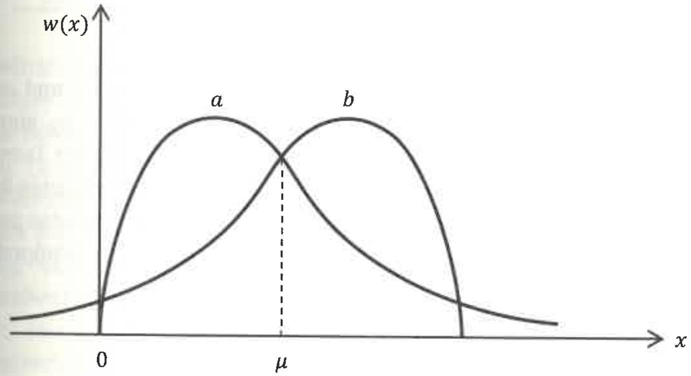

Abbildung 10.7:
Unterschiedliche Wahrscheinlichkeitsverteilungen mit übereinstimmenden Erwartungswerten und Varianzen.

Kritikwürdig ist zunächst, dass sehr unterschiedliche Verteilungen doch als äquivalent eingestuft werden. Besonders deutlich zeigt sich dieses Problem bei

Einbeziehung des dritten zentralen Moments, die ***Schiefe***, als Maß für die Symmetrie der Verteilung.[31] Die beiden Verteilungen in Abbildung 10.7 haben denselben Erwartungswert sowie dieselbe Varianz und weisen dennoch offensichtlich eine völlig andere Charakteristik auf. Bei Verteilung a sind negative Ergebnisse ausgeschlossen, und es können sehr hohe Gewinne erzielt werden. Bei Verteilung b gibt es hingegen eine niedrigere Obergrenze für Gewinne und gleichzeitig erhebliche Verlustgefahren. Trotz gleicher Parameter μ und σ^2 werden die meisten Individuen die rechtsschiefe Verteilung a vorziehen. Diese Beobachtung besagt aber letztlich nur, dass mit der Beschränkung auf zwei Parameter eine ***Einschränkung des Informationsgehalts*** verbunden ist. Das ist der Preis für die Vereinfachung der formalen Handhabung und der Herabsetzung des Informationsbedarfs. Vereinfachte Modelle sind stets ungenauer.

Ein zweiter Kritikpunkt gegenüber dem (μ,σ)-Prinzip wiegt schwerer: Bei Anwendung des (μ,σ)-Prinzips kann es zu einer ***Verletzung der Zustandsdominanz*** kommen. Das heißt, für eine Ergebnisverteilung kann ein niedrigerer Nutzen ausgewiesen werden, obwohl in allen Zuständen ein höheres Ergebnis erzielt wird als bei der anderen Aktion. Dies sei an einem einfachen Beispiel belegt. Die Nutzenfunktion möge lauten

$$u(\mu,\sigma) = \mu - 0{,}3\sigma^2.$$

Die Aktionen a_1 und a_2 sollen beurteilt werden:

	z_1 (0,5)	z_2 (0,5)	μ	σ^2	u
a_1	10	12	11	1	10,7
a_2	16	30	23	49	8,3

Tabelle 10.6: Verletzung der Zustandsdominanz.

a_2 führt in beiden Zuständen zu einem höheren Ergebnis als a_1 und ist deshalb nach dem Dominanzprinzip vorzuziehen. Gleichwohl weist a_2 einen geringeren (μ,σ)-Nutzen auf. Ursache für eine solche Verletzung der Dominanz ist, dass ein risikoaverser Entscheider die Abweichung von einem festen Erwartungswert auch dann als unangenehm empfindet, wenn es sich um eine positive Abweichung handelt. Die Überlegenheit hoher Ausprägungen wird dadurch verwischt.

4.4 Vereinbarkeit mit dem *Bernoulli*-Prinzip

Insbesondere angesichts des letzten Einwands gegen das (μ,σ)-Prinzip stellt sich die Frage, ob und unter welchen Bedingungen die jeweiligen spezifischen

[31] Die Schiefe ist definiert als $\mathrm{E}\{(\tilde{x}-\mu)^3\}$. In Abbildung 10.7 ist die Verteilung a durch eine positive Schiefe gekennzeichnet, sie ist „rechtsschief" oder auch „linkssteil"; dagegen ist die Verteilung b „linksschief" oder „rechtssteil", hat also eine negative Schiefe.

Vorteile von *Bernoulli*-Prinzip (ausgeprägte Rationalität) und (μ,σ)-Prinzip (einfache Handhabung) gemeinsam erfüllt sind. Ansatzpunkte für die Vereinbarkeit der Entscheidungsprinzipien ergeben sich zum einen bei den Nutzenfunktionen, zum anderen bei den Wahrscheinlichkeitsverteilungen.

Bei Vorliegen einer ***quadratischen Nutzenfunktion***

$$u(x) = x - \gamma x^2$$

ist der Nutzenerwartungswert stets durch Erwartungswert und Varianz der Ergebnisse eindeutig determiniert:

$$\mathrm{E}\{u(x)\} = \mathrm{E}\{x - \gamma x^2\} = \mathrm{E}\{x\} - \gamma\,\mathrm{E}\{x^2\}$$
$$= \mathrm{E}\{x\} - \gamma[(\mathrm{E}\{x\})^2 + \mathrm{Var}\{x\}] = \mu - \gamma(\mu^2 + \sigma^2).$$

Es ist allerdings zu konstatieren, dass die quadratische Nutzenfunktion mit der zunehmenden absoluten Risikoaversion eine ziemlich unplausible Risikoeinstellung impliziert. Außerdem ist zu gewährleisten, dass sämtliche Ausprägungen einer Zufallsvariablen auf dem steigenden Ast der Parabel liegen, da anderenfalls Verletzungen des Dominanzprinzips nicht ausgeschlossen werden können. Diesen Einwand kann man auch nicht dadurch entkräften, dass der Parameter γ der Nutzenfunktion variiert und auf diese Weise die Lage des Scheitelpunkts der Parabel an die Problemstellung angepasst wird. Denn das *Pratt-Arrow*-Maß als Indikator für den Grad der Risikoaversion ergibt sich im Falle der quadratischen Nutzenfunktion als

$$r_a = \frac{1}{\frac{1}{2\gamma} - x}.$$

Die individuelle Präferenz des Entscheiders schlägt sich also ausschließlich im Parameter γ nieder; eine im Parameter γ modifizierte Nutzenfunktion stünde für eine andere Risikoaversion.

Das (μ,σ)-Prinzip ist auch dann mit dem *Bernoulli*-Prinzip vereinbar, wenn die Wahrscheinlichkeitsverteilung der (μ,σ)-Klasse angehört, wenn es sich also um eine Verteilung handelt, die durch Erwartungswert und Varianz vollständig beschrieben ist. Aus den Verteilungen der (μ,σ)-Klasse hat allein die ***Normalverteilung*** die sogenannte Reproduktionseigenschaft, also die Eigenschaft, dass lineare Funktionen normalverteilter Zufallsvariablen wiederum normalverteilt sind.

Häufig werden zugleich die Normalverteilung und die exponentielle Nutzenfunktion unterstellt. Der Vorteil dieses sogenannten ***Hybrid-Modells***[32] ist die sehr einfache Gestalt der zugehörigen (μ,σ)-Nutzenfunktion, denn es gilt:[33]

[32] Der Begriff ist von *Bamberg* (1986), S. 21, übernommen.

[33] Siehe für die Umformungen *Stadler* (2003), S. 330.

$$\begin{aligned}\mathrm{E}\{u(x)\} &= \int_{-\infty}^{+\infty} [1-\exp(-\gamma x)]\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)dx \\ &= 1-\exp\left(-\gamma\left(\mu-\frac{1}{2}\gamma\sigma^2\right)\right) = u\left(\mu-\frac{1}{2}\gamma\sigma^2\right).\end{aligned}$$

Das Sicherheitsäquivalent

$$s = \mu - \frac{1}{2}\gamma\sigma^2$$

erweist sich als einfache, lineare Funktion von Erwartungswert und Varianz der Zielgröße.

Da das Hybrid-Modell auf der exponentiellen Nutzenfunktion basiert, impliziert es mit der konstanten absoluten Risikoaversion eine ebenfalls nicht sehr plausible Risikoeinstellung. Sofern der Risikoaversionsparameter γ geeignet kalibriert wird, lassen sich mit der Gestalt des Sicherheitsäquivalents die Ergebnisse der direkten Nutzenmaximierung sehr gut annähern.[34] Die genannte Formulierung ist daher recht weit verbreitet.

5. Stochastische Dominanz

5.1 Idee und Begriff

Sowohl das *Bernoulli*-Prinzip als auch das (μ,σ)-Prinzip sind für den Entscheider Instrumente zur Auswahl der individuell vorziehenswürdigen Wahrscheinlichkeitsverteilung. Dafür muss den besonderen Charakteristika des Entscheiders Rechnung getragen werden. Im *Bernoulli*-Prinzip geschieht dies durch die Heranziehung der Nutzenfunktion, im (μ,σ)-Prinzip durch Gewichtung von Erwartungswert und Standardabweichung der Zielgröße. Gerade die präzise Bestimmung der individuellen Charakteristika kann sich allerdings als besonders schwierig erweisen.

Daher stellt sich die Frage, ob auch ohne Rückgriff auf präzise bestimmte individuelle Präferenzen eine Auswahl oder wenigstens eine Vorauswahl zwischen unsicheren Ergebnissen vorgenommen werden kann. Zu diesem Zweck kann man das allgemeine Dominanzprinzip auf Wahrscheinlichkeitsverteilungen übertragen, und man kommt zu der Konzeption der ***stochastischen Dominanz*** als Kriterium zur Reihung von Wahrscheinlichkeitsverteilungen.[35] Hervorzuheben sind die stochastische Dominanz ersten Grades und die stochastische Dominanz zweiten Grades.

[34] Vgl. *Pulley* (1981) und *Kroll/Levy/Markowitz* (1984).
[35] *Hadar/Russell* (1969).

Von zwei Wahrscheinlichkeitsverteilungen A und B ist die erstgenannte genau dann ***stochastisch dominant ersten Grades*** zur zweiten, wenn der Wert der Verteilungsfunktion von A nie größer und für mindestens eine Ausprägung kleiner ist als der Wert der Verteilungsfunktion von B. Formal gilt also

$$A \succ_{SD1} B \iff \begin{cases} F_A(x) \leq F_B(x) & \text{für alle } x \text{ und} \\ F_A(x) < F_B(x) & \text{für mindestens ein } x' \end{cases}$$

wobei

$F_i(\cdot)$ Verteilungsfunktion der Zufallsvariablen i $(i = A, B)$.

Die Schreibweise „$\succ_{SD1}$" steht für „wird im Sinne der stochastischen Dominanz ersten Grades $(SD1)$ vorgezogen".

Die Verteilungsfunktion gibt bekanntlich an, mit welcher Wahrscheinlichkeit eine bestimmte Ergebnisausprägung x nicht überschritten wird. Ein niedrigerer Wert der Verteilungsfunktion ist demnach gleichbedeutend damit, dass ein größerer Teil des Wahrscheinlichkeitsgewichts auf Werte von größer als x entfällt; umgekehrt ist zu jeder Wahrscheinlichkeit der zugehörige Ergebniswert x (das zugehörige Quantil) nicht niedriger. $SD1$ steht also für im stochastischen Sinne größere Ergebniswerte.

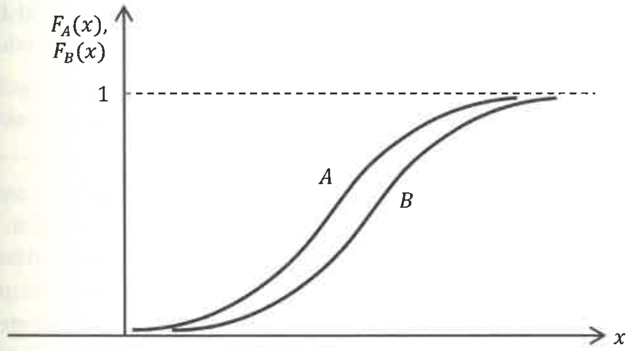

Abbildung 10.8: Stochastische Dominanz ersten Grades $(SD1)$.

Grafisch dargestellt bedeutet dies, dass eine nach $SD1$ überlegene Verteilungsfunktion stets rechts bzw. unterhalb der unterlegenen Verteilungsfunktion verläuft. Abbildung 10.8 verdeutlicht dies für einen Verteilungstyp, welcher der Normalverteilung nachempfunden ist. Bei diskreten Zufallsvariablen ergeben sich jeweils Treppenfunktionen,[36] von denen die $SD1$-dominante Verteilung unterhalb bzw. weiter rechts verläuft. In jedem Fall gilt: Besteht zwischen zwei Verteilungen die Relation der $SD1$, so schneiden sich deren Verteilungsfunktionen nicht.

[36] Vgl. Abschnitt 2.2 dieses Kapitels.

Nun ist allerdings keineswegs stets von zwei Verteilungen eine stochastisch dominant nach $SD1$; es handelt sich daher lediglich um ein partielles Reihungskriterium. Das heißt, es gibt Situationen, in denen der Entscheider alleine auf Basis der $SD1$ keine endgültige Entscheidung vornehmen kann.

Dies wirft die Frage auf, ob ein weniger strenger Dominanzbegriff hier weiterhelfen könnte. Die ***stochastische Dominanz zweiten Grades*** ($SD2$) ist eine solche weniger strenge Dominanzrelation. Eine Verteilung A erfüllt die Eigenschaft der $SD2$ gegenüber B, wenn die oben genannte Relation nicht für die Verteilungsfunktion, sondern für das Integral über die Verteilungsfunktion gilt, formal also:

$$A \succ_{SD2} B \quad \Leftrightarrow \quad \begin{cases} \int_{x_{min}}^{x} F_A(\xi)d\xi \leq \int_{x_{min}}^{x} F_B(\xi)d\xi & \text{für alle } x \text{ und} \\ \int_{x_{min}}^{x} F_A(\xi)d\xi < \int_{x_{min}}^{x} F_B(\xi)d\xi & \text{für mindestens ein } x \end{cases}$$

Ist die $SD1$ gegeben, so ist stets auch die $SD2$ erfüllt, aber nicht umgekehrt; daher ist die $SD2$ eine schwächere Bedingung. Auf Basis der $SD2$ lassen sich mehr Verteilungen in eine Reihung bringen, als es bei der $SD1$ der Fall ist. Abbildung 10.9 zeigt ein Beispiel dafür, dass zwar $SD2$ vorliegt, nicht aber $SD1$. Letzteres erkennt man daran, dass sich die Verteilungsfunktionen schneiden.

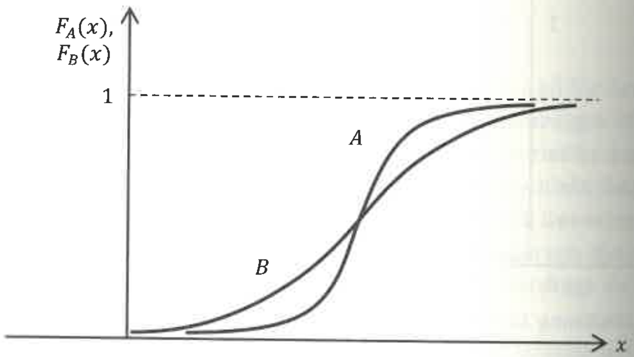

Abbildung 10.9: Stochastische Dominanz zweiten Grades ($SD2$) ohne $SD1$.

Abbildung 10.9 verdeutlicht, dass bei kleineren Ergebnissen die Verteilung A gegenüber B im Vorteil ist; das Ausmaß spiegelt sich in der Integraldifferenz, also der Fläche zwischen den Verteilungsfunktionen wider. Jenseits des Schnittpunkts der Verteilungsfunktionen holt B gegenüber A auf. Die Integralbedingung für die $SD2$ besagt gerade, dass B die Verteilung A nicht überholt. Auch diese Überlegungen lassen sich wie in Abschnitt 5.3 auf Treppenfunktionen und damit auf diskrete Verteilungen übertragen.

Für die weitere Interpretation sei festgehalten, dass die nach $SD2$ unterlegene Verteilung B ein höheres Wahrscheinlichkeitsgewicht bei sehr großen und sehr kleinen Ergebniswerten aufweist; dies ist am größeren Anstieg der Verteilungsfunktion B bei den betreffenden Ergebniswerten zu erkennen. Die Verteilung B ist daher schon nach einer ersten Intuition mit einem größeren Risiko verbunden als die Verteilung A.

5.2 Relation zum *Bernoulli*-Prinzip

Mit den beiden vorgestellten Ausprägungen der stochastischen Dominanz sind zunächst weitere Reihungskriterien eingeführt. Solange keine klare Relation zu den zuvor eingeführten Entscheidungsprinzipien hergestellt wird, bleibt unklar, was damit gewonnen ist. Tatsächlich lassen sich solche Relationen aber nachweisen.[37] In Bezug auf $SD1$ gilt

> Eine Ergebnisverteilung wird genau dann von ***allen Entscheidern*** mit stets positivem Grenznutzen vorgezogen, wenn sie $SD1$-dominant ist.

Eine ähnliche Aussage lässt sich auch für die $SD2$ ableiten:

> Eine Ergebnisverteilung wird genau dann von ***allen risikoaversen Entscheidern*** vorgezogen, wenn sie $SD2$-dominant ist.

Die strengere Bedingung der $SD1$ impliziert unmittelbar, dass jeder Entscheider, der mehr Geld weniger Geld vorzieht, die $SD1$-dominante Verteilung wählt, unabhängig davon, welche Risikopräferenz er aufweist. Die schwächere Bedingung der $SD2$ impliziert, dass bei der zusätzlichen Annahme der Risikoaversion stets eine $SD2$-dominante Verteilung präferiert wird, unabhängig von der Stärke der Risikoaversion. Sofern also die Relation der stochastischen Dominanz gegeben ist, sind darauf beruhende Entscheidungen stets vereinbar mit dem *Bernoulli*-Prinzip, einmal schlechthin, einmal mit Blick auf alle risikoaversen Entscheider.

Allerdings kann weder mit der $SD1$ noch mit der $SD2$ eine vollständige Reihung beliebiger Wahrscheinlichkeitsverteilungen vorgenommen werden. Um dies garantieren zu können, müssen mehr genauere Informationen über die Präferenzen des Entscheiders verfügbar sein. Der zwischen den beiden wünschenswerten Anforderungen an Entscheidungsprinzipien „Allgemeingültigkeit" und „Ordnungsfähigkeit" bestehende Zielkonflikt lässt sich nicht vollständig auflösen.

37 Siehe hier und im Folgenden für die Beweise *Hadar/Russell* (1969) und *Wolfstetter* (1999)

5.3 Ein Beispiel

Nachdem zuvor die stochastische Dominanz grafisch anhand einer kontinuierlichen Zufallsvariablen verdeutlicht wurde, wird nun ein konkretes Beispiel für die Entscheidungsfindung mittels stochastischer Dominanz bei einer diskreten Verteilung analysiert. Zu diesem Zweck wird das bereits oben eingeführte Beispiel[38] zur Einführung innovativer Produkte wieder aufgegriffen:

	Entwicklung des innovativen Marktsegments		
Absatzmenge	schlecht (0,2)	normal (0,6)	gut (0,2)
Produkt 1	100	300	400
Produkt 2	240	260	380
Produkt 3	360	240	120

Tabelle 10.1: Absatzmengen und Marktentwicklung.

Bei den Produkten 1 und 2 handelt es sich um unterschiedliche Innovationen. Produkt 3 stellt eine konventionelle Produktvariante dar. Nachstehende Abbildung 10.10 gibt die zugehörigen Verteilungsfunktionen wieder. Die kurz gestrichelte Linie bezieht sich dabei auf das Produkt 1, die durchgezogene Linie auf Punkt 2 sowie die lang gestrichelte Linie auf Produkt 3.

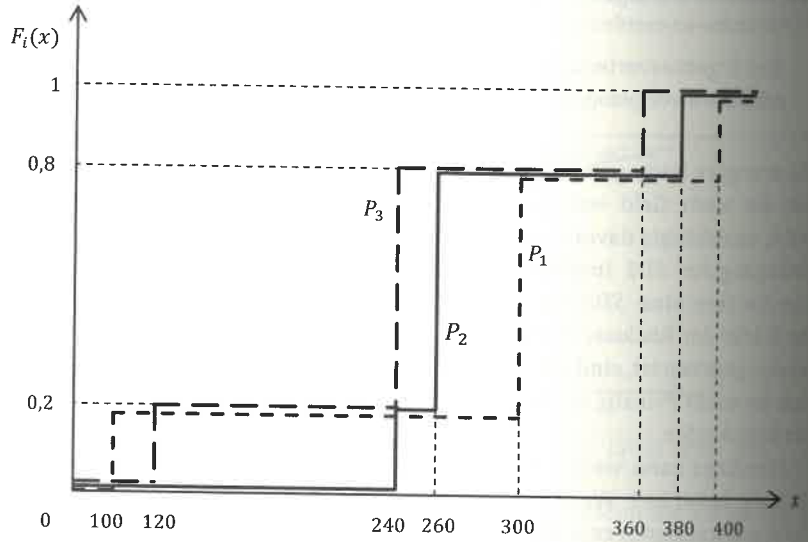

Abbildung 10.10: Verteilungsfunktionen der Absatzmengen.[39]

---

[38] Siehe Abschnitt 2.3 dieses Kapitels.

[39] Um den Funktionsverlauf deutlicher werden zu lassen, sind auch die Sprungstellen der Treppenfunktionen entsprechend gekennzeichnet; zudem sind die Verteilungsfunktionen bei den Produkten 2 und 3 jeweils etwas nach oben abgesetzt.

Der Vergleich der Verteilungsfunktionen zeigt, dass die durchgezogene Linie für Produkt 2 teilweise auf gleicher Höhe, teilweise unter, aber niemals über der lang gestrichelten Linie der Verteilungsfunktion von Produkt 3 verläuft. Die oben angegebene Gleichung für das Vorliegen von $SD1$ ist damit erfüllt. Daher zieht im Falle der Absatzmengen als Entscheidungskriterium jeder Entscheider mit einer *Bernoulli*-Nutzenfunktion die Absatzverteilung von Produkt 2 vor. Die Beibehaltung der konventionellen Produktvariante 3 scheidet also generell aus.

Vergleicht man die Verteilungsfunktionen der beiden innovativen Produktvarianten 1 und 2, so zeigt sich, dass sich die Verteilungsfunktionen schneiden. Im Intervall von $100 \leq x < 240$ verläuft die durchgezogene Verteilungsfunktion von Produkt 2 unterhalb der gestrichelten Funktion von Produkt 1; in den Intervallen $260 \leq x < 300$ sowie $380 \leq x < 400$ verhält es sich umgekehrt. Zwischen den Absatzverteilungen besteht also nicht die Relation der $SD1$. Demnach kann *ohne weiteres* keine Entscheidung über die Vorteilhaftigkeit der beiden innovativen Produktvarianten getroffen werden.

Allerdings können die Verteilungen auf das Kriterium der $SD2$ überprüft werden. Zu diesem Zweck müssen die Integrale über die jeweiligen Verteilungsfunktionen miteinander verglichen werden. Aus den Treppenfunktionen resultieren abschnittsweise linear ansteigende Funktionen $\int F_i(\xi)d\xi$. Analytisch erhält man

$$\int_0^x F_1(\xi)d\xi = \begin{cases} 0 & \text{für} \quad 0 \leq x < 100 \\ 0{,}2(x-100) & \text{für} \quad 100 \leq x < 300 \\ 40 + 0{,}8(x-300) & \text{für} \quad 300 \leq x < 400 \\ 120 + (x-400) & \text{für} \quad 400 \leq x \end{cases},$$

$$\int_0^x F_2(\xi)d\xi = \begin{cases} 0 & \text{für} \quad 0 \leq x < 240 \\ 0{,}2(x-240) & \text{für} \quad 240 \leq x < 260 \\ 4 + 0{,}8(x-260) & \text{für} \quad 260 \leq x < 380 \\ 100 + (x-380) & \text{für} \quad 380 \leq x \end{cases}.$$

Abbildung 10.11 gibt die entsprechenden Funktionen grafisch wieder. Gleichung und Abbildung zeigen, dass das P_2-Integral im Intervall $100 < x < 400$ unterhalb des P_1-Integrals verläuft; außerhalb dieses Intervalls haben die Integrale einen identischen Verlauf. Somit sind die Bedingungen für $SD2$ erfüllt: Jeder risikoaverse Entscheider zieht die Absatzverteilung von Produkt 2 der Absatzverteilung von Produkt 1 vor. Kenntnisse über den Grad der Risikoaverse über den Grad der Risikoaversion sind für diese Folgerung nicht erforderlich.

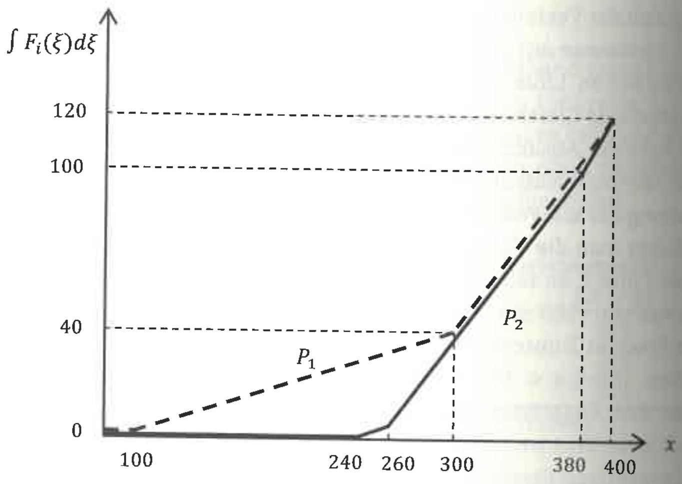

Abbildung 10.11: Integrale über die Verteilungsfunktionen.

5.4 Relation zum (μ, σ)-Prinzip

Im Abschnitt über die Parameter von Wahrscheinlichkeitsverteilungen sind für die in Tabelle 10.1 angegebenen Verteilungen auch die Erwartungswerte und Varianzen angegeben.[40] Es gilt:

	Erwartungswert	Varianz
Produkt 1	280	9.600
Produkt 2	280	2.560
Produkt 3	240	5.760

Tabelle 10.7: Verteilungsparameter der Absatzmengen.

In diesem Beispiel wird die nach $SD2$ für alle risikoaversen Entscheider überlegene Absatzverteilung von Produkt 2 auch nach der (μ, σ)-Dominanz eindeutig vorgezogen: Bei keiner anderen Verteilung ist der Erwartungswert höher und bei keiner anderen Verteilung ist die Varianz niedriger als bei den Absatzmengen von Produkt 2.

Über den Zusammenhang zwischen der stochastischen Dominanz und der Relation von Verteilungsparametern lassen sich einige allgemeingültige Zusammenhänge festhalten:[41]

[40] Vgl. Abschnitt 2.3 dieses Kapitels.

[41] Siehe auch *Neus* (2007), Sp. 1774 f.

Bei einer $SD1$-dominanten Verteilung ist stets der Erwartungswert höher.
Der Erwartungswert einer $SD2$-dominanten Verteilung ist nicht kleiner als der einer $SD2$-dominierten Verteilung.
Ist eine Verteilung $SD2$-dominant und hat den gleichen Erwartungswert, so hat sie eine kleinere Varianz als die $SD2$-dominierte Verteilung.

Die Gültigkeit der Aussagen lässt sich für das vorliegende Beispiel unschwer überprüfen. Mit Blick auf die Reichweite der Aussagen ist aber zu beachten, dass der Umkehrschluss nicht gilt:

Aus einem höheren Erwartungswert folgt nicht zwingend die $SD1$.
Eine bei gleichem Erwartungswert geringere Varianz impliziert nicht zwingend die $SD2$.

Zusammenfassend lässt sich daher noch einmal festhalten, dass die stochastische Dominanz relativ zum (μ, σ)-Prinzip gegebenenfalls die zuverlässigeren Entscheidungen ermöglicht. Die mit der vereinfachenden Heranziehung nur zweier Verteilungsparameter verbundenen Schwächen[42] lassen sich nur durch differenziertere Instrumente wie die stochastische Dominanz oder – wenn Nutzenfunktionen bestimmt werden können – das *Bernoulli*-Prinzip vermeiden.

Wiederholungsfragen und Übungsaufgaben

Lösungshinweise *https://online.mohr.de/elib/neus.*

Aufgabe 10.1

Gegeben sind zwei Unternehmensstrategien 1 und 2, deren Gewinnaussichten vom Unternehmer wie folgt eingeschätzt werden:

$$g_1 = \begin{cases} 30 & \text{mit } 50\% \\ 170 & \text{mit } 50\% \end{cases}; \quad g_2 = \begin{cases} 50 & \text{mit } 98\% \\ 590 & \text{mit } 2\% \end{cases}.$$

a) Welche Strategie zieht der Unternehmer vor, wenn er risikoindifferent ist?

[42] Vgl. Abschnitt 4.3 dieses Kapitels.

b) Wie würde der Unternehmer entscheiden, wenn er risikoavers ist und sich nur an Erwartungswert und Varianz des Gewinns orientiert?
c) Beurteilen Sie das (μ, σ)-Prinzip angesichts Ihrer Antwort unter b).

Aufgabe 10.2

Ein vorsichtiger Beamter muss eine Entscheidung treffen, aufgrund derer er möglicherweise zu einer Amtshaftpflicht herangezogen werden kann. Seine Schadensersatzpflicht hängt vom Eintritt ihm unbekannter und nicht beeinflussbarer Sachverhalte sowie von der Sorgfalt der Erledigung seiner Aufgabe ab. Über die Höhe des Schadens und die Eintrittswahrscheinlichkeiten sind dem Beamten folgende Informationen bekannt:

	Zustand (Wahrscheinlichkeit)		
Schaden	z_1 (95%)	z_2(4%)	z_3(1%)
Sorgfalt	0	0	500
Schlamperei	0	500	500

Wenn der Beamte den Schaden ersetzen muss, wird sein Einkommen von 4.000 um den Schadensbetrag gekürzt. Der Beamte möchte seinen Nutzenerwartungswert maximieren, wobei er sein Einkommen x mit der folgenden Nutzenfunktion bewertet:

$$u(x) = x - \frac{x^2}{10.000}.$$

a) Wäre der Beamte, der sich selbst für unbedingt zuverlässig hält, bereit, eine Prämie von 10 für eine Haftpflichtversicherung zu bezahlen?
b) Wenn es zu einem Versicherungsabschluss kommt: Womit muss die Versicherung rechnen, für die der Beamte ein Neukunde ist?

Aufgabe 10.3

Ein Unternehmer schätzt die Gewinnaussichten seines Unternehmens für das folgende Jahr wie folgt ein:

$$g = \begin{cases} 0 & \text{mit } 10\% \\ 1.000.000 & \text{mit } 40\% \\ 4.000.000 & \text{mit } 40\% \\ 9.000.000 & \text{mit } 10\% \end{cases}.$$

Er bewertet den Gewinn mit der Nutzenfunktion $u(g) = \sqrt{g}$.
a) Wie hoch ist der Nutzenerwartungswert des Unternehmers?
b) Welchen Wert nimmt das Sicherheitsäquivalent des Gewinns an?
c) Wie hoch ist die Risikoprämie?

Aufgabe 10.4

Eine Versicherung möchte ihre Produktpalette durch eine Rechtsschutzversicherung abrunden. Vor der Produkteinführung werden ausgiebige Marktuntersuchungen mit den folgenden Ergebnissen vorgenommen:

- Die potentielle Kundschaft umfasst 95% wenig streitlustige Kunden, aber 5% notorische Querulanten, die wegen jeder Kleinigkeit vor Gericht ziehen.
- Von je 100.000 wenig streitlustigen Kunden werden es voraussichtlich jährlich 20 zu einem Prozess kommen lassen.
- Bei je 100.000 Querulanten kommt es jährlich voraussichtlich zu 1.500 Prozessen.
- Im Durchschnitt wird ein Prozess für die Versicherung zu Kosten von 20.000 führen.

a) Welche jährliche Versicherungsprämie muss die Versicherung zur Abdeckung der erwarteten Prozesskosten kalkulieren, wenn die tatsächliche Kundschaft genauso zusammengesetzt ist wie die potentielle Kundschaft?
b) Welche Gefahr besteht für die Versicherung bei dieser Kalkulation?

Aufgabe 10.5

Ein Investor steht vor der Entscheidung, das ihm zur Verfügung stehende Budget in Höhe von $b = 2{,}5$ Mio. € auf risikolose Staatspapiere mit der Rendite $i = 6\%$ und auf eine Beteiligung an einem Unternehmen aufzuteilen. Die Beteiligung erbringt eine unsichere Rendite r mit dem Erwartungswert $\mu = 8\%$ und der Standardabweichung von $\sigma = 4\%$. Der Investor bewertet sein Endvermögen y mit der (μ, σ)-Nutzenfunktion

$$u = \mathrm{E}\{y\} - \frac{\mathrm{Var}\{y\}}{250.000}.$$

a) Wie hoch ist der optimale in das Unternehmen investierte Betrag I?
Nun wird eine einfache lineare Gewinnsteuer mit konstantem Steuersatz $s = 50\%$ einbezogen, die gegebenenfalls einen vollen Verlustausgleich vorsieht. Bemessungsgrundlage für die Gewinnsteuer ist das Endvermögen y abzüglich des Budgets b. Es gilt also für die Steuerzahlung $S = s(y - b)$. Der Nutzen bezieht sich nun auf das Endvermögen nach Steuern $y_s = y - S$.
b) Wieviel wird jetzt in das Unternehmen investiert?
c) Erklären Sie den zunächst vielleicht überraschenden Effekt.

Aufgabe 10.6

In der Controllingabteilung eines Unternehmens wird über die richtige Bemessungsgrundlage für kurzfristig wirksame Entscheidungen gestritten. Es stehen zwei Handlungsalternativen zur Disposition, von denen die eine einen sicheren

Gewinn von 500 €, die andere mit je 50% Wahrscheinlichkeit einen Gewinn von 200 € oder 1.200 € verspricht. Bei der Ermittlung des Gewinns wurden jeweils Fixkosten in Höhe von 300 € in Abzug gebracht. Einigkeit besteht darüber, dass der Nutzenerwartungswert maximiert werden soll und die Nutzenfunktion des Entscheiders lautet

$$u(x) = x - \frac{x^2}{2.800}.$$

Strittig ist dagegen, ob auf Vollkostenbasis oder auf Teilkostenbasis gerechnet werden soll, ob also die Gewinne oder die Deckungsbeiträge der Entscheidung zugrunde liegen sollen.

a) Welche Alternative ist besser, wenn auf Basis von Gewinnen gerechnet wird, welche, wenn auf Basis der Deckungsbeiträge kalkuliert wird?
b) Erklären Sie das Ergebnis.
c) Lässt sich aus Sicht des *Bernoulli*-Prinzips sagen, welcher Kalkül vorzuziehen ist?

Aufgabe 10.7

a) Definieren Sie in Worten (also ohne Zuhilfenahme von Gleichungen) die Begriffe Sicherheitsäquivalent und Risikoprämie.

Ein Entscheider mit der Nutzenfunktion $u(x) = (x + 6)^2$ soll die folgende Lotterie bewerten:

$$x = \begin{cases} 0 & \text{mit} \quad 40\% \\ 10 & \text{mit} \quad 60\% \end{cases}.$$

b) Wie hoch ist die von ihm verlangte Risikoprämie?
c) Welche Risikoeinstellung besitzt der Entscheider?

Aufgabe 10.8

Zwei Eigenschaften interessieren Sie an Ihrem Geschäftspartner ganz besonders: Ist er ehrlich oder nicht? Und: Ist er fähig oder nicht?

Die Wahrscheinlichkeit dafür, dass er ehrlich ist, schätzen Sie mit 60% ein. Mit 15% Wahrscheinlichkeit gehen Sie davon aus, dass er zwar ehrlich, aber unfähig ist.

- Bestimmen Sie die Wahrscheinlichkeit dafür, dass er unfähig ist, für den Fall, dass Ihnen ein Dritter unzweifelhaft bestätigt, Ihr Geschäftspartner sei ehrlich.

Aufgabe 10.9

- Berechnen Sie in allgemeiner Form die Varianz einer Zufallsvariablen, die im Intervall $[0; z]$ stetig gleichverteilt ist.

Aufgabe 10.10

Sie haben Aktien zweier Unternehmen a und b erworben. Mit 20% Wahrscheinlichkeit steigt nur der Kurs von Aktie a. Mit 50% Wahrscheinlichkeit steigen beide Aktienkurse, mit 30% Wahrscheinlichkeit steigt keiner der Kurse.

a) Sie erfahren, dass Kurs a gestiegen ist. Wie hoch ist die Wahrscheinlichkeit dafür, dass auch Kurs b gestiegen ist?
b) Sie erfahren, dass Kurs b gestiegen ist. Wie hoch ist die Wahrscheinlichkeit dafür, dass auch Kurs a gestiegen ist?
c) Mit welcher Wahrscheinlichkeit steigt nur Kurs b?

Aufgabe 10.11

Beurteilen Sie, ob die folgenden Aussagen richtig oder falsch sind.

a) Das *Bernoulli*-Prinzip ist allgemeiner als das (μ, σ)-Prinzip, weil ersteres nicht nur für Risikosituationen, sondern auch für den Fall der Ungewissheit geeignet ist.
b) Die Risikoprämie ist definiert als die Differenz zwischen Erwartungswert und Sicherheitsäquivalent.
c) Die empirische Entscheidungsforschung zeigt, dass bei sehr hohen Ergebniswerten und bei sehr geringen Wahrscheinlichkeiten bisweilen Entscheidungen getroffen werden, die mit dem *Bernoulli*-Prinzip nicht vereinbar sind.
d) Die Kovarianz ist auf den Bereich zwischen Null und Eins normiert.
e) Bei der quadratischen Nutzenfunktion ist zwar generell eine Vereinbarkeit von *Bernoulli*-Prinzip und (μ, σ)-Prinzip gegeben, infolge des fallenden Astes der Nutzenfunktion kann dennoch eine Verletzung des Dominanzprinzips nicht ausgeschlossen werden.
f) Risikoaverse Entscheider ziehen bei normalverteilten Gewinnen mit gleichem Erwartungswert stets die Lösungen mit der geringsten Varianz vor.
g) Bei Risikoaversion wird stets die Lösung mit der geringsten Varianz vorgezogen.
h) Empirischen Untersuchungen zufolge riskieren Menschen mit geringem Vermögen mehr als Menschen mit hohem Vermögen.
i) Mit Hilfe des Satzes von *Bayes* kann man rationales Lernen bei unsicheren Erwartungen quantifizieren.
j) Im Rahmen des *Bernoulli*-Prinzips führt eine positiv-lineare Transformation der Nutzenfunktion keinesfalls dazu, dass nach der Transformation eine andere Alternative gewählt wird als vorher.

k) Nur die lineare Nutzenfunktion weist die Eigenschaft der konstanten absoluten Risikoaversion auf.

Literaturhinweise

Die Wahrscheinlichkeitsrechnung gehört zur Grundausbildung in Statistik. Empfehlenswert sind die Darstellungen bei *Bamberg/Baur/Krapp* (2012), Teil II, und *Poddig/Dichtl/Petersmeier* (2008), insb. Kapitel 2, letztere mit Anwendungsbezug in Richtung Finanzmärkte.

Das *Bernoulli*-Prinzip und das (μ, σ)-Prinzip werden in Lehrbüchern zur Entscheidungstheorie vorgestellt und diskutiert. Hier sei auf *Eisenführ/Weber/Langer* (2010) und *Laux/Gillenkirch/Schenk-Mathes* (2014) verwiesen. Eine umfassende Kritik der Theorie rationaler Entscheidungen aus Sicht der verhaltensorientierten Entscheidungslehre bietet *Kahneman* (2012); dieses exzellent geschriebene Buch kann durchaus auch zuhause auf dem Sofa mit Genuss gelesen werden. Einen verglichen mit dem vorliegenden Abschnitt 5 weiter gehenden, aber gut zugänglichen Zugang zur Stochastischen Dominanz vermitteln *Kruschwitz/Husmann* (2012), Abschnitt 2.4.

Schlüsselbegriffe

Bernoulli-Prinzip
(μ, σ)-Prinzip
Risikoaversion
Risikoprämie
Satz von *Bayes*
Sicherheitsäquivalent
Stochastische Dominanz

Kapitel 11

Theorie nicht-kooperativer Spiele

1. Grundelemente und Darstellungsformen

Die Entscheidungstheorie stellt das angemessene Instrumentarium zur Untersuchung der Entscheidungen einzelner Menschen bereit. Der überwiegende Teil der Betriebswirtschaftslehre beschäftigt sich jedoch, wie gesehen, mit der Interdependenz von Entscheidungen mehrerer Individuen; dafür gibt die Entscheidungstheorie nicht sehr viel her. Vielmehr ist dies Gegenstand der Spieltheorie, die sich somit als die angemessene Methodik zur Untersuchung der ***Interaktion rational handelnder Menschen*** erweist.

Es gibt einige Gemeinsamkeiten zwischen Entscheidungstheorie und Spieltheorie: Die Konsequenzen einer Entscheidung hängen nicht alleine von der Entscheidung selbst ab, sondern auch von äußeren Umständen. Im Fall der Entscheidungstheorie ist dies der Zufall, also eine exogene Unsicherheit. Im Fall der Spieltheorie wird dagegen die Mehrwertigkeit der Entscheidungsfolgen dadurch hervorgerufen, dass die Entscheidungen anderer, ebenfalls rational handelnder Individuen das Ergebnis beeinflussen. Dies schließt natürlich nicht aus, dass zusätzlich der von keinem Individuum beeinflussbare Zufall mitwirkt. In ökonomischen Anwendungen ist dies sogar die Regel, auch wenn die Zufallswirkungen in spieltheoretischen Ansätzen nicht immer expliziert werden.

Der Begriff der Spieltheorie war zunächst der mathematischen Theorie von Spielen wie Schach oder Skat vorbehalten. Auch militärische Fragen wie die richtige Strategie bei der Androhung von Nuklearschlägen wurden zu Zeiten des Kalten Krieges mit Hilfe der Spieltheorie untersucht. Die Fruchtbarkeit der Übertragung auf ökonomische Fragestellungen ist offensichtlich, insbesondere bei einer Untersuchung der Gestaltung von Kooperationen. Insbesondere in ***institutionenökonomischen Ansätzen*** findet die Spieltheorie daher eine breite wirtschaftswissenschaftliche Anwendung.

Ein Spiel wird durch seine Regeln beschrieben. Sie umfassen die Festlegung der teilnehmenden Spieler, der einzelnen Spielzüge und der Ergebnisse. Daher sind diese Elemente bei der Modellierung einer ökonomischen Interaktion als Spiel vorab präzise zu beschreiben. ***Spieler*** sind alle Individuen, die eigenständige Entscheidungen treffen können. Daneben kann der Zufall als weiterer Spieler einbezogen werden, dessen „Handlungen" jedoch nicht der Rationalität folgen. In Bezug auf die ***Spielzüge*** ist genau zu bestimmen, welcher Spieler wann