3 Prämienkalkulationsprinzipien und Risikomaße

Ausgehend von der in den vorhergehenden Kapiteln gegebenen Beschreibung des Risikos eines Versicherungsportfolios muss das Versicherungsunternehmen eine ausreichende Prämie bestimmen. Zudem ist festzulegen, wie diese Gesamtprämie auf die einzelnen Risiken umzulegen ist. In der Lebensversicherungsmathematik wurde dazu stets die Äquivalenzprämie verwendet. Ignoriert man die Unterschiede bei den Zahlungszeitpunkten, so führt dies zur Nettorisikoprämie.

3.1 Definition *Bezeichnet $X \geqslant 0$ die zufällige Schadenhöhe (eines Portfolios oder eines Einzelrisikos), so heißt $E(X)$ die zugehörige* **Nettorisikoprämie**. □

Für das Versicherungsunternehmen ist die Nettorisikoprämie auf die Dauer *nicht* ausreichend, um den technischen Ruin zu vermeiden. Bezeichnen u das Anfangskapital, X_i, $1 \leqslant i \leqslant n$, die als unabhängig und identisch verteilt angenommenen Schadenhöhen in n Versicherungsperioden und π die konstante Prämie, so ist

$$Z_n := u - \sum_{i=1}^{n}(X_i - \pi)$$

das Kapital nach n Perioden, wenn man jegliche Kapitalgewinne vernachlässigt. Man spricht dann von einem technischen Ruin in den ersten n Perioden, falls $\min_{1\leqslant i\leqslant n} Z_i < 0$, das Unternehmen also zu einem Zeitpunkt alles Kapital aufgebraucht hat und die Zahlungsverpflichtungen rechnerisch nicht mehr erfüllen kann. Wir wollen nun mit Hilfe des Zentralen Grenzwertsatzes eine untere Schranke für die Wahrscheinlichkeit berechnen, dass irgendwann der technische Ruin eintritt. Dazu benötigen wir das folgende Hilfsresultat.

3.2 Lemma *Sind F_n, $n \in \mathbb{N}$, Vf. und F eine stetige Vf., so dass $\lim_{n\to\infty} F_n(x) = F(x)$ für alle $x \in \mathbb{R}$, so gilt sogar $\lim_{n\to\infty} F_n(x_n) = F(x)$ für jede Folge $x_n \in \mathbb{R}$ mit $\lim_{n\to\infty} x_n = x$.* □

Beweis. Für $x \in \mathbb{R}$ und $\varepsilon > 0$ gilt wegen der Monotonie der Vf. F_n für jede Folge $x_n \in \mathbb{R}$ mit $\lim_{n\to\infty} x_n = x$

$$\begin{aligned}\liminf_{n\to\infty} F_n(x_n) &\geqslant \liminf_{n\to\infty} F_n(x-\varepsilon) = F(x-\varepsilon)\\ \limsup_{n\to\infty} F_n(x_n) &\leqslant \limsup_{n\to\infty} F_n(x+\varepsilon) = F(x+\varepsilon).\end{aligned}$$

Da die rechten Seiten für $\varepsilon \downarrow 0$ gegen $F(x)$ konvergieren, folgt daraus die Behauptung. □

3.3 Bemerkung Die Folgerung von Lemma 3.2 ist äquivalent zur lokalen Gleichmäßigkeit der Konvergenz, d.h. für alle $x \in \mathbb{R}$ kann man eine Umgebung von x finden, auf der die F_n gleichmäßig gegen F konvergieren. In der Tat kann man sogar zeigen, dass unter den genannten Voraussetzungen die F_n auf ganz $\mathbb{R}$ gleichmäßig gegen F konvergieren. □

Wählt man nun speziell $\pi = E(X_1)$ und nimmt $\sigma^2 := Var(X_1) \in (0, \infty)$ an, so zeigt der Zentrale Grenzwertsatz zusammen mit dem obigen Lemma, dass

$$\begin{aligned} P\Big\{ \min_{1 \leqslant i \leqslant n} Z_i < 0 \Big\} &\geqslant P\Big\{ \sum_{i=1}^{n} (X_i - \pi) > u \Big\} \\ &= P\Big\{ \frac{1}{\sqrt{n\sigma^2}} \sum_{i=1}^{n} (X_i - E(X_i)) > \frac{u}{\sqrt{n\sigma^2}} \Big\} \\ &\longrightarrow 1 - \Phi(0) = \frac{1}{2}, \end{aligned}$$

für $n \to \infty$, da $u/\sqrt{n\sigma^2} \to 0$. Es tritt also wenigstens mit der Wahrscheinlichkeit 1/2 irgendwann Ruin ein. Man kann sogar (mit Hilfe des Satzes vom iterierten Logarithmus) zeigen, dass f.s. Ruin eintritt und dass dies auch dann gilt, wenn die Varianz von X_1 nicht endlich ist! Anschaulich ist dies plausibel, da die Nettorisikoprämie im Mittel gerade ausreicht, die eintretenden Schäden zu bezahlen, mit steigender Anzahl n der betrachteten Perioden die zufälligen Abweichungen der summierten Schadenhöhen von ihrem Erwartungswert aber absolut gesehen immer größer werden. (Gemäß dem zentrale Grenzwertsatz sind diese Abweichungen von der Größenordnung $\sqrt{n}$, falls die Varianz endlich ist.) Irgendwann wird daher der Fall eintreten, dass die summierten zufälligen Abweichungen das Anfangskapital übersteigen und somit der technische Ruin eintritt.

Diese Überlegungen zeigen, dass die Versicherungsprämie für das durch X beschriebene Risiko größer als $E(X)$ gewählt werden muss.

3.4 Definition *Die Differenz zwischen der Nettoprämie für das durch X beschriebene Risiko (d.h. der Prämie bei Vernachlässigung von Verwaltungskosten u.ä.) und $E(X)$ wird als* **Sicherheitszuschlag** *(engl.:* **safety loading***) bezeichnet.* □

3.5 Bemerkung (i) In der Lebensversicherungsmathematik werden keine expliziten Sicherheitszuschläge für die Prämien angesetzt. Das oben skizzierte Ruinproblem wird hier dadurch vermieden, dass in die Rechnungsgrundlagen (Sterbetafeln, Zinsen, ...) Sicherheitszu- bzw. -abschläge eingearbeitet werden.

(ii) In der Praxis müssen auch noch Kapitalgewinne berücksichtigt werden. Tritt wie üblich der Versicherungsnehmer mit seinen Prämienzahlungen in Vorleistung, so kann durchaus die Nettoprämie niedriger als $E(X)$ angesetzt werden, wenn zwischen dem Zeitpunkt der Prämienzahlung und dem Leistungszeitpunkt hinreichend hohe Kapitalgewinne aus der Prämie generiert werden können. Somit wäre es eigentlich angemessener, die Nettoprämie mit dem erwarteten *Barwert* der Versicherungsleistungen zu vergleichen. Da die Spanne zwischen Prämienzahlungszeitpunkt und dem Leistungszeitpunkt im Mittel üblicherweise deutlich kürzer als bei Lebensversicherungen ist (und daher die Verzinsungseffekte weniger gravierend), werden wir uns aber im Folgenden auf den vereinfachenden direkten Vergleich von Nettoprämie und Nettorisikoprämie beschränken. □

Es gibt nun zahlreiche Ansätze für die Berechnung des Sicherheitszuschlags. Dabei sollte dieser nur von der Verteilung der Zv. X abhängen (und nicht von der genauen Form der Zv. als Abbildung selber).

3.6 Definition *Sei $\mathcal{X}_H \subset \mathcal{X} := \{X \mid X \geqslant 0 \text{ Zv.}\}$ eine Menge von zulässigen Zv. Ein* **Prämienkalkulationsprinzip** *oder kurz* **Prämienprinzip** *für diese Menge ist eine Abbildung $H : \mathcal{X}_H \to [0, \infty)$ mit der Eigenschaft $H(X) = H(Y)$, falls $P^X = P^Y$.*
Risiken $X \in \mathcal{X} \backslash \mathcal{X}_H$ heißen **nicht versicherbar** *bei Verwendung des Prämienprinzips H.* □

3.7 Definition *(i)* **Nettorisikoprinzip***:*

$$\begin{aligned} \mathcal{X}_H &:= \{X \in \mathcal{X} \mid E(X) < \infty\} \\ H(X) &:= E(X) \end{aligned}$$

(ii) **Erwartungswertprinzip***:*

$$\begin{aligned} \mathcal{X}_H &:= \{X \in \mathcal{X} \mid E(X) < \infty\} \\ H(X) &:= (1 + \delta)E(X) \end{aligned}$$

für ein $\delta > 0$.

(iii) **Varianzprinzip***:*

$$\begin{aligned} \mathcal{X}_H &:= \{X \in \mathcal{X} \mid E(X^2) < \infty\} \\ H(X) &:= E(X) + \delta Var(X) \end{aligned}$$

für ein $\delta > 0$.

(iv) **Standardabweichungsprinzip***:*

$$\begin{aligned} \mathcal{X}_H &:= \{X \in \mathcal{X} \mid E(X^2) < \infty\} \\ H(X) &:= E(X) + \delta\sqrt{Var(X)} \end{aligned}$$

für ein $\delta > 0$.

(v) **Semistandardabweichungsprinzip***:*

$$\begin{aligned} \mathcal{X}_H &:= \{X \in \mathcal{X} \mid E(X^2) < \infty\} = \mathcal{X} \cap \mathcal{L}_2 \\ H(X) &:= E(X) + \delta\sqrt{E\big(((X - E(X))^+)^2} = E(X) + \delta\|(X - E(X))^+\|_2 \end{aligned}$$

für ein $\delta > 0$. Hierbei bezeichnet

$$\|Z\|_2 := \sqrt{E(Z^2)}$$

die übliche Seminorm auf $\mathcal{L}_2 := \{Z \mid E(Z^2) < \infty\}$.

(vi) **Perzentilprinzip***:*

$$\begin{aligned} \mathcal{X}_H &:= \mathcal{X} \\ H(X) &:= F_X^{\leftarrow}(1 - \varepsilon) \end{aligned}$$

für ein $\varepsilon \in (0, 1)$, wobei $F_X^{\leftarrow}$ die Quantilfunktion von X bezeichne.

(vii) **Maximalschadenprinzip**:

$$\begin{aligned} \mathcal{X}_H &:= \{X \in \mathcal{X} \mid F_X^{\leftarrow}(1) < \infty\} \\ H(X) &:= F_X^{\leftarrow}(1) \end{aligned}$$ □

3.8 Bemerkung (i) Wie oben dargelegt, ist das Nettorisikoprinzip aus Sicht des Versicherungsunternehmens unsinnig. Ebenso ist das Maximalschadenprinzip aus Sicht des Versicherungsnehmers sinnlos.

(ii) In den obigen Beispielen sind die Definitionsbereiche jeweils (in gewissem Sinne) maximal gewählt. Für die nicht versicherbaren Risiken wäre daher jeweils eine unendlich hohe Prämie zu verlangen.

(iii) Die Bezeichnung Prämienkalkulationsprinzip ist eigentlich etwas hoch gegriffen, da es sich schlicht um eine formelmäßige Beschreibung einer Berechnungsmethode für Prämien handelt. Zugrunde liegende grundsätzliche Prinzipien sind bei den in Definition 3.7 genannten Beispielen nicht zu erkennen. Dagegen werden wir in Definition 3.12 ein Beispiel für eine Klasse von Prämienprinzipien kennen lernen, bei denen wenigstens im Ansatz ein zugrunde liegendes ökonomisches "Prinzip" erkennbar ist. □

Es gibt keine gut zu begründenden Gütekriterien für Prämienprinzipien, wohl aber gewisse wünschenswerte Eigenschaften solcher Prinzipien.

3.9 Definition *Sei stets H ein Prämienkalkulationsprinzip für eine Menge $\mathcal{X}_H$ von zulässigen Zv. Dann heißt H*

(i) **erwartungswertübersteigend**, *falls*

$$H(X) \geqslant E(X) \qquad \forall\, X \in \mathcal{X}_H.$$

(ii) **maximalschadenbegrenzt**, *falls*

$$H(X) \leqslant F_X^{\leftarrow}(1) \qquad \forall\, X \in \mathcal{X}_H.$$

(iii) **monoton**, *falls*

$$H(X) \leqslant H(Y) \qquad \forall\, X, Y \in \mathcal{X}_H \text{ mit } X \leqslant Y \ P-f.s.$$

(iv) **translationsinvariant** *(oder* **konsistent***), falls*

$$H(X + c) = H(X) + c \qquad \forall\, X \in \mathcal{X}_H, c > 0 \text{ mit } X + c \in \mathcal{X}_H.$$

(v) **skaleninvariant** *(oder* **homogen***), falls*

$$H(cX) = cH(X) \qquad \forall\, X \in \mathcal{X}_H, c > 0 \text{ mit } cX \in \mathcal{X}_H.$$

(vi) **subadditiv**, *falls*

$$H(X + Y) \leqslant H(X) + H(Y) \qquad \forall\, X, Y \in \mathcal{X}_H \text{ mit } X + Y \in \mathcal{X}_H.$$

(vii) **additiv**, *falls*

$$H(X+Y) = H(X)+H(Y) \quad \textit{für alle unabhängigen } X, Y \in \mathcal{X}_H \textit{ mit } X+Y \in \mathcal{X}_H. \quad \square$$

3.10 Bemerkung (i) Ist H maximalschadenbegrenzt, so sagt man auch, dass kein "rip off" statt findet.

(ii) Statt von Translationsinvarianz bzw. Skaleninvarianz sollte eigentlich besser von Translations*äqui*varianz bzw. Skalen*äqui*varianz gesprochen werden, da ja die Prämie bei einer Translation bzw. Skalierung der Schadenhöhe in gleicher Weise transformiert wird, nicht aber unverändert bleibt. Wir halten uns hier aber an die in der Literatur übliche Terminologie.

(iii) Die Translationsinvarianz sichert, dass ein deterministischer Kapitaltransfer fair für Versicherungsunternehmen und Versicherungsnehmer ist.

(iv) Die Skaleninvarianz sorgt dafür, dass die Prämie unabhängig von der verwendeten Geldeinheit ist. Diese Motivation ist allerdings schon daher nur bedingt überzeugend, weil es keinen zwingenden Grund gibt, bei Verwendung unterschiedlicher Geldeinheiten das Prämienprinzip beizubehalten. (Außerdem kann die Übernahme sehr hoher Risiken für das Versicherungsunternehmen ungünstig sein, wenn keine hinreichenden Kapitalreserven zur Verfügung stehen; daher mag aus betriebswirtschaftlicher Sicht ein Prämienprinzip sinnvoll sein, bei dem auch $H(cX) > cH(X)$ für $c > 1$ möglich ist. Solche unternehmensabhängigen Überlegungen werden im folgenden immer vernachlässigt.)

(v) Die Subadditivität soll sichern, dass der Versicherungsnehmer nicht dadurch seine Prämie senken kann, dass er ein zu versicherndes Risiko "künstlich" aufteilt und die Teilrisiken getrennt versichert.

Im übrigen ist damit zu rechnen, dass sich die zufälligen Schwankungen der Risiken X und Y teilweise ausgleichen, so dass es plausibel ist, dass der Sicherheitszuschlag $H(X+Y) - E(X+Y)$ des Gesamtrisikos zumindest nicht größer als die Summe der einzelnen Sicherheitszuschläge sein sollte, was gerade die Subadditivität impliziert. Um diese Überlegungen mathematisch zu präzisieren, sei der Einfachheit halber angenommen, dass X und Y die gleiche Verteilung besitzen. Man kann dann zeigen, dass $E(\phi(X+Y)) \leqslant E(\phi(2X))$ für *alle* konvexen Funktionen $\phi : \mathbb{R} \to \mathbb{R}$; insbesondere ist $E(|X+Y-2E(X)|^\alpha) \leqslant E(|2X-2E(X)|^\alpha)$ und $E(((X+Y-2E(X))^+)^\alpha) \leqslant E(((2X-2E(X))^+)^\alpha)$ für alle $\alpha \geqslant 1$. Man wird daher die Schadenhöhenzv. $2X$ als riskanter als die Schadenhöhe $X+Y$ ansehen und eine wenigstens ebenso hohe Prämie für die Versicherung dieses Risikos verlangen. Setzt man die Skaleninvarianz voraus, so würde man für $2X$ gerade die Prämie $2H(X)$ verlangen, für die Versicherung des Risikos $X+Y$ sollte man daher eine Prämie $H(X+Y)$ verlangen, die maximal $2H(X) = H(X) + H(Y)$ beträgt.

Diese Begründung setzt allerdings die nicht völlig unkritische Skaleninvarianz voraus. Eine etwas schwächere Forderung als die Subadditivität, die vom gedanklichen Ansatz her nicht mit der Skaleninvarianz verknüpft ist, werden wir im Absatz über Risikomaße kennen lernen.

(vi) Die Additivität schließlich wird dadurch motiviert, dass sich die Sicherheitszuschläge verschiedener unabhängiger Versicherungsnehmer addieren sollen, ist aber aus ökonomischer und stochastischer Sichtweise sicherlich fragwürdig, da in größeren Portfolios von unabhängigen Risiken ein besserer Risikoausgleich statt findet und somit auch niedrigere Sicherheitszuschläge erhoben werden können. (Die obige Argumentation legt in der Tat nahe, dass $H(X) + H(Y)$ eher dann eine angemessene Prämie für $X + Y$ ist, wenn X und Y komonoton sind, also z.B. Y als monoton steigende Funktion von X darstellbar ist.) □

3.11 Satz *Die in Definition 3.7 genannten Prämienprinzipien erfüllen stets die in der folgenden Tabelle mit ✓ markierten Eigenschaften aus Definition 3.9, während die übrigen Eigenschaften i.allg. nicht erfüllt sein müssen:*

	erw.wertüberst.	max.schadenbegr.	monoton	translationsinv.	skaleninv.	subadditiv	additiv
Nettorisikoprinzip	✓	✓	✓	✓	✓	✓	✓
Erwartungswertprinzip	✓	(1)	✓		✓	✓	✓
Varianzprinzip	✓	(2)	(2)	✓		(3)	✓
Standardabweichungspr.	✓	(4)	(4)	✓	✓	✓(5)	(6)
Semistandardabw.pr. ($\delta \leqslant 1$)	✓	✓(8)	✓(8)	✓	✓	✓(7)	(9)
Perzentilprinzip	(10)	✓	✓	✓	✓	(11)	(11)
Maximalschadenprinzip	✓	✓	✓	✓	✓	✓	✓

□

Beweis. Die nicht mit einer Nummer versehenen Einträge sind offensichtlich.

(1) Ist $X = a > 0$ P-f.s., so gilt $H(X) = (1 + \delta)a > F_X^{\leftarrow}(1) = a$.

(2) Ist $P\{X = 1 + b\} = \alpha$ und $P\{X = 1\} = 1 - \alpha$ für ein $\alpha \in (0, 1)$ und ein $b > 0$, so gilt $H(X) = 1 + \alpha b + \delta\alpha(1 - \alpha)b^2$ und $F_X^{\leftarrow}(1) = 1 + b$. Für hinreichend großes b gilt offensichtlich $H(X) > F_X^{\leftarrow}(1)$.

(3) Für $X = Y$ mit $Var(X) > 0$ erhält man $H(X + Y) = H(2X) = 2E(X) + 4\delta Var(X) > 2E(X) + 2\delta Var(X) = H(X) + H(Y)$.

(4) Für die unter (2) geschilderte Verteilung P^X ist nun $H(X) = 1+\alpha b+\delta\sqrt{\alpha(1 - \alpha)}b$. Für $\alpha = (1 + 1/\sqrt{1 + \delta^2})/2$ zeigen direkte Rechnungen, dass $\alpha + \delta\sqrt{\alpha(1 - \alpha)} = (1 + \sqrt{1 + \delta^2})/2 > 1$ und daher $H(X) > F_X^{\leftarrow}(1)$.

(5) Die Cauchy-Schwarz'sche Ungleichung liefert $Cov(X, Y) \leqslant \sqrt{Var(X)Var(Y)}$. Somit gilt $Var(X + Y) \leqslant Var(X) + Var(Y) + 2\sqrt{Var(X)Var(Y)} = \big(\sqrt{Var(X)} + \sqrt{Var(Y)}\big)^2$, woraus die Behauptung sofort folgt. (Die letzte Ungleichung ist übrigens i.W. gerade die Dreiecksungleichung im Raum $\mathcal{L}_2$ der quadratintegrablen Zv.)

(6) folgt aus $\sqrt{Var(X+Y)} = \sqrt{Var(X)+Var(Y)} < \sqrt{Var(X)} + \sqrt{Var(Y)}$, falls X und Y stochastisch unabhängig sind mit $Var(X), Var(Y) > 0$.

(7) folgt aus $(u+v)^+ \leqslant u^+ + v^+$ und der Dreiecksungleichung für die $\mathcal{L}_2$-Seminorm:

$$\begin{aligned} H(X+Y) &\leqslant E(X+Y) + \delta\|(X-E(X))^+ + (Y-E(Y))^+\|_2 \\ &\leqslant E(X) + \delta\|(X-E(X))^+\|_2 + E(Y) + \delta\|(Y-E(Y))^+\|_2 \\ &= H(X) + H(Y). \end{aligned}$$

(8) Ist $X \leqslant Y$ P-f.s., so gilt $(X-Y-E(X-Y))^+ \leqslant E(Y-X)$ P-f.s. Daher folgt wie in (7) mit der Dreiecksungleichung für die $\mathcal{L}_2$-Seminorm

$$\begin{aligned} H(X) &= E(X) + \delta\|(Y-E(Y) + ((X-Y)-E(X-Y)))^+\|_2 \\ &\leqslant E(X) + \delta\|(Y-E(Y))^+\|_2 + \delta\|((X-Y)-E(X-Y))^+\|_2 \\ &\leqslant E(X) + \delta E(Y-X) + \delta\|(Y-E(Y))^+\|_2 \\ &\leqslant H(Y) \end{aligned}$$

für $\delta \in (0,1]$. Insbesondere folgt für $Y = F_X^{\leftarrow}(1)$, dass H maximalschadenbegrenzt ist, da $H(Y) = F_X^{\leftarrow}(1)$.

(9) Seien X und Y unabhängige $\mathcal{B}_{(1,1/2)}$-verteilte Zufallsvariablen. Dann ist $\|(X-E(X))^+\|_2 = \sqrt{P\{X=1\}/4} = 1/(2\sqrt{2})$ und somit $H(X) = H(Y) = 1/2 + \delta/(2\sqrt{2})$. Ferner ist $X+Y$ $\mathcal{B}_{(2,1/2)}$-verteilt und daher $\|(X+Y-E(X+Y))^+\|_2 = \sqrt{P\{X+Y=2\}} = 1/2$, so dass $H(X+Y) = 1 + \delta/2 < H(X) + H(Y)$.

(10) Es sei $P\{X=1\} = 1-\varepsilon/2$ und $P\{X=2\} = \varepsilon/2$. Dann ist $F_X^{\leftarrow}(1-\varepsilon) = 1 < E(X)$.

(11) Es seien X, Y unabhängige Zv. mit $P\{X=1\} = P\{Y=1\} = 1-\tau$ und $P\{X=2\} = P\{Y=2\} = \tau$ für ein $\tau < \varepsilon$ mit $2\tau - \tau^2 > \varepsilon$. (Ein solches τ existiert stets, da $\lim_{\tau\uparrow\varepsilon} 2\tau - \tau^2 = 2\varepsilon - \varepsilon^2 > \varepsilon$ für $\varepsilon \in (0,1)$.) Dann gilt $F_X^{\leftarrow}(1-\varepsilon) = F_Y^{\leftarrow}(1-\varepsilon) = 1$, also $H(X) + H(Y) = 2$, aber $H(X+Y) = F_{X+Y}^{\leftarrow}(1-\varepsilon) > 2$, da $P\{X+Y \leqslant 2\} = P\{X=1\}\cdot P\{Y=1\} = (1-\tau)^2 = 1-(2\tau-\tau^2) < 1-\varepsilon$. □

Man sieht also, dass nur die ökonomisch unsinnigen Prämienprinzipien (d.h. das Nettorisikoprinzip und das Maximalschadenprinzip) alle Bedingungen erfüllen. Das Semistandardabweichungsprinzip (und analoge Prämienprinzipien, bei denen die $\mathcal{L}_2$-Seminorm durch die $\mathcal{L}_p$-Seminorm $\|Z\|_p := (E(Z^p))^{1/p}$ für ein $p > 1$ ersetzt wird) erfüllt im Fall $\delta \leqslant 1$ alle Bedingungen bis auf die (ohnehin problematisch zu sehende) Additivität. Alle weiteren hier diskutierten Prämienprinzipien verletzen dagegen noch wenigstens eine weitere Bedingung. (Etwas wird das letztgenannte negative Resultat allerdings dadurch relativiert, dass für "sinnvolle" Werte für δ bzw. ε und "realistische" Verteilungen P^X die Prämienprinzipien oft doch erwartungswertübersteigend, maximalschadenbegrenzt, monoton und subadditiv sind.)

Einen etwas systematischeren, ökonomisch begründeten Ansatz liefert das Nullnutzenprinzip, das besagt, dass die Prämie dann fair ist, wenn der Nutzen, den das Versicherungsunternehmen den saldierten Zahlungen zuweist und der mit Hilfe einer sog. Nutzenfunktion gemessen wird, im Mittel gleich 0 ist.

3.12 Definition *(i) Eine* **(Standard-)Nutzenfunktion** *ist eine zweimal differenzierbare Abbildung* $u : \mathbb{R} \to \mathbb{R}$ *mit*

$$u'(x) \geqslant 0, \quad u''(x) \leqslant 0 \quad \forall\, x \in \mathbb{R}, \qquad u(0) = 0, \quad u'(0) = 1.$$

(ii) $H : \mathcal{X}_H \to [0, \infty)$ *heißt ein* **Nullnutzenprinzip** *zur Nutzenfunktion* u*, falls*

$$E(u(H(X) - X)) = 0 \qquad \forall\, X \in \mathcal{X}_H. \qquad \square$$

3.13 Bemerkung (i) $u(x)$ beschreibt den Nutzen, den das Versicherungsunternehmen dem Kapital x zumisst. Die geforderte Monotonie von u besagt also einfach, dass einem höheren Kapital ein größerer Wert beigemessen wird. Die Forderung $u''(x) \leqslant 0$, die insbesondere die Konkavität von u impliziert, besagt, dass der Nutzen*zuwachs* eines festen zusätzlichen positiven Betrags mit steigendem bereits vorhandenen Kapital abnimmt. Die beiden letzten Bedingungen stellen nur eine Normierung dar.

(ii) Das Nullnutzenprinzip besagt, dass der erwartete Nutzen des Versicherungsvertrags aus Sicht des Versicherungsunternehmens 0 ist.

(iii) Wenn u nicht streng monoton ist, so kann es mehrere Nullnutzenprinzipien zu vorgegebener Nutzenfunktion und vorgegebenem Definitionsbereich $\mathcal{X}_H$ geben. $\square$

3.14 Beispiel Für die *Exponentialnutzenfunktion*

$$u(x) = \frac{1}{\beta}(1 - \mathrm{e}^{-\beta x}), \quad x \in \mathbb{R},$$

mit $\beta > 0$ ergibt sich das zugehörige Nullnutzenprinzip durch die Bedingung

$$\begin{aligned} & E(u(H(X) - X)) = \frac{1}{\beta}\big(1 - E(\mathrm{e}^{\beta X})\mathrm{e}^{-\beta H(X)}\big) = 0 \\ \iff\quad & \psi_X(\beta) = E(\mathrm{e}^{\beta X}) = \mathrm{e}^{\beta H(X)} \\ \iff\quad & H(X) = \frac{1}{\beta}\log \psi_X(\beta), \end{aligned}$$

vorausgesetzt dass $\psi_X(\beta)$ wohldefiniert ist. (Formal heißt dies, dass wir H nur auf $\mathcal{X}_H := \{X \in \mathcal{X} \mid E(\mathrm{e}^{\beta X}) < \infty\}$ definieren können.) H heißt dann auch *Exponentialprinzip.* $\square$

Unter schwachen Zusatzannahmen, sind gewisse der Eigenschaften aus Definition 3.9 immer bzw. nie erfüllt.

3.15 Satz *(i) Ist u streng monoton steigend, so ist ein zugehöriges Nullnutzenprinzip stets erwartungswertübersteigend, maximalschadenbegrenzt, monoton und translationsinvariant.*

(ii) Ist u strikt konkav, so ist ein zugehöriges Nullnutzenprinzip H nicht skaleninvariant und nicht subadditiv, falls es $X \in \mathcal{X}_H$ mit $P\{X \neq H(X)\} > 0$ gibt und ein $c \in (0, 1)$, so dass $cX, (1 - c)X \in \mathcal{X}_H$.

(iii) *Enthält $\mathcal{X}_H$ (wenigstens) alle nicht-negativen Zufallsvariablen, die nur maximal vier verschiedene Werte annehmen, so sind das Exponentialprinzip und das Nettorisikoprinzip die einzigen additiven Nullnutzenprinzipien zu einer streng monoton steigenden, zweimal stetig differenzierbaren Nutzenfunktion.* □

Beweis. (i) Da u konkav ist, ist $-u$ konvex. Die Jensensche Ungleichung liefert daher

$$\begin{aligned} & 0 = E\big(-u(H(X)-X)\big) \geqslant -u\big(E(H(X)-X)\big) \\ \Longrightarrow\ & u(H(X)-E(X)) \geqslant 0 \\ \Longrightarrow\ & H(X) \geqslant E(X), \end{aligned}$$

da u streng monoton steigend ist mit $u(0)=0$, also nicht-negative Funktionswerte nur für nicht-negative Argumente angenommen werden.

Für den Nachweis, dass H maximalschadenbegrenzt ist, nehmen wir an, dass $H(X) > F_X^{\leftarrow}(1)$ für ein $X \in \mathcal{X}_H$ gilt, d.h. $H(X) > X$ P-f.s. Dann folgt auch

$$E\big(u(H(X)-X)\big) > E(u(0)) = 0$$

im Widerspruch zur Definition von H.

Die Monotonie von H ergibt sich wiederum aus der strengen Monotonie von u. Gäbe es nämlich Zv. $X, Y \in \mathcal{X}_H$ mit $X \leqslant Y$ P-f.s., aber $H(X) > H(Y)$, so wäre

$$0 = E(u(H(X)-X)) > E(u(H(Y)-Y)) = 0.$$

Schließlich gilt nach Definition von H

$$E\big(u((H(X)+c)-(X+c))\big) = E(u(H(X)-X)) = 0,$$

d.h. $H(X)+c$ ist eine Nullstelle von $a \mapsto E(u(a-(X+c)))$. Da mit u auch diese Funktion von a streng monoton ist, ist dies die einzige Nullstelle, so dass nach Definition von H die Behauptung $H(X+c) = H(X)+c$ folgt.

(ii) Es sei nun u als strikt konkav vorausgesetzt. Insbesondere gilt

$$u(cy) > u(0) + c(u(y)-u(0)) = cu(y)$$

für alle $0 < c < 1$ und $y \neq 0$. Es folgt für $X \in \mathcal{X}_H$ mit $P\{X \neq H(X)\} > 0$, $0 < c < 1$ mit $cX \in \mathcal{X}_H$

$$E\big(u(cH(X)-cX)\big) > cE(u(H(X)-X)) = 0 = E(u(H(cX)-cX))$$

und somit wegen der Monotonie von u auch $cH(X) > H(cX)$. Folglich ist H nicht skaleninvariant.

Außerdem folgt $H(X) = cH(X) + (1-c)H(X) > H(cX) + H((1-c)X)$, d.h. H ist auch nicht subadditiv.

(iii) Das Erwartungswertprinzip ist ein Nullnutzenprinzip zur Nutzenfunktion $u(x) = x$ und additiv. Die Additivität des Exponentialprinzip ergibt sich sofort aus Satz 2.7.

Für den Nachweis, dass dies die einzigen additiven Nullnutzenprinzipien sind, betrachte unabhängige Zv. X und Y mit $P\{X = 0\} = 1 - \alpha$ und $P\{X = x\} = \alpha$, wobei $\alpha = u(c)/(u(c) - u(c-x))$ für gewisse $0 < c < x$, und $P\{Y = 0\} = 1 - \beta$ und $P\{Y = 2y\} = \beta$, wobei $\beta = u(y)/(u(y) - u(-y))$. Dann sind wegen $E(u(c - X)) = E(u(y - Y)) = 0$ die eindeutigen Prämien gemäß dem Nullnutzenprinzip gerade $H(X) = c$ und $H(Y) = y$. Ist H nun additiv, so muss für

$$X + Y = \begin{cases} 0 & & (1-\alpha)(1-\beta) \\ x & \text{mit W.} & \alpha(1-\beta) \\ 2y & & \beta(1-\alpha) \\ x + 2y & & \alpha\beta \end{cases}$$

die Identität $E\big(u(c + y - (X + Y))\big) = 0$ gelten.

Wir betrachten nun zunächst den Fall $c = x/2 = y$, so dass $\alpha = \beta = u(x/2)/(u(x/2) - u(-x/2))$. Direkte Rechnungen liefern die Bedingung

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

Die Funktion $g(x) := \log(u(x)/(-u(-x)))$ erfüllt daher die Gleichung $g(x) = 2g(x/2)$ für alle $x > 0$. Sie ist zweimal stetig differenzierbar mit Ableitung

$$g'(x) = \frac{u'(x)}{u(x)} + \frac{u'(-x)}{u(-x)}.$$

Eine Taylor-Entwicklung von Zähler und Nenner in 0 zeigt, dass für geeignete $x_i \in (0, x)$

$$\begin{aligned} g'(x) &= \frac{1 + xu''(x_1)}{x + u''(x_2)x^2/2} + \frac{1 - xu''(x_3)}{-x + x^2u''(x_4)/2} \\ &= \frac{-x^2(u''(x_1) - u''(x_4)/2 + u''(x_3) - u''(x_2)/2) + O(x^3)}{-x^2 + O(x^3)} \\ &\longrightarrow u''(0) =: a \end{aligned}$$

für $x \downarrow 0$. Es folgt wegen $g'(x) = g'(x/2)$ induktiv $g'(x) = g'(2^{-k}x) \to a$ für $k \to \infty$ für alle $x > 0$, d.h. $g'(x) = a$ für alle $x > 0$. Es folgt $g(x) = ax + b$, wobei wegen $g(x) = 2g(x/2)$ außerdem $b = 0$ gelten muss, also schließlich

$$u(-x) = -\mathrm{e}^{-ax}u(x)$$

für alle $x > 0$.

Für beliebige $0 < c < x$ und $y > 0$ folgt nun wiederum aus der Bedingung $E\big(u(c + y - (X + Y))\big) = 0$ mit direkter Rechnung

$$u(c - x)\big(u(c + y)\mathrm{e}^{-ay} + u(c - y)\big) = u(c)\big(u(c - x + y)\mathrm{e}^{-ay} + u(c - x - y)\big).$$

Leitet man nun zweimal nach y ab und lässt dann y gegen 0 konvergieren, so folgt

$$u(c-x)\Big(2u''(c) - 2au'(c) + a^2u(c)\Big) = u(c)\Big(2u''(c-x) - 2au'(c-x) + a^2u(c-x)\Big),$$

d.h. die Funktion $z \mapsto (2u''(z) + 2au'(z) + a^2u(z))/u(z)$ ist konstant, bzw. äquivalent dazu

$$u''(z) - au'(z) + du(z) = 0$$

für alle $z \in \mathbb{R}$ und eine geeignete Konstante d. Um die Lösungen dieser homogenen linearen Differentialgleichung mit konstanten Koeffizienten zu ermitteln, betrachtet man am einfachsten die Nullstellen des zugehörigen charakteristischen Polynoms $h(t) = t^2 - at + d$, also $t_\pm = a/2 \pm \sqrt{a^2/4 - d}$ (mit $\sqrt{x} := i\sqrt{|x|}$ für $x < 0$). Es sind 3 Fälle zu unterscheiden:

1. Fall: $a^2/4 > d \implies t_\pm \in \mathbb{R}$, $t_+ > t_-$
   In dem Fall hat die allgemeine Lösung der obigen Differentialgleichung die Form $u(z) = \lambda_+ e^{t_+z} + \lambda_- e^{t_-z}$. Die Normierungsbedingung $u(0) = 0$ impliziert $\lambda_- = -\lambda_+$, die Bedingung $u'(0) = 1$ daher $\lambda_+ = 1/(t_+ - t_-) > 0$. Ist $t_+ \neq 0$, so ist $u''(z) = \lambda_+(t_+^2 e^{t_+z} - t_-^2 e^{t_-z})$ für hinreichend große Werte von z sicherlich strikt positiv, im Widerspruch zur Definition einer Standardnutzenfunktion. Es muss also $t_+ = 0$ gelten und u ist eine exponentielle Nutzenfunktion.

2. Fall: $a^2/4 < d \implies t_\pm \in \mathbb{C}\backslash\mathbb{R}$
   In dem Fall sind t_+ und t_- konjugiert komplex mit Imaginär- und Realteil $\pm s$ und r und die allgemeine Lösung der obigen Differentialgleichung hat die Form $u(z) = \lambda_1 e^{rz}\cos(sz) + \lambda_2 e^{rz}\sin(sz)$. Aus $u(0) = 0$ folgt $\lambda_1 = 0$. Man sieht nun unmittelbar, dass diese Funktion nicht streng monoton sein kann, im Widerspruch zur Annahme.

3. Fall: $a^2/4 = d \implies t_+ = t_- \in \mathbb{R}$
   In dem Fall lautet die allgemeine Lösung der obigen Differentialgleichung $u(z) = (\lambda_1 + \lambda_2 z)e^{t_+z}$, wegen der Normierungsbedingungen also $u(z) = ze^{t_+z}$. Diese Funktion ist aber nur dann streng monoton steigend und konkav, wenn $t_+ = 0$, was gerade das Nettorisikoprinzip liefert. □

3.16 Bemerkung Die im Teil (ii) des Beweises hergeleitete Beziehung $H(X+Y) > H(X) + H(Y)$ für gewisse $X, Y \in \mathcal{X}_H$ impliziert, dass bei der Aggregation von Risiken der Sicherheitszuschlag des Gesamtrisikos größer ist als die Summe der einzelnen Sicherheitszuschläge, obwohl in größeren Portfolios i.d.R. ein besserer Risikoausgleich erfolgt. (Ein Stetigkeitsargument zeigt, dass die Ungleichung $H(X+Y) > H(X) + H(Y)$ nicht nur für die im Beweis verwendeten total abhängigen Zv. vom Typ $X = cZ$ und $Y = (1-c)Z$ gilt, sondern auch für gewisse positiv abhängige Zv., die einen Risikoausgleich erlauben.) Offensichtlich ist dieses wenig plausible Verhalten durch die strikte Konkavität von u bedingt. Betrachtet man umgekehrt konvexe Nutzenfunktionen, so würde $H(X+Y) \leqslant H(X) + H(Y)$ für diese Zv. folgern. Allerdings würde die Konvexität andererseits implizieren, dass das Versicherungsunternehmen risikobevorzugend ist, es also einen zufälligen Gewinn einem sicheren Gewinn in Höhe des Erwartungswerts vorziehen würde, was unplausibel ist. Es gälte dann nämlich wegen der Jensenschen Ungleichung

$E(u(X)) \geqslant u(E(X))$ für jede Zv. $X \geqslant 0$ mit $E(X), E(u(X)) < \infty$, d.h. im Mittel hätte X einen höheren Nutzen als $E(X)$. Die hier geforderte Konkavität von u sichert dagegen die üblicherweise unterstellte Risikoaversion des Versicherungsunternehmens, d.h. ein sicherer Gewinn wird einem zufälligen Gewinn mit gleichem Erwartungswert vorgezogen.

Im Abschnitt über Risikomaße werden wir jedoch sehen, dass Nullnutzenprinzipien oft eine abgeschwächte Bedingung erfüllen, die ebenfalls Diversifikationseffekte berücksichtigt. □

Als Fazit ergibt sich also, dass die zunächst theoretisch fundierter erscheinenden Nullnutzenprinzipien ebenfalls nicht alle wünschenswerten Eigenschaften erfüllen.

Eine wichtige Anwendung von Prämienprinzipien ist die Aufteilung des für das Portfolio als ausreichend erachteten Sicherheitszuschlags auf die einzelnen Policen. Sei dazu im folgenden $S_n = \sum_{i=1}^n X_i$ der Gesamtschaden eines Portfolios aus n Policen mit Einzelschadenhöhen X_i, die i.allg. weder unabhängig noch identisch verteilt sein müssen.

Sind die X_i unabhängig und wird für die Berechnung der Prämie für das Portfolio ein additives Prämienprinzip H verwendet, so gilt $H(S_n) = \sum_{i=1}^n H(X_i)$ und die Summe der Sicherheitszuschläge, die sich für die einzelnen Policen bei Verwendung des gleichen Prämienprinzips ergeben, ist gerade gleich dem Sicherheitszuschlag für das Gesamtportfolio. Sind die X_i dagegen nicht unabhängig, so gilt bei Subadditivität i.allg. nur $H(S_n) \leqslant \sum_{i=1}^n H(X_i)$, d.h. die Summe der einzelnen Sicherheitszuschläge ist i.allg. zu groß, wenn man auch für die Berechnung der Prämien der einzelnen Policen das Prinzip H anwendet. Die einzelnen Prämien wären in dem Fall also zu hoch und das angebotene Versicherungsprodukt daher eventuell nicht konkurrenzfähig.

Nun gibt es keinen zwingenden Grund, für die Berechnung der Prämie des Gesamtportfolios und der Prämien für die einzelnen Policen das gleiche Prämienprinzip anzuwenden. Bei dem folgenden alternativen Ansatz wird der Sicherheitszuschlag $\Delta := H(S_n) - E(S_n)$ des Portfolios proportional zu dem Anteil aufgeteilt, den das Einzelrisiko zur Gesamtvarianz beisteuert.

3.17 Definition *In der oben geschilderten Situation sei $Var(X_i) < \infty$ für alle $1 \leqslant i \leqslant n$ und $Var(S_n) > 0$. Die Prämien B_i der Einzelrisiken werden nach dem* **Kovarianzprinzip** *berechnet, falls mit $\Delta := H(S_n) - E(S_n)$ gilt*

$$B_i = E(X_i) + \frac{Cov(X_i, S_n)}{Var(S_n)}\Delta \qquad \forall\, 1 \leqslant i \leqslant n.$$ □

3.18 Bemerkung (i) Das Kovarianzprinzip ist *kein* Prämienprinzip im Sinne der Definition 3.6, da B_i nicht nur von P^{X_i}, sondern von $P^{(X_i,S_n)}$ abhängt.

(ii) Das Kovarianzprinzip definiert in der Tat eine Aufteilung der Gesamtprämie $H(S_n)$, denn

$$\begin{aligned}\sum_{i=1}^n B_i &= \sum_{i=1}^n E(X_i) + \sum_{i=1}^n Cov(X_i, S_n)\frac{\Delta}{Var(S_n)}\\ &= E(S_n) + Cov(S_n, S_n)\frac{\Delta}{Var(S_n)}\\ &= H(S_n).\end{aligned}$$

(iii) Da beim Kovarianzprinzip der Beitrag, den die einzelne Police zur Gesamtvarianz beisteuert, als Maßstab für die Bewertung des mit dieser Police verbundenen Schwankungsrisikos innerhalb des Portfolios verwendet wird, ist es am ehesten dann sinnvoll, das Kovarianzprinzip zu benutzen, wenn der Sicherheitszuschlag Δ für das Portfolio als Funktion von $Var(S_n)$ berechnet wurde.

(iv) Im Fall unkorrelierter Schadenhöhen X_i wird wegen $Cov(X_i, S_n) = Var(X_i)$ die Prämie B_i formal nach dem Varianzprinzip berechnet, wobei allerdings der Faktor $\delta = \Delta / Var(S_n)$ von P^{S_n} abhängt.

(v) Eine Einzelprämie B_i kann auch kleiner als $E(X_i)$ sein. Insbesondere ist das der Fall, wenn $Cov(X_i, S_n) < 0$, d.h. das entsprechende Risiko mit dem Gesamtrisiko des Portfolios negativ korreliert. In dem Fall verringert X_i in der Tat die Gesamtvarianz, da für die Varianz der Schadenhöhe des Portfolios ohne das Risiko X_i gilt

$$Var(S_n - X_i) = Var(S_n) + Var(X_i) - 2Cov(S_n, X_i) > Var(S_n). \qquad \square$$

Ein gravierender Nachteil des Kovarianzprinzips ist, dass i.d.R. jede einzelne Prämie nur innerhalb eines festen Portfolios zu kalkulieren ist. Jeder Ab- oder Zugang von bzw. zu dem Portfolio hat nämlich i.d.R. zur Folge, dass *alle* anderen Prämien neu zu berechnen sind. Das Kovarianzprinzip kann daher in der Praxis nur als *Richtschnur* für die Entwicklung neuer Tarife (auf der Basis von Modellportfolios) oder bei der notwendigen Modifikation vorhandener Tarife (z.B. aus Anlass einer Tariferhöhung) dienen.

Da die Versicherungsprämie das vom Versicherungsunternehmen übernommene Risiko "angemessen" quantifizieren sollte, lässt sie sich auf natürliche Weise auch als Risikomaß interpretieren.

Risikomaße

In diesem Abschnitt soll ein systematischer Zugang zur Bewertung von Risiken vorgestellt werden. Dabei soll die Sichtweise einer Aufsichtsbehörde (oder eines internen Risikokontrolleurs) eingenommen werden, die ein (minimales) Risikokapital bestimmen will, das zur Absicherung eines vom Versicherungsunternehmen eingegangenen Risikos zur Verfügung stehen muss. Im folgenden verstehen wir unter einem Risiko eine beliebige Abbildung $X : \Omega \to \mathbb{R}$, wobei Ω als Menge möglicher Szenarien interpretiert werden kann. Wir gehen also (anders als im ersten Teil dieses Kapitels) zunächst *nicht* mehr von einem vorgegebenen Wahrscheinlichkeitsraum aus und lassen auch zu, dass das Risiko negative Werte annimmt. Mit diesem erweiterten Risikobegriff lassen sich insbesondere auch Verluste (und Gewinne) aus Finanzanlagen mit behandeln, die sowohl positiv als auch negativ ausfallen können. Außerdem wird bei diesem Zugang nicht vorausgesetzt, dass der Regulator eine realistische Einschätzung des Zufallsgeschehens (in Form einer Verteilung P^X) geben kann.

Zunächst ist es nicht selbstverständlich, dass überhaupt eine einzige Kennzahl geeignet ist, ein vom Versicherungsunternehmen übernommenes Risiko zu bewerten. So ist es nicht unmittelbar einsichtig, warum zwei Risiken X und Y vergleichbar sein sollen, wenn etwa X mit hoher Wahrscheinlichkeit größer ist als Y, das zweite Risiko aber mit kleiner Wahrscheinlichkeit einen viel größeren Wert annimmt. Zur Motivation reeller Maßzahlen zur Quantifizierung von Risiken sei angenommen, dass der Regulator eine Menge akzeptabler Risiken vorgibt. Im folgenden gehen wir stets von der folgenden Situation aus:

3.19 (Generalvoraussetzung) Sei die Menge $\mathfrak{X}$ aller möglichen Risiken ein konvexer Kegel in $\{X : \Omega \to \mathbb{R}\}$, d.h. für alle $X, Y \in \mathfrak{X}$ und alle $a, b > 0$ gilt $aX + bY \in \mathfrak{X}$. Die Menge $\mathfrak{X}$ enthalte insbesondere alle konstanten Abbildungen; diese werden wir der Einfachheit halber mit den konstanten Werten identifizieren. □

Die Menge $\mathfrak{X}$ kann z.B. die Menge aller reellwertigen Abbildungen auf Ω sein oder weitere Einschränkungen vorsehen, z.B. dass nur beschränkte Risiken betrachtet werden sollen.

3.20 Definition *Eine Menge* $\mathfrak{A} \subset \mathfrak{X}$ *heißt* **Akzeptanzmenge***, falls Konstanten* $a \in \mathfrak{A}$ *und* $b \in \mathfrak{X} \backslash \mathfrak{A}$ *existieren und gilt*

$$X \in \mathfrak{A}, Y \in \mathfrak{X} \text{ mit } Y \leqslant X \implies Y \in \mathfrak{A}.$$

□

Es gibt also akzeptable konstante Risiken, aber nicht beliebig große sichere Verluste sind akzeptabel. Ferner sind Risiken, die in jedem Szenario höchstens so groß sind wie ein festes akzeptables Risiko, ebenfalls akzeptabel.

Das benötigte Risikokapital wird nun als das minimale Kapital definiert, um das das Risiko reduziert werden muss, um es akzeptabel werden zu lassen:

3.21 Definition *Das* **zu einer gegebenen Akzeptanzmenge $\mathfrak{A}$ gehörige Risikomaß** $\rho_{\mathfrak{A}} : \mathfrak{X} \to [-\infty, \infty]$ *wird definiert durch*

$$\rho_{\mathfrak{A}}(X) := \inf\{t \in \mathbb{R} \mid X - t \in \mathfrak{A}\}.$$

□

Wir setzen im folgenden stets voraus, dass $\rho_{\mathfrak{A}}$ nur Werte in $\mathbb{R}$ annimmt. Dies ist insbesondere dann gesichert, wenn $\mathfrak{X}$ nur beschränkte Abbildungen enthält, da wegen $X - \sup X(\Omega) + a \leqslant a \in \mathfrak{A}$ und der Definition einer Akzeptanzmenge offensichtlich $\rho_{\mathfrak{A}}(X) \leqslant \sup X(\Omega) - a < \infty$ und ebenso $\rho_{\mathfrak{A}}(X) \geqslant \inf X(\Omega) - b > -\infty$. (Während die Einschränkung auf beschränkte Risiken aus ökonomischer Sicht sicherlich unproblematisch ist, werden damit allerdings Risiken ausgeschlossen, die in Kapitel 1 vorgestellten Schadenhöhenverteilungen besitzen.)

3.22 Satz und Definition *Ist das zu einer gegebenen Akzeptanzmenge $\mathfrak{A}$ gehörige Risikomaß $\rho_{\mathfrak{A}}$ $\mathbb{R}$-wertig, so besitzt es die folgenden Eigenschaften:*

(i) Monotonie: Für alle $X, Y \in \mathfrak{X}$ mit $X \leqslant Y$ gilt $\rho_{\mathfrak{A}}(X) \leqslant \rho_{\mathfrak{A}}(Y)$.

(ii) Translationsinvarianz: Für alle $X \in \mathfrak{X}$ und $c \in \mathbb{R}$ gilt $\rho_{\mathfrak{A}}(X + c) = \rho_{\mathfrak{A}}(X) + c$.

Eine Abbildung $\rho : \mathfrak{X} \to \mathbb{R}$, die diese Eigenschaften besitzt, wird **monetäres Risikomaß** *genannt.* □

Beweis. Die Monotonie folgt sofort aus der entsprechenden Eigenschaft der Akzeptanzmenge $\mathfrak{A}$, die Translationsinvarianz unmittelbar aus der Definition von $\rho_{\mathfrak{A}}$. □

Umgekehrt definiert jedes monetäre Risikomaß auch eine Akzeptanzmenge:

3.23 Satz und Definition *Sei ρ ein monetäres Risikomaß.*

(i) Durch

$$\mathfrak{A}_\rho := \{X \in \mathfrak{X} \mid \rho(X) \leqslant 0\}$$

wird eine Akzeptanzmenge definiert, die **zu ρ gehörende Akzeptanzmenge**.

(ii) Es gilt

$$\rho_{\mathfrak{A}_\rho} = \rho.$$

(iii) Enthält $\mathfrak{X}$ nur beschränkte Risiken, so ist ρ Lipschitz-stetig bzgl. der Supremumsnorm $\|X\|_\infty := \sup_{\omega \in \Omega} |X(\omega)|$ auf $\mathfrak{X}$, d.h. für alle $X, Y \in \mathfrak{X}$ gilt

$$|\rho(X) - \rho(Y)| \leqslant \|X - Y\|_\infty,$$

und die zu ρ gehörende Akzeptanzmenge besitzt die folgende Abgeschlossenheitseigenschaft:

Für alle $X \in \mathfrak{A}_\rho$ und alle $Y \in \mathfrak{X}$ ist $\left\{\lambda \in [0,1] \mid \lambda X + (1-\lambda)Y \in \mathfrak{A}_\rho\right\}$ abgeschlossen in $[0,1]$.

(iv) Für alle Akzeptanzmengen $\mathfrak{A}$ gilt

$$\mathfrak{A} \subset \mathfrak{A}_{\rho_{\mathfrak{A}}}.$$

Wenn $\mathfrak{A}$ die Abgeschlossenheitseigenschaft aus (iii) erfüllt, so gilt sogar

$$\mathfrak{A} = \mathfrak{A}_{\rho_{\mathfrak{A}}}.$$ □

Beweis. (i) Für $X \in \mathfrak{A}_\rho, Y \in \mathfrak{X}$ mit $Y \leqslant X$ gilt $\rho(Y) \leqslant \rho(X) \leqslant 0$, also $Y \in \mathfrak{A}_\rho$.

Ferner gilt wegen der Translationsinvarianz $\rho(c - \rho(0)) = \rho(0) + (c - \rho(0)) = c$ für alle $c \in \mathbb{R}$ und somit $-1 - \rho(0) \in \mathfrak{A}_\rho$ und $1 - \rho(0) \notin \mathfrak{A}_\rho$.

(ii) Aufgrund der Translationsinvarianz von ρ gilt für alle $X \in \mathfrak{X}$

$$\begin{aligned} \rho_{\mathfrak{A}_\rho}(X) &= \inf\{t \in \mathbb{R} \mid X - t \in \mathfrak{A}_\rho\} \\ &= \inf\{t \in \mathbb{R} \mid \rho(X) - t = \rho(X - t) \leqslant 0\} \\ &= \rho(X). \end{aligned}$$

(iii) Da $X \leqslant Y + \|X - Y\|_\infty$, folgt aus der Monotonie und Translationsinvarianz von ρ, dass $\rho(X) \leqslant \rho(Y + \|X - Y\|_\infty) = \rho(Y) + \|X - Y\|_\infty$. Vertauscht man die Rollen von X und Y, so erhält man ebenso $\rho(Y) \leqslant \rho(X) + \|X - Y\|_\infty$, zusammen also die behauptete Lipschitz-Stetigkeit.

Folglich ist die Abbildung $h : [0,1] \to \mathbb{R}$, $h(\lambda) := \rho(\lambda X + (1 - \lambda)Y)$ als Hintereinanderausführung zweier stetiger Funktionen stetig und die Menge

$$\begin{aligned} h^{-1}\big((-\infty, 0]\big) &= \big\{\lambda \in [0,1] \mid \rho(\lambda X + (1-\lambda)Y) \leqslant 0\big\} \\ &= \big\{\lambda \in [0,1] \mid \lambda X + (1-\lambda)Y \in \mathfrak{A}_\rho\big\} \end{aligned}$$

ist als Urbild einer abgeschlossenen Menge abgeschlossen in $[0,1]$.

(iv) Es gilt stets

$$\begin{aligned} \mathfrak{A}_{\rho_{\mathfrak{A}}} &= \big\{X \in \mathfrak{X} \mid \rho_{\mathfrak{A}}(X) = \inf\{t \in \mathbb{R} \mid X - t \in \mathfrak{A}\} \leqslant 0\big\} \\ &\supset \big\{X \in \mathfrak{X} \mid X \in \mathfrak{A}\big\} = \mathfrak{A}. \end{aligned}$$

Zum Nachweis der umgekehrten Inklusion beachte man, dass für alle $X \in \mathfrak{A}_{\rho_{\mathfrak{A}}}$ und alle $t > 0$ nach obiger Rechnung $X - t \in \mathfrak{A}$ gelten muss, d.h.

$$(1/2, 1] \subset \big\{\lambda \in [0,1] \mid \lambda(X - 1) + (1 - \lambda)(X + 1) = X + 1 - 2\lambda \in \mathfrak{A}\big\}.$$

Da aber die rechte Menge abgeschlossen in $[0,1]$ ist, liegt auch $1/2$ in dieser Menge, d.h. $X \in \mathfrak{A}$. □

Die Konzepte “Akzeptanzmengen” und “monetäre Risikomaße” sind noch zu allgemein gefasst, als dass tiefer liegende interessante Resultate für diese Größen hergeleitet werden könnten. Insbesondere spiegelt die Definition einer Akzeptanzmenge noch nicht wider, dass Konvexkombinationen von akzeptablen Risiken wiederum akzeptabel sein sollten, da sie nicht als riskanter eingestuft werden sollten als die ursprünglichen Risiken (wegen eines evtl. eintretenden Risikoausgleichs aber durchaus weniger riskant sein können als jedes der einzelnen Risiken).

3.24 Satz und Definition *Sei ρ ein monetäres Risikomaß.*

(i) Die Akzeptanzmenge $\mathfrak{A}_\rho$ ist genau dann konvex, wenn ρ ein **konvexes Risikomaß** *ist, d.h. ein monetäres Risikomaß, so dass für alle $X, Y \in \mathfrak{X}$ und alle $\lambda \in [0,1]$*

$$\rho\big(\lambda X + (1 - \lambda)Y\big) \leqslant \lambda\rho(X) + (1 - \lambda)\rho(Y).$$

(ii) *Die Akzeptanzmenge $\mathfrak{A}_\rho$ ist genau dann ein konvexer Kegel (d.h. für $X, Y \in \mathfrak{A}_\rho$, $\lambda \in [0,1]$ und $c > 0$ gilt $\lambda X + (1-\lambda)Y \in \mathfrak{A}_\rho$ und $cX \in \mathfrak{A}_\rho$), wenn ρ ein* **kohärentes Risikomaß** *ist, d.h. ein konvexes Risikomaß, das* **positiv homogen** *(oder* **skaleninvariant***) ist, für das also gilt*

$$\rho(cX) = c\rho(X) \quad \forall\, X \in \mathfrak{X}, c > 0.$$ □

Beweis. (i) Die Konvexität von $\mathfrak{A}_\rho$ folgt unmittelbar aus der Konvexität von ρ: Für $X, Y \in \mathfrak{A}_\rho$ und $\lambda \in [0,1]$ gilt

$$\rho\big(\lambda X + (1-\lambda)Y\big) \leqslant \lambda \underbrace{\rho(X)}_{\leqslant 0} + (1-\lambda) \underbrace{\rho(Y)}_{\leqslant 0} \leqslant 0.$$

Ist umgekehrt $\mathfrak{A}$ eine beliebige konvexe Akzeptanzmenge, so gilt für alle $\lambda \in [0,1]$ und alle $\varepsilon > 0$

$$\begin{aligned} & X - \rho_{\mathfrak{A}}(X) - \varepsilon \in \mathfrak{A}, \quad Y - \rho_{\mathfrak{A}}(Y) - \varepsilon \in \mathfrak{A} \\ \Longrightarrow\ & \lambda X + (1-\lambda)Y - \big(\lambda\rho_{\mathfrak{A}}(X) + (1-\lambda)\rho_{\mathfrak{A}}(Y) + \varepsilon\big) \in \mathfrak{A} \end{aligned}$$

und folglich

$$\begin{aligned} \rho_{\mathfrak{A}}(\lambda X + (1-\lambda)Y) &= \inf\big\{t \in \mathbb{R} \mid \lambda X + (1-\lambda)Y - t \in \mathfrak{A}\big\} \\ &\leqslant \lambda\rho_{\mathfrak{A}}(X) + (1-\lambda)\rho_{\mathfrak{A}}(Y), \end{aligned}$$

d.h. $\rho_{\mathfrak{A}}$ ist konvex. Ist $\mathfrak{A}_\rho$ konvex, so folglich auch $\rho = \rho_{\mathfrak{A}_\rho}$.

(ii) Ist ρ positiv homogen, so folgt unmittelbar, dass $\mathfrak{A}_\rho$ ein Kegel ist. Ist umgekehrt $\mathfrak{A}$ eine Akzeptanzmenge, die ein Kegel ist, so folgt für $X \in \mathfrak{X}$, $c > 0$

$$\begin{aligned} \rho_{\mathfrak{A}}(cX) &= \inf\big\{cu \in \mathbb{R} \mid cX - cu = \underbrace{c(X-u) \in \mathfrak{A}}_{\Leftrightarrow X - u \in \mathfrak{A}}\big\} \\ &= c \cdot \inf\{u \in \mathbb{R} \mid X - u \in \mathfrak{A}\} \\ &= c\rho_{\mathfrak{A}}(X). \end{aligned}$$

Die Behauptung folgt aus wiederum aus $\rho = \rho_{\mathfrak{A}_\rho}$. □

3.25 Beispiel Sei $(\Omega, \mathcal{A}, P)$ ein W.raum und $\alpha \in (0,1)$. Der **Value at Risk** (VaR) zum Niveau α ist definiert als

$$VaR_\alpha(X) = F_X^{\leftarrow}(\alpha) = \inf\big\{t \in \mathbb{R} \mid P\{X \leqslant t\} \geqslant \alpha\big\}$$

für $X \in \mathcal{L}(\Omega, \mathcal{A}) := \{Y : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathbb{B})\}$, wobei F_X die Verteilungsfunktion zu X bezeichne und $F_X^{\leftarrow}$ die zugehörige Quantilfunktion. Offensichtlich ist VaR_α ein positiv homogenes monetäres Risikomaß. Da VaR_α eingeschränkt auf nicht-negative Zv. gerade ein Perzentilprinzip ist, ist VaR_α nicht subadditiv und daher auch nicht konvex als Risikomaß auf $\mathcal{L}(\Omega, \mathcal{A})$ (vgl. Übungen).

Im folgenden betrachten wir VaR_α als Risikomaß eingeschränkt auf einen linearen Unterraum $\mathfrak{X}$ von $\mathcal{L}_2(\Omega, \mathcal{A})$, so dass P^X für alle $X \in \mathfrak{X}$ zu einer Lokations- und Skalenfamilie gehört, d.h. es existiert ein $X_0 \in \mathfrak{X}$ mit $E(X_0^2) < \infty$, so dass zu jedem $X \in \mathfrak{X}$ Konstanten $a \geqslant 0$ und $b \in \mathbb{R}$ existieren mit $P^X = P^{aX_0+b}$. Man kann o.E. $E(X_0) = 0$ und $Var(X_0) = 1$ annehmen. Es gilt dann $E(X) = b$ und $Var(X) = a^2$ und daher

$$VaR_\alpha(X) = F_X^{\leftarrow}(\alpha) = aF_{X_0}^{\leftarrow}(\alpha) + b = E(X) + \delta\sqrt{Var(X)}$$

mit $\delta = F_{X_0}^{\leftarrow}(\alpha)$. Ist $\delta \geqslant 0$, so ist VaR_α also gerade eine Verallgemeinerung des Standardabweichungsprinzips (bzw. im Fall $\delta = 0$ des Nettorisikoprinzips) auf nicht notwendig positive Zv. Der Beweis von Satz 3.11 (5) zeigt daher, dass VaR_α (eingeschränkt auf diese Familie von Risiken) subadditiv ist, also ein kohärentes Risikomaß. □

Ebenso wie den VaR kann man die eingangs vorgestellten Prämienprinzipien zu monetären Risikomaßen fortsetzen, wenn sie monoton und translationsinvariant sind. Sind sie zusätzlich skaleninvariant und subadditiv, so erhält man kohärente Risikomaße.

Im Folgenden soll ein kohärentes Risikomaß vorgestellt werden, das sich nicht aus einem der üblichen Prämienprinzip ergibt.

3.26 Beispiel In der Situation von Beispiel 3.25 wird der **Average Value at Risk** (AVaR) zum Niveau $\alpha \in (0, 1)$ definiert als

$$AVaR_\alpha(X) := \frac{1}{1-\alpha} \int_\alpha^1 VaR_\tau(X)\, d\tau \tag{3.1}$$

für $X \in \mathfrak{X} := \{X : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathbb{B}) \mid E(|X|) < \infty\}$. Die Monotonie und die Translationsinvarianz folgt unmittelbar aus den entsprechenden Eigenschaften von VaR_τ; Gleiches gilt für die positive Homogenität. Im Folgenden soll gezeigt werden, dass $AVaR_\alpha$ (im Gegensatz zum VaR) konvex ist. Dazu leiten wir zunächst eine alternative Darstellung her. Bekanntlich gilt für die Quantilfunktion $F_X^{\leftarrow}$ von X und die Gleichverteilung $U[0, 1]$ auf $[0, 1]$

$$F_X^{\leftarrow}(\tau) > x \iff F_X(x) < \tau \quad \text{und} \quad (U_{[0,1]})^{F_X^{\leftarrow}} = P^X.$$

Aus $F_X(F_X^{\leftarrow}(\alpha)) \geqslant \alpha$ folgt somit insbesondere $F_X^{\leftarrow}(\tau) = F_X^{\leftarrow}(\alpha)$ für alle $\tau \in [\alpha, F_X(F_X^{\leftarrow}(\alpha))]$ und daher

$$\begin{aligned}
AVaR_\alpha(X) &= \frac{1}{1-\alpha}\bigg(\int 1_{(\alpha, F_X(F_X^{\leftarrow}(\alpha))]}(\tau) F_X^{\leftarrow}(\tau)\, d\tau + \int 1_{(F_X(F_X^{\leftarrow}(\alpha)),1]}(\tau) F_X^{\leftarrow}(\tau)\, d\tau\bigg) \\
&= \frac{1}{1-\alpha}\bigg(F_X^{\leftarrow}(\alpha) \cdot \big(F_X(F_X^{\leftarrow}(\alpha)) - \alpha\big) + \int_{[0,1]} 1_{(F_X^{\leftarrow}(\alpha),\infty)}(F_x^{\leftarrow}(\tau)) F_X^{\leftarrow}(\tau)\, d\tau\bigg) \\
&= \frac{1}{1-\alpha}\bigg(F_X^{\leftarrow}(\alpha) \cdot \big(F_X(F_X^{\leftarrow}(\alpha)) - \alpha\big) + \int 1_{(F_X^{\leftarrow}(\alpha),\infty)}(u) u\, \underbrace{(U_{[0,1]})^{F_X^{\leftarrow}}}_{=P^X}(du)\bigg) \\
&= \frac{1}{1-\alpha}\bigg(E\Big(X 1_{(F_X^{\leftarrow}(\alpha),\infty)}(X)\Big) + F_X^{\leftarrow}(\alpha) \cdot \Big(F_X(F_X^{\leftarrow}(\alpha)) - \alpha\Big)\bigg).
\end{aligned} \tag{3.2}$$

Wenn $F_X(F_X^{\leftarrow}(\alpha)) = \alpha$, dann gilt also

$$AVaR_\alpha(X) = E(X \mid X > F_X^{\leftarrow}(\alpha)); \tag{3.3}$$

i.allg. gilt dagegen

$$\begin{aligned} AVaR_\alpha(X) &= \frac{1}{1-\alpha}\Big(E\Big((X-F_X^{\leftarrow}(\alpha))1_{(F_X^{\leftarrow}(\alpha),\infty)}(X)\Big) \\ &\qquad + F_X^{\leftarrow}(\alpha)\cdot\big(1-F_X(F_X^{\leftarrow}(\alpha))+F_X(F_X^{\leftarrow}(\alpha))-\alpha\big)\Big) \\ &= \frac{E\big((X-F_X^{\leftarrow}(\alpha))^+\big)}{1-\alpha}+F_X^{\leftarrow}(\alpha) \qquad (3.4) \end{aligned}$$

Wegen (3.3) wird daher der $AVaR$ auch als **Expected Shortfall** (oder Conditional Value at Risk) bezeichnet.

Aus dieser Darstellung des AVaR soll als Nächstes gefolgert werden, dass

$$AVaR_\alpha(X) = \sup_{Q\in\mathcal{Q}_\alpha} E_Q(X), \qquad (3.5)$$

wobei

$$\mathcal{Q}_\alpha := \big\{Q \text{ W.maß auf } (\Omega,\mathcal{A}) \mid Q \text{ besitzt } P\text{-Dichte } f \leqslant 1/(1-\alpha)\big\}.$$

(Hierbei bezeichnet wie üblich E_Q den Erwartungswert unter dem W.maß Q.) Die Kohärenz von $AVaR_\alpha$ ergibt sich dann mit dem nachfolgenden Satz 3.27.

Für den Nachweis von "$\geqslant$" in (3.5) sei $Q\in\mathcal{Q}_\alpha$ mit P-Dichte $f\leqslant 1/(1-\alpha)$. Dann gilt wegen $E_P(f)=1$

$$\begin{aligned} &AVaR_\alpha(X)-E_Q(X) \\ &\overset{(3.4)}{=} \frac{1}{1-\alpha}\Big(E_P\big((X-F_X^{\leftarrow}(\alpha))1_{\{X>F_X^{\leftarrow}(\alpha)\}}\big)+(1-\alpha)F_X^{\leftarrow}(\alpha)\Big)-E_P(fX) \\ &= \frac{1}{1-\alpha}E_P\Big[(X-F_X^{\leftarrow}(\alpha))1_{\{X>F_X^{\leftarrow}(\alpha)\}}-(1-\alpha)fX+(1-\alpha)fF_X^{\leftarrow}(\alpha)\Big] \\ &= \frac{1}{1-\alpha}E_P\Big[(X-F_X^{\leftarrow}(\alpha))\cdot\big(1_{\{X>F_X^{\leftarrow}(\alpha)\}}-(1-\alpha)f\big)\Big] \\ &\geqslant 0, \end{aligned}$$

da beim letzten Erwartungswert beide Faktoren nicht negativ sind auf der Menge $\{X > F_X^{\leftarrow}(\alpha)\}$ und nicht positiv auf der Komplementmenge $\{X\leqslant F_X^{\leftarrow}(\alpha)\}$.

Für den Nachweis der umgekehrten Ungleichung definieren wir

$$\kappa := \frac{F_X(F_X^{\leftarrow}(\alpha))-\alpha}{F_X(F_X^{\leftarrow}(\alpha))-F_X(F_X^{\leftarrow}(\alpha)-)},$$

falls $0<F_X(F_X^{\leftarrow}(\alpha))-F_X(F_X^{\leftarrow}(\alpha)-)=P\{X=F_X^{\leftarrow}(\alpha)\}$, und $\kappa=0$ andernfalls. In beiden Fällen gilt $\kappa\in[0,1]$, da $F_X(F_X^{\leftarrow}(\alpha))\geqslant\alpha$ und $F_X(F_X^{\leftarrow}(\alpha)-)\leqslant\alpha$. Folglich nimmt

$$f := \frac{1}{1-\alpha}\big(1_{\{X>F_X^{\leftarrow}(\alpha)\}}+\kappa\cdot 1_{\{X=F_X^{\leftarrow}(\alpha)\}}\big)$$

nur Werte zwischen 0 und $1/(1-\alpha)$ an und wegen

$$\begin{aligned} E_P(f) &= \frac{1}{1-\alpha}\big(P\{X>F_X^{\leftarrow}(\alpha)\}+\kappa\cdot P\{X=F_X^{\leftarrow}(\alpha)\}\big) \\ &= \frac{1}{1-\alpha}\big(1-F_X(F_X^{\leftarrow}(\alpha))+F_X(F_X^{\leftarrow}(\alpha))-\alpha\big) \\ &= 1 \end{aligned}$$

ist f die P-Dichte eines W.maßes $Q \in \mathcal{Q}_\alpha$. Ferner gilt

$$\begin{aligned} E_Q(X) &= E_P(fX) \\ &= \frac{1}{1-\alpha}\Big(E_P\big(X1_{\{X > F_X^{\leftarrow}(\alpha)\}}\big) + \kappa F_X^{\leftarrow}(\alpha) P\{X = F_X^{\leftarrow}(\alpha)\}\Big) \\ &\overset{(3.2)}{=} AVaR_\alpha(X) \end{aligned}$$

und folglich

$$AVaR_\alpha(X) \leqslant \sup_{Q \in \mathcal{Q}_\alpha} E_Q(X). \qquad \square$$

3.27 Satz *Sei $\mathcal{Q}$ eine nicht-leere Menge von W.maßen auf $(\Omega, \mathcal{A})$ und*

$$\mathfrak{X} := \big\{X : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathbb{B}) \mid E_Q(X) \text{ existiert in } \mathbb{R}\ \forall\, Q \in \mathcal{Q},\ \sup_{Q \in \mathcal{Q}} E_Q(X) < \infty\big\}.$$

(i) Ist $\gamma : \mathcal{Q} \to \mathbb{R}$ eine Abbildung mit $\inf_{Q \in \mathcal{Q}} \gamma(Q) > -\infty$, so definiert

$$\rho(X) := \sup_{Q \in \mathcal{Q}} (E_Q(X) - \gamma(Q))$$

ein konvexes Risikomaß auf $\mathfrak{X}$.

(ii)

$$\rho(X) := \sup_{Q \in \mathcal{Q}} E_Q(X)$$

definiert ein kohärentes Risikomaß auf $\mathfrak{X}$. □

Beweis. Übungen. □

3.28 Bemerkung Werden nur beschränkte Risiken betrachtet, d.h.

$$\mathfrak{X} := \big\{X : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathbb{B}) \mid \sup_{\omega \in \Omega} |X(\omega)| < \infty\big\},$$

so gilt unter der folgenden Stetigkeitsbedingung an ρ

$$\rho(X_n) \downarrow \rho(X) \quad \text{für alle } X_n \in \mathfrak{X}, n \in \mathbb{N}, \text{ mit } X_n \downarrow X$$

auch die Umkehrung:

(i) Ist ρ konvex, so existiert eine nicht-leere Menge $\mathcal{Q}$ von W.maßen und eine Abbildung $\gamma : \mathcal{Q} \to \mathbb{R}$, so dass

$$\rho(X) := \sup_{Q \in \mathcal{Q}} (E_Q(X) - \gamma(Q)) \quad \forall\, X \in \mathfrak{X}.$$

Die Wahl von γ und $\mathcal{Q}$ ist nicht eindeutig, aber eine minimale Wahl von γ ist gegeben durch

$$\gamma_{\min}(Q) = \sup_{X \in \mathfrak{X}} \big(E_Q(X) - \rho(X)\big)$$

für $Q \in \mathcal{Q} := \big\{\tilde{Q} \text{ W.maß auf } (\Omega, \mathcal{A}) \mid \sup_{X \in \mathfrak{X}} \big(E_{\tilde{Q}}(X) - \rho(X)\big) < \infty\big\}$. (Ist nämlich $\tilde{\gamma}$ eine weitere Wahl, so folgt aus $\rho(X) \geqslant E_Q(X) - \tilde{\gamma}(Q)$ für alle $X \in \mathfrak{X}$ gerade $\tilde{\gamma}(Q) \geqslant E_Q(X) - \rho(X)$, also $\tilde{\gamma}(Q) \geqslant \gamma_{\min}(Q)$.)

(ii) Ist ρ sogar kohärent, so existiert eine nicht-leere Menge $\mathcal{Q}$ von W.maßen, so dass

$$\rho(X) = \sup_{Q \in \mathcal{Q}} E_Q(X) \quad \forall\, X \in \mathfrak{X}. \qquad \square$$

3.29 Bemerkung Die Darstellung eines kohärenten Risikomaßes aus Satz 3.27 und Bemerkung 3.28(ii) lässt sich so interpretieren, dass das Risikomaß als "worst case" des erwarteten Verlusts aus einer Menge von stochastischen Szenarien, also W.maßen auf $(\Omega, \mathcal{A})$, berechnet wird. Dabei wird jedes stochastische Szenario gleich behandelt.

Liegt dagegen nur ein konvexes Risikomaß vor, so kann durch die Funktion γ zusätzlich berücksichtigt werden, dass manche stochastische Szenarien Q plausibler erscheinen als andere und somit der "Strafterm" $\gamma(Q)$ für die plausibleren Szenarien kleiner gewählt wird.

Wenn man beispielsweise davon ausgeht, dass das stochastische Geschehen durch ein W.maß aus einer parametrischen Familie $(Q_\vartheta)_{\vartheta \in \Theta}$ beschrieben wird und ein Punktschätzer einen Wert ϑ_0 liefert, so könnte man das Risikomaß

$$\rho(X) := \sup_{\vartheta \in \Theta} (E_{Q_\vartheta}(X) - \gamma(\vartheta))$$

betrachten, wobei $\gamma(\vartheta)$ "klein" ist für Werte von ϑ nahe ϑ_0 und $\gamma(\vartheta)$ "groß" ist für ϑ mit "großem Abstand" zu ϑ_0. $\square$

3.30 Beispiel Sei $\ell : \mathbb{R} \to \mathbb{R}$ streng monoton steigend und konvex; ℓ wird interpretiert als Verlustfunktion, d.h. $\ell(X(\omega))$ wird als negativer Wert interpretiert, dem man dem Risiko X im Szenario ω zumisst. (Beispielsweise könnte man $\ell(x) = u(x_0) - u(x_0 - x)$ für eine streng monotone Nutzenfunktion u definieren, also den negativen Nutzen des Risikos bei Anfangskapital x_0 betrachten.) Sei nun $(\Omega, \mathcal{A}, P)$ ein W.raum und

$$\mathfrak{X} := \{X : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathbb{B}) \mid \sup_{\omega \in \Omega} |X(\omega)| < \infty\}.$$

Zu vorgegebener Konstante $c_0 > \lim_{x \to -\infty} \ell(x)$ definieren wir eine Akzeptanzmenge

$$\mathfrak{A} := \{X \in \mathfrak{X} \mid E(\ell(X)) \leqslant c_0\},$$

d.h. ein Risiko wird dann als akzeptabel angesehen, wenn der erwartete Verlust eine vorgegebene Schranke nicht überschreitet. Das zugehörige Risikomaß $\rho = \rho_{\mathfrak{A}}$ ergibt sich zu

$$\rho(X) := \inf\{t \in \mathbb{R} \mid E(\ell(X - t)) \leqslant c_0\};$$

da $t \mapsto E(\ell(X - t))$ streng monoton fallend und stetig ist[1] mit $\lim_{t \to \infty} E(\ell(X - t)) = \lim_{u \to -\infty} l(u) < c_0$ und $\lim_{t \to -\infty} E(\ell(X - t)) = \lim_{u \to \infty} l(u) = \infty$, ist $\rho(X)$ also die eindeutige Lösung der Gleichung (in t)

$$E(\ell(X - t)) = c_0,$$

[1] Letzteres ergibt sich, da ℓ als konvexe Funktion stetig ist, mit dem Satz von der monotonen Konvergenz

falls $\lim_{t\to\infty} E(\ell(X-t)) < c_0$. Ist $\ell(x) = -u(-x)$ für eine Nutzenfunktion u und ist $c_0 = 0$, so ist ρ gerade das zugehörige Nullnutzenprinzip auf $\mathfrak{X}$.

Da ℓ konvex ist, ist offensichtlich auch $\mathfrak{A}$ konvex und somit gemäß Satz 3.24(i) auch ρ. Somit existiert gemäß Bemerkung 3.28 eine Darstellung

$$\rho(X) = \sup_{Q\in\mathcal{Q}}(E_Q(X) - \gamma(Q)).$$

(Die Stetigkeitsbedingung an ρ aus Bemerkung 3.28 folgt wiederum mit dem Satz von der monotonen Konvergenz.) Diese Darstellung soll nun für die exponentielle Verlustfunktion

$$\ell(x) = \mathrm{e}^{\beta x}$$

($\beta > 0$) und $c_0 > 0$ bestimmt werden.

Es gilt dann

$$\begin{aligned}
&E\big(\ell(X-\rho(X))\big) \overset{(!)}{=} c_0 \\
\iff\quad & E(\mathrm{e}^{\beta X})\mathrm{e}^{-\beta\rho(X)} = c_0 \\
\iff\quad & \rho(X) = \frac{1}{\beta}\Big(\log E(\mathrm{e}^{\beta X}) - \log c_0\Big) = \frac{1}{\beta}\Big(\log \psi_X(\beta) - \log c_0\Big),
\end{aligned}$$

d.h. bis auf eine additive Konstante entspricht ρ dem Exponentialprinzip, von dem wir in Satz 3.15 bereits gesehen haben, dass es nicht subadditiv und nicht positiv homogen (skaleninvariant) ist; also ist ρ auch nicht kohärent. Das Risikomaß ρ ist jedoch konvex!

Die minimale Wahl der Funktion γ erhält man gemäß Bemerkung 3.28(i) als

$$\begin{aligned}
\gamma_{\min}(Q) &= \sup_{X\in\mathfrak{X}}(E_Q(X) - \rho(X)) \\
&= \frac{1}{\beta}\Big(\sup_{X\in\mathfrak{X}}\big(E_Q(\beta X) - \log E(\mathrm{e}^{\beta X})\big) + \log c_0\Big) \\
&= \frac{1}{\beta}\Big(\sup_{Y\in\mathfrak{X}}\big(E_Q(Y) - \log E(\mathrm{e}^{Y})\big) + \log c_0\Big) \\
&\overset{(*)}{=} \frac{1}{\beta}\big(H(Q|P) + \log c_0\big),
\end{aligned}$$

wobei $H(Q|P)$ die **relative Entropie von Q bzgl. P** (oder Kullback-Leibler-Divergenz) ist:

$$H(Q|P) := \begin{cases} E_Q(\log f), & \text{falls } Q \text{ die } P\text{-Dichte } f \text{ besitzt}, \\ \infty, & \text{falls } Q \text{ nicht absolutstetig bzgl. } P \text{ ist}. \end{cases}$$

Die Gleichung $(*)$ folgt dabei aus dem unten angegebenen Satz 3.31. Die relative Entropie lässt sich als Maß für die Ähnlichkeit von Q zu P interpretieren. (Allerdings definiert sie *keine* Metrik, denn i.allg. gilt $H(Q|P) \neq H(P|Q)$.) Man erhält also die Darstellung

$$\rho(X) = \sup_{Q\in\mathcal{Q}}\Big(E_Q(X) - \frac{H(Q|P) + \log c_0}{\beta}\Big)$$

mit $\mathcal{Q} := \{Q \ll P \mid H(Q|P) < \infty\}$; einem "stochastischem Szenario" Q wird umso weniger Bedeutung zugemessen, je unähnlicher er dem vorgegebenen W.maß P ist. $\square$

3.31 Satz *Für die relative Entropie gilt*

$$H(Q|P) = \sup_{X \in \mathfrak{X}} \big(E_Q(X) - \log E_P(\mathrm{e}^X) \big).$$

mit

$$\mathfrak{X} := \{X : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathbb{B}) \mid \sup_{\omega \in \Omega} |X(\omega)| < \infty\}. \qquad \square$$

Beweis. Wir zeigen zunächst $H(Q|P) \geqslant E_Q(X) - \log E_P(\mathrm{e}^X)$ für alle $X \in \mathfrak{X}$. Sei dazu o.E. $H(Q|P) < \infty$ und $E_P(\mathrm{e}^X) < \infty$; insbesondere besitzt Q eine P-Dichte f. Sei Q_X das W.maß mit P-Dichte $\mathrm{e}^X / E_P(\mathrm{e}^X)$. Dann hat Q die Q_X-Dichte $g = f/(\mathrm{e}^X/E_P(\mathrm{e}^X)) = \mathrm{e}^{-X} E_P(\mathrm{e}^X) f$ und es folgt

$$H(Q|Q_X) = E_Q(\log g) = E_Q(\log f) - E_Q(X) + \log E_P(\mathrm{e}^X).$$

Da $E_Q(\log f) = H(Q|P)$ und $H(Q|Q_X)$ als relative Entropie nicht negativ ist (Übungen), folgt die behauptete Ungleichung.

Zum Nachweis der umgekehrten Ungleichung

$$H(Q|P) \leqslant \sup_{X \in \mathfrak{X}} \big(E_Q(X) - \log E_P(\mathrm{e}^X) \big) \tag{3.6}$$

sei zunächst angenommen, dass Q nicht absolutstetig ist bzgl. P, d.h. es existiert ein $A \in \mathcal{A}$ mit $P(A) = 0$, aber $Q(A) > 0$. Dann gilt für $X_n = n 1_A \in \mathfrak{X}$

$$E_Q(X_n) - \log E_P(\underbrace{\mathrm{e}^{X_n}}_{=1\ P-f.s.}) = nQ(A) \to \infty,$$

also (3.6).

Sei nun $Q \ll P$ mit P-Dichte f. Ist f beschränkt und von 0 weg beschränkt, so ist $X = \log f \in \mathfrak{X}$ mit

$$E_Q(X) - \log E_P(\mathrm{e}^X) = H(Q|P) - \log \underbrace{E_P(f)}_{=1} = H(Q|P).$$

Andernfalls betrachte

$$X_n := \begin{cases} -n & & X \leqslant -n \\ X & \text{falls} & -n < X < n \\ n & & X \geqslant n \end{cases} \in \mathfrak{X}.$$

Dann gilt für $X \leqslant 0$

$$f X_n \geqslant f \log f \geqslant -\mathrm{e}^{-1}$$

(da $\inf_{x>0} x \log x = -1/e$) und für $X \geqslant 0$ sogar $f X_n \geqslant 0$, d.h. $f X_n$ ist nach unten beschränkt. Das Lemma von Fatou liefert daher

$$\liminf_{n\to\infty} E_Q(X_n) = \liminf_{n\to\infty} E_P(f X_n) \geqslant E_P(\liminf_{n\to\infty} f X_n) = E_P(f \log f) = H(Q|P).$$

Andererseits gilt nach dem Satz von der monotonen Konvergenz

$$E_P(\mathrm{e}^{X_n} 1_{\{X \geqslant 0\}}) \to E_P(\mathrm{e}^X 1_{\{X \geqslant 0\}})$$

und nach dem Satz von der majorisierten Konvergenz

$$E_P(\mathrm{e}^{X_n} 1_{\{X<0\}}) \to E_P(\mathrm{e}^X 1_{\{X<0\}}).$$

Zusammen erhält man also

$$\liminf_{n\to\infty} (E_Q(X_n) - \log E_P(\mathrm{e}^{X_n})) \geqslant H(Q|P) - \log E_P(\mathrm{e}^X) = H(Q|P),$$

d.h. (3.6). □

Das Exponentialprinzip berücksichtigt also doch in gewissem Umfang Diversifikationseffekte, obwohl es nicht subadditiv ist. Ebenso lassen sich weitere konvexe, aber nicht subadditive Risikomaße konstruieren (z.B. Nullnutzenprinzipien). Ob man nun nur die Konvexität des Risikomaßes fordern sollte oder aber die Subadditivität (oder gar die Kohärenz), hängt z.T. von dem konkreten Zweck ab, zu dem das Risikomaß berechnet werden soll. So steht es einem Kapitalanleger oft frei, ein Portfolio aus (nahezu) beliebigen Anteilen eines Risikos zusammen zu setzen. (Ausnahmen gibt es aber auch hier, etwa bei dem Erwerb von Immobilien oder von "Private Equity".) Die Konvexitätsbedingung spiegelt daher die Diversifikationsmöglichkeiten gut wider. Ein Versicherungsunternehmen kann dagegen oft nicht nur einen Bruchteil eines Risikos X versichern, d.h. Konvexkombinationen von einzelnen Risiken sind nicht in jedem Fall realisierbar. (Allerdings kann das Versicherungsunternehmen einen Teil des Risikos über Rückversicherungen evtl. doch weiter geben oder ein Konsortium von Versicherungsunternehmen kann das Risiko gemeinsam zeichnen und dann auf die beteiligten Unternehmen aufteilen.) In dem Fall kann Diversifikation nur durch Hinzunahme ganzer Risiken erfolgen und die Subadditivität kann eine angemessene Bedingung darstellen.