

Mathematische Methoden in den Wirtschaftswissenschaften

Grundkurs

Josef Leydold

Department für Statistik und Mathematik – WU Wien

SS 2006

Übersicht

Grundlagen

- Renditen
- Elementare Wahrscheinlichkeitsrechnung
- Zufallsvariablen, Erwartungswert und Varianz
- Versicherungsmathematik
- Exkurs: Multivariate Analysis
- Erwarteter Nutzen
- Taylorreihen
- Kovarianzen und Korrelation

- Grundbegriffe
- Portfolio Management
- Derivative
- Das Binomialmodell
- Das stochastische Modell für Aktienkurse
- Stochastische Analysis
- Das Black-Scholes Modell

Prognose und Zeitreihenanalyse

- Daten und Indizes
- Beschreibung von Zeitreihen
- Zerlegung von Zeitreihen, Glättung und Saisonbereinigung
- Regression
- *Autoregressive moving average* (ARMA) Modelle

Literatur

- Cuthbertson, K. (1996): *Quantitative Financial Economics*, Wiley, Chichester. Kapitel 1
Renditen, Nutzenfunktion
- Spremann, K. (1996): *Wirtschaft, Investition und Finanzierung*, Oldenbourg, München.
deutsche Synonyme für die englischen Ausdrücke
- Bosch, Karl(1990): *Finanzmathematik*, Oldenbourg. Kapitel 7
Versicherungsmathematik
- Wilmott, P. (2001): *Wilmott Introduces Quantitative Finance*, Wiley, Chichester. Kapitelauswahl
Quantitative Finance
- Makridakis, Wheelwright and Hyndman (1998): *Forecasting: Methods and Applications* (3rd edition), Wiley. Kapitelauswahl
Zeitreihenanalyse

Renditen

Lernziele

Basiskonzepte in der Finanzierung

- Renditen
- Zinseszinsrechnung
- Gegenwartswert (*PV*) auch Barwert, Kapitalwert oder Zeitwert, Endwert
- Renditen auf Aktien, Bonds und realen Kapitalanlagen
- Einfacher, effektiver Zinssatz, stetige Zinsen, interne Ertragsrate

Einfacher und effektiver Zinssatz

Im Gegensatz zum theoretischen Konzept der **(Zinseszins)** werden stets „**einfache**“ (*simple*) **Zinssätze** publiziert.

Beispiel:

Ein Zinssatz von 5%, der alle sechs Monate zu bezahlen ist, wird als „einfacher“ Zinssatz von 10% pro Jahr angegeben.

Der **effektive** Jahreszinssatz bei zwei aufeinanderfolgenden Sechs-Monatsanleihen ist durch den Zinseszinseffekt

$$(1.05)^2 = 1.1025 \quad \text{d.s.} \quad 10.25\% > 2 \times 5\% = 10\%$$

- Der einfache Zinssatz ist in diesem Beispiel 10%,
- Der **effektive** (konforme, äquivalente) Jahreszinssatz 10.25%.
- Die **Zinsperiode** ist 6 Monate.
- Die Anzahl der Zinsperioden ist 2 (pro Jahr).

Endwert und Zinseszinsen

Angenommen wir investieren am 1. Jänner x €. Wie hoch ist der **Endwert** dieser Investition nach n Jahren, wenn sich die Zinsperioden zu denen die Zinsen bezahlt werden ändern?

x sei ein Betrag, der für n Jahre mit einer Rendite R (in % pro Jahr) investiert wird.

Wenn die Zinsperioden jeweils die Länge von einem Jahr haben, und sich beginnend mit dem 1. Jänner nahtlos aneinanderreihen, so ist der zukünftige Wert (Endwert), FV_n (*future value*),

$$FV_n = x(1 + R)^n$$

Endwert und Zinsperioden

Werden die Zinsen jedoch m Mal pro Jahr bezahlt, dann ist der Endwert nach n Jahren

$$FV_n^m = x \left[\left(1 + \frac{R}{m} \right)^m \right]^n = x \left(1 + \frac{R}{m} \right)^{m \cdot n}$$

R/m wird als **Periodenzinssatz** oder „relativer“ (*periodic*) Zinssatz bezeichnet.

Formel: $a^{b^c} = (a^b)^c = a^{b \cdot c}$ falls definiert. a, b, c reell.

Beispiel:

$$5^{3^2} = (5^3)^2 = 125^2 = 15625$$

$$5^{3 \cdot 2} = 5^6 = 15625$$

Stetige Verzinsung

Wenn die Zinsperioden kürzer werden – das ist eine Erhöhung der Anzahl der Zinseszinsperioden pro Jahr – nähert sich die Verzinsung einer **stetigen** Verzinsung. Der zukünftige Wert wird zu

$$FV_n^c = \lim_{m \rightarrow \infty} x \left[\left(1 + \frac{R}{m} \right)^m \right]^n = x \exp(R)^n = x \exp(R \cdot n)$$

FV_n^c steht für Endwert bei stetiger Verzinsung (*future value in case of continuous compounding*) nach n Jahren.

Formel:

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n} \right)^n = \exp(x) \quad [= e^x]$$

$$\exp(x)^y = (e^x)^y = e^{x \cdot y} = \exp(x \cdot y)$$

Stetige Verzinsung / Beispiel

Der Wert von 100 € am Ende des ersten Jahres ($n = 1$) bei einem einfachen jährlichen Zinssatz R von 10% und einer Frequenz (Anzahl der Perioden pro Jahr) m .

Periodizität	m	Wert in €
jährlich	1	110.000
vierteljährlich	4	110.381
wöchentlich	52	110.506
täglich	250	110.515
stetig	∞	110.517

Zusammenhang zwischen Zinssätzen

Beziehung zwischen **einfachem** Zinssatz R , **Periodenzinssatz** R_p , **effektiven** jährlichen Rate R_f , und **stetigem** Zinssatz R_c (c wie *continuous*).

- Einfacher und Periodenzinssatz:

$$R = R_p m \quad \text{bzw.} \quad R_p = \frac{R}{m}$$

- Effektive Rate R_f durch $x(1 + R_f) = x \left(1 + \frac{R}{m}\right)^m$ implizit gegeben ($n = 1$):

$$R_f = \left(1 + \frac{R}{m}\right)^m - 1$$

Zusammenhang zwischen Zinssätzen / 2

- Der (implizite) stetige Zinssatz R_c liefert dieselbe jährliche Verzinsung wie der effektive Zinssatz:

$$x(1 + R_f) = x \left(1 + \frac{R}{m}\right)^m = x \exp(R_c) \quad (n = 1)$$

$$R_c = m \log \left(1 + \frac{R}{m}\right)$$

- Umgekehrt kann der einfache Zinssatz R durch R_c ausgedrückt werden:

$$R = m \left[\exp \left(\frac{R_c}{m} \right) - 1 \right]$$

$\log(x)$ bezeichnet stets den **natürlichen** Logarithmus.

Zusammenhang zwischen Zinssätzen / Beispiel

Eine Investition liefere alle 6 Monate einen Periodenzinssatz von 5%:

$$R_p = 0.05, \quad m = 2$$

$$R = R_p \cdot m = 0.05 \cdot 2 = 0.10$$

$$R_f = \left(1 + \frac{R}{m}\right)^m - 1 = \left(1 + \frac{0.10}{2}\right)^2 - 1 = 0.1025$$

$$R_c = m \log\left(1 + \frac{R}{m}\right) = 2 \log\left(1 + \frac{0.10}{2}\right) = 0.0976$$

- Effektive Rate größer als einfache (Zinseszinsseffekt).
(5% Zinsen der ersten Periode werden wieder mit 5% verzinst.)
- Stetiger Zinssatz kleiner als effektiver, da durch die unendlichmalige Aneinanderreihung der Perioden zusätzliche Zinseszinsen anfallen.
- Der stetige Zinssatz ist sogar kleiner als der einfache.

Barwert (Gegenwartswert, Zeitwert)

Wir investieren einen Betrag von $x \text{ €}$ mit einem sicheren, fixen, jährlichen Zinssatz $rs^{(n)}$ auf n Jahre. Der zukünftige Wert (Endwert) ergibt sich als

$$FV_n = x(1 + rs^{(n)})^n$$

Umgekehrt sind wir für eine Investition dieser Art (erhalte $FV_n \text{ €}$ in n Perioden) bereit, heute $x \text{ €}$ zu bezahlen.

Der heutige Wert von FV_n in n Perioden (**Barwert**, (diskontierter) Gegenwartswert, Kapitalwert, Zeitwert, (*discounted*) *present value*) ist daher

$$PV = \frac{FV_n}{(1 + rs^{(n)})^n}$$

$\frac{1}{(1 + rs^{(n)})^n}$ wird als Diskontfaktor oder Abzinsungsfaktor bezeichnet.

Zinsstruktur(kurve)

Betrachten den Zinsertrag für Investitionen, die unterschiedlich lange dauern, i.e., für unterschiedlich lange gebundenes Kapital.

Der Zinssatz für Investitionen mit einer Laufzeit von i Jahren bezeichnen wir mit $rs^{(i)}, i = 1, \dots, n$.

Die **Zinsstrukturkurve** (*term structure of interest*) fasst diese Zinssätze zusammen

$$\{rs^{(i)}, i = 1, \dots, n\} \quad (= \{rs^{(1)}, rs^{(2)}, \dots, rs^{(n)}\})$$

Sie gibt den Zinssatz in Abhängigkeit von der Investitionsdauer an.

$rs^{(i)}$ wird auch als **spot** Zinssatz bezeichnet:

$rs^{(1)}$ ist die Rate für Geld, das heute für ein Jahr angelegt wird,

$rs^{(2)}$ ist die Rate für Geld von Zeitpunkt 0 bis Zeitpunkt 2.

Flache Zinsstrukturkurve

Nehmen wir nun für das Folgende eine **flache** Zinsstruktur an:

$$r_s^{(i)} = r \quad \text{für alle } i.$$

Angenommen wir erhalten einen Auszahlungsstrom FV_i für die Jahre $i = 1, \dots, n$. Dann ist der Gegenwartswert PV die Summe aller Erträge abgezinst mit den Diskontfaktoren $\frac{1}{(1+r)^i}$:

$$PV = \sum_{i=1}^n \frac{FV_i}{(1+r)^i}$$

PV ist eine monoton fallende Funktion von r : (sofern alle $FV_i > 0$)

$$PV = PV(r) \quad \text{mit} \quad PV(r)' = \frac{dPV}{dr} < 0$$

Physisches Investitionsprojekt

Betrachten wir nun eine physische Investition, wie den Bau einer neuen Fabrik, von der wir Nettoauszahlungen FV_i erwarten. Angenommen die Kapitalkosten des Projekts zur Zeit $t = 0$ sind KC (*costs*).

Wir investieren, wenn der Gegenwartswert der Auszahlungen zumindest die Kosten erreicht:

$$PV \geq KC \quad \text{bzw.} \quad \text{Netto-PV} = PV - KC \geq 0$$

Interne Ertragsrate (IRR)

Steigt der Zinssatz r , so fällt der PV , wie auch der Netto- PV . Bei einem bestimmten Wert wird der Netto- PV Null. Dieser Zinssatz wird **interne Ertragsrate** IRR (*internal rate of return*) genannt.

$$\text{Netto-PV}(IRR) = 0$$

Die interne Ertragsrate kann als konstantes y aus

$$\text{Netto-PV}(y) = \sum_{i=1}^n \frac{FV_i}{(1+y)^i} - KC = 0$$

numerisch berechnet werden (z.B.: Newton-Verfahren).
(Bei mehreren positiven Lösungen wird der kleinere Wert genommen.)

Interne Ertragsrate (IRR) / 2

Wir investieren, wenn

- $IRR \geq r$ und $IRR > 0$
bzw.
Interne Rate \geq Rate für alternative Investitionen

oder wenn

- der PV eines Investitionsprojektes A positiv und größer ist als der PV eines Projekts B.

Gegenwartswert und Unsicherheit

Bei einer nicht-flachen Zinsstrukturkurve erhalten wir als Barwert

$$PV = \sum_{i=1}^n \frac{FV_i}{(1 + rs^{(i)})^i} = \sum_{i=1}^n \delta_i FV_i$$

mit den Diskontfaktoren

$$\delta_i = \frac{1}{(1 + rs^{(i)})^i}$$

Zukünftige Erträge sind oft mit Unsicherheit behaftet. Dieser wird durch die Einführung einer **Risikoprämie** $rp^{(i)}$ Rechnung tragen. r wird dann zu $rs^{(i)} + rp^{(i)}$ und wir erhalten die Diskontfaktoren

$$\delta_i = \frac{1}{(1 + rs^{(i)} + rp^{(i)})^i}$$

Endfällige Bonds ohne Coupons

Ein endfälliger Bond ohne Coupons hat einen fixen Tilgungskurs, M , eine bekannte Laufzeit und zahlt keine Coupons. Er wird zu einem bestimmten Preis P_t zum Zeitpunkt t gekauft, $P_t < M$.

Endfällige Bonds ohne Coupons heißen auch Zero-Coupon Bonds (*pure discount bonds, zero coupon bonds*, bei kurzer Laufzeit bis ein Jahr *bills*).

Wir investieren nun in endfällige Bonds ohne Coupons. Für eine einjährige Laufzeit erhalten wir als Ertragsrate

$$rs_t^{(1)} = \frac{M_1 - P_{1t}}{P_{1t}}$$

Endfällige Bonds ohne Coupons / 2

Aus der Sicht des PV und dem IRR ist $P_{1t} = \frac{M_1}{(1 + y_{1t})}$

nach y zu lösen: $y_{1t} = rs_t^{(1)}$.

Für einen zweijährigen endfälligen Bond finden wir

$$P_{2t} = \frac{M_2}{(1 + rs_t^{(2)})^2}$$

für den spot Zinssatz $rs_t^{(2)}$.

Die Beziehung zwischen Preis und Ertragsrate ist für Bonds invers:

Je höher der Preis P_t , desto niedriger ist die spot Rate $rs_t^{(j)}$.

Coupon Bonds

Ein Coupon Bond (Bond mit Zinsschein; kann nicht vorzeitig eingelöst werden) zahlt einen fixen Betrag C zu fixen Zeitpunkten (nehmen wir hier an jährlich), hat eine fixe Laufzeit n und einen fixen Tilgungskurs M_n . Für einen Bond mit n Jahren Laufzeit bis zur Fälligkeit sei der aktuelle Preis $P_t^{(n)}$.

Die interne Ertragsrate des Bonds, R_t^y , kann aus

$$P_t^{(n)} = \frac{C}{(1 + R_t^y)} + \frac{C}{(1 + R_t^y)^2} + \dots + \frac{C + M_n}{(1 + R_t^y)^n}$$

berechnet werden.

Halteperioden und Renditen von Aktien

Angenommen P_t ist der Preis einer Aktie zum Zeitpunkt t . Wir kaufen sie und halten sie eine Periode. Die zugehörige Rendite H_{t+1} ist dann

$$H_{t+1} = \frac{P_{t+1} - P_t}{P_t} + \frac{D_{t+1}}{P_t}$$

Der erste Term ist der anteilige Gewinn/Verlust über eine Periode auf Grund der Kursänderung, der zweite der anteilige Dividendenertrag.

Ex post (im Nachhinein) sind P_{t+1} und D_{t+1} bekannt, aber **ex ante** (im Vorhinein) unsicher.

Halteperioden und Renditen von Aktien / 2

H_{t+i+1} sei die 1-Perioden Rendite zwischen $(t+i)$ und $(t+i+1)$.

$$1 + H_{t+i+1} = \frac{P_{t+i+1} + D_{t+i+1}}{P_{t+i}}$$

Eine Investition von A_t € zum Zeitpunkt t wird demnach nach n Perioden

$$A_{t+n} = A_t (1 + H_{t+1}) (1 + H_{t+2}) \cdots (1 + H_{t+n})$$

wert sein.

Kapitel 2

Wahrscheinlichkeitstheorie

Lernziele

- Experimente, Ereignisse und Ereignisraum
- Wahrscheinlichkeit
- Rechnen mit Wahrscheinlichkeiten
- Bedingte Wahrscheinlichkeit
- Stochastische Unabhängigkeit
- Satz von totalen Wahrscheinlichkeit
- Satz von Bayes

Problem

- **Problem 1:**

Wie groß ist die Wahrscheinlichkeit, dass beim Werfen einer Münze „Kopf“ kommt?

Verwenden Sie eine Skala von

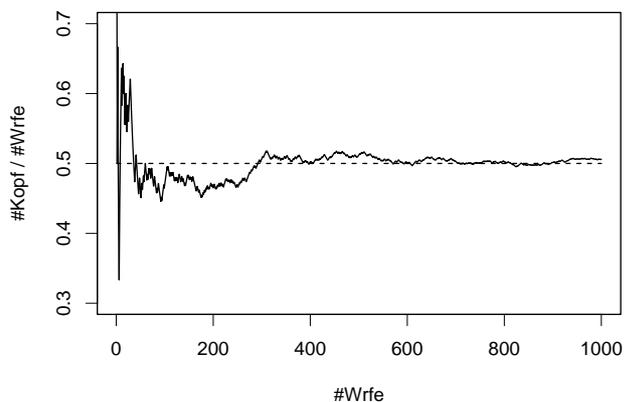
0 („sicher nicht“) bis **1 („sicher“)**.

- **Problem 2:**

Werfen Sie nun eine Münze zweimal!

Haben Sie genau einmal „Kopf“ und einmal „Zahl“ geworfen?

Wiederholungen des Münzwurfs



Experiment und Ereignis

- **(Zufalls-) Experiment**

- Ein Verfahren um eine Beobachtung zu erhalten.
- Spezifikation des Merkmals:
Was interessiert mich an dem Experiment?
Was wird beobachtet?

- **Ereignis**

- Ein mögliches Ergebnis eines Experiments.
- Ereignisse werden mit Großbuchstaben, A, B, C, \dots , bezeichnet.

- **Elementarereignis**

- Elementares (einfachstes) Ergebnis eines Experiments.

- **Ereignisraum (S)**

- Menge aller möglichen Elementarereignisse.

Experimente / Beispiel

- **Experiment:**
Ziehe Spielkarte. Beobachte Farbe und Typ der Karte.
- **Elementarereignisse:** Herz-2, . . . , Pik-König, Pik-Ass .
- **Ereignisse:**
„schwarze Karte“, „As“, „Herz-König“, „Pik“, „Bild“, „rote 5“,
- **Ereignisraum:** Alle möglichen Kombinationen von Karten.

Josef Leydold © 2006

Mathematische Methoden – II – Wahrscheinlichkeitstheorie – 6 / 24

Experimente / Beispiel

- **Experiment:**
Werfe 2 Münzen. Beobachtet wird Kopf/Zahl.
- **Elementarereignisse:** KK, KZ, ZK, ZZ
- **Ereignisse:**

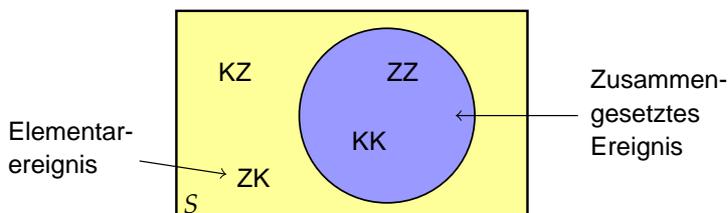
(Zusammengesetztes) Ereignis	Menge der zugehörigen Elementarereignisse
Ereignisraum	KK, KZ, ZK, ZZ
1 Kopf und 1 Zahl	KZ, ZK
Kopf auf 1. Münze	KK, KZ
zumindest einmal Kopf	KK, KZ, ZK
Kopf auf beiden Münzen	KK

Josef Leydold © 2006

Mathematische Methoden – II – Wahrscheinlichkeitstheorie – 7 / 24

Venn-Diagramm

Experiment: Werfen von 2 Münzen. Beobachtet wird Kopf/Zahl.



Ereignisraum: $S = \{KK, KZ, ZK, ZZ\}$

Josef Leydold © 2006

Mathematische Methoden – II – Wahrscheinlichkeitstheorie – 8 / 24

Zusammengesetzte Ereignisse

Zusammengesetzte Ereignisse erhält man durch Bildung von

- **Durchschnitt**

- Alle Elementarereignisse, die in beiden Ereignissen A **und** B enthalten sind.
- Symbol: \cap (d.h., $A \cap B$)

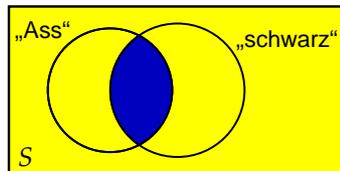
- **Vereinigung**

- Alle Elementarereignisse, die in Ereignis A **oder** B enthalten sind.
- Symbol: \cup (d.h., $A \cup B$)

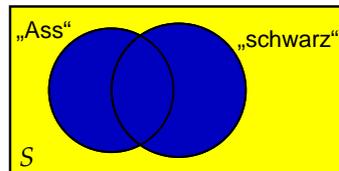
- **Komplement**

- Alle Elementarereignisse, die in **nicht** im Ereignis A enthalten sind.
- Symbol: \bar{A}

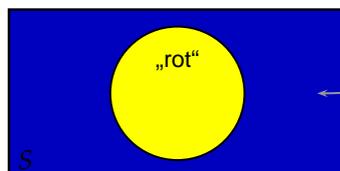
Zusammengesetzte Ereignisse / Venn-Diagramme



Durchschnitt



Vereinigung



Komplement

← Ereignis: nicht-„rot“

Wahrscheinlichkeit

Die **Wahrscheinlichkeit** ist . . .

- Numerisches Maß für die Chance, dass ein Ereignis eintritt
 - $P(\text{Ereignis})$, $P(A)$, $\text{Probability}(A)$
- Liegt zwischen 0 (sicher nicht) und 1 (sicher).
- Summe der Wahrscheinlichkeiten aller Elementarereignisse ist 1.

Wahrscheinlichkeiten spezieller Ereignisse

- Unmögliches Ereignis A
Ereignis mit der Wahrscheinlichkeit 0.
 - $P(A) = 0$
- Sicheres Ereignis S
Ereignis mit der Wahrscheinlichkeit 1.
 - $P(S) = 1$
- Komplementäreignis zu A , \bar{A}
 - $P(\bar{A}) = 1 - P(A)$ bzw. $P(A) + P(\bar{A}) = 1$
- Einander ausschließende Ereignisse A und B
 - $P(A \cap B) = 0$
 - $P(A \cap \bar{A}) = 0$ (Gilt für jede Wahl von A .)

Zuordnung von Wahrscheinlichkeiten

- *A priori* Methode
- Empirische Methode
- Subjektive Methode

A priori Methode

- Struktur des Experiments muß im **Vorhinein** bekannt sein.

Beispiel: Würfeln (idealer Würfel)

Jede Augenzahl ist gleichwahrscheinlich:

$$P(\{1\}) = P(\{2\}) = \dots = P(\{6\}) = \frac{1}{6}$$

- Regel für gleichwahrscheinliche Elementarereignisse:

$$P(\text{Ereignis}) = \frac{\text{Anzahl der günstigen Fälle}}{\text{Anzahl der möglichen Fälle}} = \frac{G}{M}$$

Beispiel: Würfeln (idealer Würfel)

$$P(\{1\}) = \frac{G}{M} = \frac{1}{6} \quad P(\{1,2\}) = \frac{G}{M} = \frac{2}{6}$$

$$P(\text{Gerade Augenzahl}) = P(\{2,4,6\}) = \frac{G}{M} = \frac{3}{6}$$

Empirische Methode

- Daten werden bei Experiment gesammelt.
- Auswertung:

$$P(\text{Ereignis}) = \frac{\text{Anzahl mit Eigenschaft}}{\text{Anzahl der Wiederholungen}} = \frac{X}{N}$$

Beispiel: Ausschußwahrscheinlichkeit

1000 Teile werden auf Fehler kontrolliert. Es werden 20 defekte Teile festgestellt.

$$P(\text{„defekt“}) = \frac{X}{N} = \frac{20}{1000} = 0.02 = 2\%$$

- **Annahme:** Es gibt keine Änderung der Anteile.

Subjektive Methode

- Die Wahrscheinlichkeit wird vor dem Experiment erhoben.
- Basiert auf individuellem Wissen, Erfahrung.
- Die Antwort unterscheidet sich je nachdem, wen man fragt.

Beispiele:

- Frage an den Experten:
Wo wird der Aktienmarkt im Dezember stehen?
 $P(\text{DAX} \leq 2500) = ?$
- Frage an den Fußballfan:
Wer wird nächster Fußballmeister?
 $P(X \text{ wird Meister}) = ?$
(Wettbüros „messen“ subjektive Wahrscheinlichkeiten.)

Zuordnung von Wahrscheinlichkeiten / Beispiel

Welche Methoden sind auf folgende Problemstellungen anzuwenden?

- Werfen einer Münze
- Lotto spielen
- Aktien veranlagen
- Sportwetten
- Wie groß ist die Wahrscheinlichkeit, dass ein Teilnehmer dieser LV die Note „Gut“ bekommt?
- *Risk Management*
- ...

Additionsregel

- Die **Additionsregel** wird verwendet, um Wahrscheinlichkeiten von Vereinigungen von Ereignissen, $A \cup B$, zu berechnen.

$$P(A \text{ oder } B) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- Für **einander ausschließende** Ereignisse ($P(A \cap B) = 0$) gilt

$$P(A \text{ oder } B) = P(A \cup B) = P(A) + P(B)$$

Bedingte Wahrscheinlichkeit

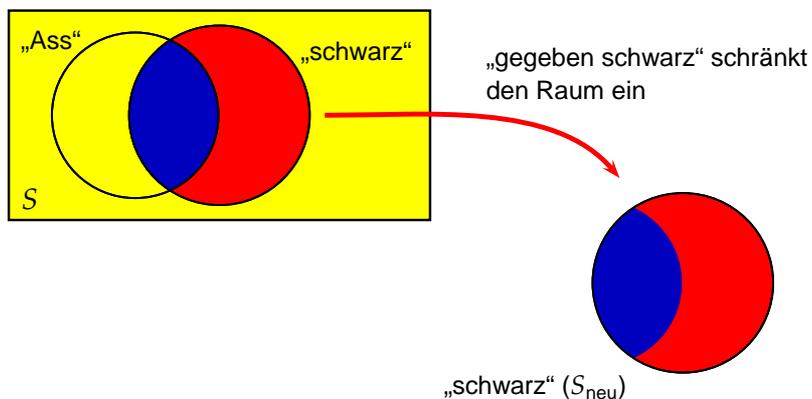
- Wahrscheinlichkeit eines Ereignisses, **gegeben** dass ein anderes Ereignis eingetreten ist.
- Schränkt die Grundgesamtheit auf den Teil ein, der zur **neuen Information** passt. (Einige Elementarereignisse scheiden aus.)
- Notation und Definition:**

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

- Sprechweise:**
„Wahrscheinlichkeit von A unter der Bedingung B “,
„Wahrscheinlichkeit von A gegeben B “

Bedingte Wahrscheinlichkeit / Venn-Diagramm

Experiment: Ziehen einer Karte. Beobachtet wird Art und Farbe.



Statistische Unabhängigkeit

- Das Eintreten eines Ereignisses A hat keine Auswirkung auf die Wahrscheinlichkeit des Eintretens eines anderen Ereignisses B . A und B sind dann **(stochastisch) unabhängig**.
 - Beispiel: Werfen von 2 Münzen
Das Ergebnis des 2. Wurfs ist vom Ergebnis des 1. Wurfs unabhängig.
- Keine Kausalität!
(Der Storch bringt nicht die Kinder.)
- Überprüfung, ob A und B unabhängig sind: Es gilt

$$P(A|B) = P(A) \quad \text{und} \quad P(A \cap B) = P(A) P(B)$$

Multiplikationsregel

- Die **Multiplikationsregel** wird verwendet, um Wahrscheinlichkeiten von Durchschnitten von Ereignissen, $A \cap B$, zu berechnen.

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

- Für **unabhängige** Ereignisse, A, B , gilt

$$P(A \text{ und } B) = P(A \cap B) = P(A) P(B)$$

Satz von der totalen Wahrscheinlichkeit

Seien A_1, A_2, \dots, A_n gegenseitig ausschließende Ereignisse, die den Ereignisraum S ganz ausfüllen (*Partition*), i.e.,

$$A_1 \cup \dots \cup A_n = S \quad \text{und} \quad A_i \cap A_j = \emptyset \quad \text{für } i \neq j$$

Jedes beliebige Ereignis E lässt sich darstellen als

$$E = (E \cap A_1) \cup \dots \cup (E \cap A_n)$$

Nach dem Additionssatz und den Multiplikationssatz gilt

$$P(E) = \sum_{i=1}^n P(E|A_i) P(A_i)$$

Satz von Bayes

Nach dem Multiplikationssatz gilt

$$P(E \cap A_i) = P(A_i|E) P(E) = P(E|A_i) P(A_i)$$

Zusammen mit dem Satz von der totalen Wahrscheinlichkeit erhält man den **Satz von Bayes**:

$$P(A_i|E) = \frac{P(E|A_i) P(A_i)}{\sum_{i=1}^n P(E|A_i) P(A_i)}$$

Kapitel 3

Zufallsvariable

Lernziele

- Diskrete und stetige Zufallsvariable
- Wahrscheinlichkeitsfunktion, Dichte und Verteilungsfunktion
- Erwartungswert, Varianz und Standardabweichung
- Binomial- und Poissonverteilung
- Normal- und Exponentialverteilung

Problem

Sie möchten einen Multiple-Choice-Test mit 20 Fragen bestehen. Für jede Frage gibt es 5 Antwortmöglichkeiten, wobei immer nur (genau) eine richtig ist.

Die erste Frage können Sie nicht beantworten und müssen raten. Wie groß ist die Wahrscheinlichkeit, dass Sie richtig raten?

Wenn Sie keine einzige Frage beantworten können und daher bei jeder Antwort raten, wie groß ist dann die Wahrscheinlichkeit, dass Sie den Test bestehen?

Mindestens 10 Antworten müssen richtig sein.

Zufallsvariable

Eine **Zufallsvariable** ist eine Abbildung, die jedem Ergebnis eines Zufallsexperiments eine reelle Zahl zuordnet.

- Experiment: Frage nach Anzahl PKW im Haushalt
Die ZV ordnet jedem Haushalt die Anzahl PKW zu.
Werte: $0, 1, 2, \dots$ – **diskrete ZV**
- Experiment: Gewichtsbestimmung von Äpfeln
Gewicht eines Apfels(g): $123, 245, 301, \dots$ – **stetige ZV**

Zufallsvariable werden mit Großbuchstaben bezeichnet, z.B. X , mögliche Realisationen mit Kleinbuchstaben, hier x .

Diskrete Zufallsvariable

- Nur ganze Zahlen sind als Ergebnisse möglich.
Z.B.: $0, 1, 2, \dots$
- Tritt als Ergebnis von Zählexperimenten auf.
- Hat meistens nur eine endliche Anzahl an Werten.
- Wird durch die **Wahrscheinlichkeitsfunktion** oder **Verteilungsfunktion** beschrieben.

Wahrscheinlichkeitsfunktion

- Liste **aller** Paare $(x, f(x))$
 - x ... Wert der Zufallsvariablen X
 - $f(x)$... Wahrscheinlichkeit, dass Wert x eintritt.

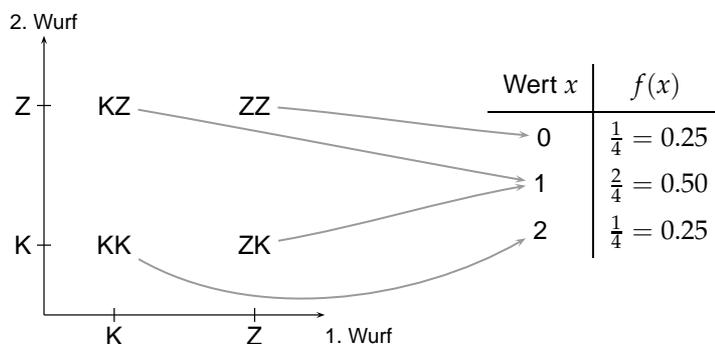
$$f(x) = P(X = x)$$

- Jeder x Wert kommt nur einmal vor. (Funktion!)
- $0 \leq f(x) \leq 1$
- $\sum_{x \in S} f(x) = 1$

Wahrscheinlichkeitsfunktion / Beispiel

Experiment: Werfe zwei Münzen; Zähle Anzahl „Kopf“.

Zufallsvariable heißt: „Anzahl Köpfe“



Wahrscheinlichkeitsfunktion / Darstellung

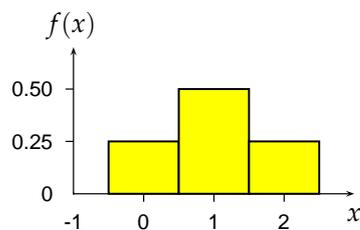
- **Liste**

$$\{(0, 0.25), (1, 0.50), (2, 0.25)\}$$

- **Tabelle**

x	Anzahl	$f(x)$
0	1	$\frac{1}{4} = 0.25$
1	2	$\frac{2}{4} = 0.50$
2	1	$\frac{1}{4} = 0.25$

- **Grafik**



- **Formel**

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

Verteilungsfunktion

Gibt die Wahrscheinlichkeit an, dass die diskrete Zufallsvariable nicht größer als ein vorgegebener Wert ist:

$$F(x) = P(X \leq x) = \sum_{y \leq x} f(y)$$

Die Verteilungsfunktion ist monoton in x .

Erwartungswert

Der **Erwartungswert** (*expectation*) einer diskreten Zufallsvariable ist die gewichtete Summe der möglichen Realisationen.

$$E(X) = \mu_x = \sum_x x f(x) = \sum_x x P(X = x)$$

Achtung!

Der Erwartungswert muss nicht immer existieren.

Erwartungswert / Beispiel

Angenommen sie arbeiten für eine Versicherung, und verkaufen Lebensversicherungen mit einer Vertragssumme von 10 000 €. Die Jahresprämie dafür beträgt 290 €.

Sterbetafeln geben für einen Kunden in Abhängigkeit von Alter, Geschlecht, Gesundheitszustand, etc. eine Wahrscheinlichkeit von 0.001 an, dass er in diesem Jahr verstirbt.

Was ist der erwartete jährliche Gewinn für Polizen dieser Art?

Die ZV Nettogewinn, Einzahlung minus Auszahlung, bezeichnen wir mit X . Wahrscheinlichkeitsfunktion:

Gewinn, x	Ereignis	Wahrscheinlichkeit
290 €	Kunde lebt	0.999
-9710 €	Kunde stirbt	0.001

Erwartungswert / Beispiel – Fortsetzung

Wenn die Versicherung 1000 Polizzen davon verkauft, so wird – vereinfacht gesprochen – in 999 Fällen von den 1000 der Kunde das Jahresende erleben, in einem von den 1000 Fällen wird er innerhalb dieses Jahres sterben.

290 € wird die Versicherung in 999 Fällen ohne monetäre Gegenleistung erhalten, also einen Nettoertrag von 290 €. In einem Fall von 1000 hat sie einen Nettoabgang von 9710 € (= 290 – 10000).

Der durchschnittliche Gewinn ist daher 280 €:

$$\frac{999}{1000} 290 + \frac{1}{1000} (-9710) = 0.999 \cdot 290 + 0.001 \cdot (-9710) = 280$$

Erwartungswert / Beispiel – Fortsetzung

Die Rechenschritte, die in der Formel

$$E(X) = \sum_x x P(X = x) = \mu_x$$

angegeben sind, kann man in die Tabelle der Wahrscheinlichkeitsverteilung leicht integrieren.

x	$P(X = x)$	$x P(X = x)$	
290	0.999	$290 \cdot 0.999 =$	289.71
-9710	0.001	$-9710 \cdot 0.001 =$	-9.71
Summe	1.000		280.00
			= $E(X)$

Erwartung einer Funktion einer ZV

Der Erwartungswert einer Funktion g einer diskreten Zufallsvariable X , $g(X)$, ist

$$E[g(X)] = \sum_x g(x) f(x) = \sum_x g(x) P(X = x)$$

Angenommen Sie erhalten 10% Gewinnbeteiligung obiger Versicherung abzüglich 10 € Bearbeitungsgebühr. Ihr erwarteter Gewinn beträgt dann

$$\begin{aligned} E(G) &= E[g(X)] = E[0.1 X - 10] = \\ &= (0.1 \cdot 290 - 10) \cdot 0.999 + (-0.1 \cdot 9710 - 10) \cdot 0.001 = 18 \end{aligned}$$

Erwartungswert / Rechenregeln

Sei $Y = a + bX$, dann gilt allgemein für den Erwartungswert von Y

$$E(Y) = E(a + bX) = a + bE(X)$$

Die Erwartung einer Linearkombination von ZVen ist gleich die Linearkombination der Erwartungswerte der einzelnen ZVen.

Bemerkung: Der Erwartungswert von X muss existieren.

Achtung!

Ist f nicht linear, kann Erwartung und Funktion nicht vertauscht werden. Im allgemeinen gilt

$$E[f(X)] \neq f(E[X])$$

Varianz

Die **Varianz** einer ZV X , $V(X) = \sigma_x^2$, ist definiert als die erwartete quadratische Abweichung vom Erwartungswert (mittlere quadratische Abweichung vom Mittel).

$$V(X) = E([X - E(X)]^2) = E([X - \mu]^2) = \sigma_x^2$$

Die **Standardabweichung** ist ein Maß für die Streuung:

$$\sqrt{V(X)} = \sqrt{\sigma_x^2} = \sigma_x$$

Die Standardabweichung wird oft als Maß für die Unsicherheit über den Ausgang eines Experiments verwendet.

Varianz / Verschiebungssatz

Der **Verschiebungssatz** erleichtert die Berechnung der Varianz. Er lautet mit $E(X) = \mu$:

$$V(X) = E([X - \mu]^2) = E(X^2) - \mu^2$$

bzw.

$$V(X) = \sum_x ([x - \mu]^2) P(X = x) = \sum_x x^2 P(X = x) - \mu^2$$

$$V(X) = \sum_x ([x - \mu]^2) f(x) = \sum_x x^2 f(x) - \mu^2$$

Varianz / Rechenregeln

Sei $Y = a + bX$, dann gilt allgemein für die Varianz von Y

$$V(Y) = V(a + bX) = b^2 V(X)$$

Die Varianz ist verschiebungsinvariant. Sie hängt nicht von der Konstanten a ab.

- Die Varianz von Y steigt linear mit b^2 ,
- die Standardabweichung von Y steigt linear mit b .

Bemerkung: Erwartungswert und Varianz von X müssen existieren.

Stetige Zufallsvariable

- Ganze Zahlen (gerundet), Bruchzahlen oder reelle Zahlen.
- Erhält man bei Messungen.
- Unendlich viele mögliche Werte in einem Intervall.
Zu viele, um sie wie bei diskreten Zufallsvariablen aufschreiben zu können.
- Wird durch die **Dichtefunktion** oder **Verteilungsfunktion** beschrieben.

Dichtefunktion

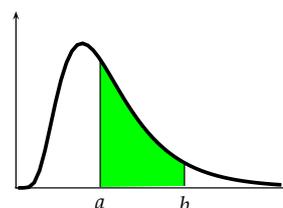
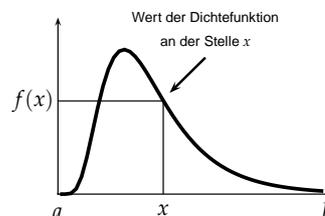
- Ein Graph, der alle x die „Häufigkeit“ $f(x)$ zeigt.
 $f(x)$ ist keine Wahrscheinlichkeit.

- Eigenschaften:

- $\int_{-\infty}^{\infty} f(x) dx = 1$.
Die Fläche unter der Kurve ist 1.
- $f(x) \geq 0$.

- Wahrscheinlichkeiten können nur für Intervalle angegeben werden:

$$P(a < X \leq b) = \int_a^b f(x) dx$$



Verteilungsfunktion

Gibt die Wahrscheinlichkeit an, dass die Zufallsvariable nicht größer als ein vorgegebener Wert ist:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(y) dy$$

Die Verteilungsfunktion ist monoton in x .

Mit Hilfe der Verteilungsfunktion lassen sich Wahrscheinlichkeiten ausrechnen:

$$P(a < X \leq b) = F(b) - F(a)$$

Erwartungswert und Varianz

- Erwartungswert

$$E(X) = \mu_x = \int_{-\infty}^{\infty} x f(x) dy$$

- Varianz

$$V(X) = \sigma_x^2 = \int_{-\infty}^{\infty} (x - \mu_x)^2 f(x) dy = \int_{-\infty}^{\infty} x^2 f(x) dy - \mu_x^2$$

- Standardabweichung

$$\sigma_x = \sqrt{V(X)}$$

Spezielle Verteilungen

Viele Merkmale folgen der gleichen (oder ähnlichen) Verteilungen. Diese sind daher besonders wichtig und gut untersucht.

- Jede Verteilung beschreibt ein zugrundeliegendes Phänomen (**Modell**).
- Für Varianten derselben Fragestellung sind nur die Parameter der Verteilung anzupassen.
- Für bestimmte Fragestellungen können die Verteilungen – die mathematische Formeln – explizit angegeben werden.
- Wahrscheinlichkeiten, Erwartungswert, Varianz, . . . , können dann exakt berechnet werden.

Binomialverteilung

Beschreibt die Zahl der „Erfolge“ in einer Stichprobe von Umfang n

- Zufallsvariable heißt: Zahl der „Erfolge“ unter n (unabhängigen) Beobachtungen (Versuchen).

Beispiele:

- Anzahl „Kopf“ bei zehnmalem Werfen einer Münze.
- Anzahl der richtigen Antworten bei MC-Test mit 20 Fragen.
- Anzahl defekter Teile bei Qualitätskontrolle in einer Kiste mit 50 Stück.
- Anzahl der erfolgreichen Verkaufsgespräche bei 100 geführten.

Binomialverteilung / Modellvoraussetzungen

- Folge von n **identischen, unabhängigen** Versuchen.
- Jeder Versuch hat genau zwei mögliche Ausgänge: „**Erfolg**“ oder „**Mißerfolg**“.
- Die Wahrscheinlichkeit für „Erfolg“, bzw. „Mißerfolg“, ist für alle Versuche gleich.
- Zwei Stichprobenauswahlverfahren dazu:
 - **Unendliche** Grundgesamtheit **ohne** Zurücklegen, oder
 - **Endliche** Grundgesamtheit **mit** Zurücklegen

Binomialverteilung / Wahrscheinlichkeitsfunktion

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

$f(x)$... Wahrscheinlichkeit für x „Erfolge“, $P(X = x) = f(x)$

n ... Anzahl der Wiederholungen (Stichprobengröße)

p ... Wahrscheinlichkeit für „Erfolg“

x ... Anzahl an „Erfolgen“ ($x = 0, 1, 2, \dots, n$)

$\binom{n}{x}$... Binomialkoeffizient. Man sagt: „ n über x “.

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

Binomialverteilung / Erwartungswert

- Erwartungswert

$$\mu = E(X) = np$$

- Standardabweichung

$$\sigma = \sqrt{np(1-p)}$$

Binomialverteilung / Beispiel

Experiment: Münze wird 5 mal geworfen. Beobachte Anzahl „Kopf“.

Wie groß ist die Wahrscheinlichkeit für 3 mal „Kopf“?

Antwort: $n = 5$, $p = 0.5$, $x = 3$, $P(X = 3) = f(3)$,

$$\begin{aligned} f(x) &= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\ f(3) &= \frac{5!}{3!(5-3)!} 0.5^3 (1-0.5)^{5-3} \\ &= 10 \cdot 0.5^3 \cdot 0.5^2 \\ &= 0.3125 \end{aligned}$$

Binomialverteilung / Beispiel

Bei einem Multiple-Choice-Test gibt es 20 Fragen mit je 5 Antwortmöglichkeiten, wobei immer nur (genau) eine richtig ist.

Wenn Sie keine einzige Frage beantworten können und daher bei jeder Antwort raten, wie groß ist dann die Wahrscheinlichkeit, dass Sie den Test bestehen? Mindestens 10 Antworten müssen richtig sein.

Wie lauten Erwartungswert und Standardabweichung für die Anzahl an richtigen Antworten?

Binomialverteilung / Lösung

Die Zufallsvariable, X , heißt: „Anzahl der Erfolge unter 20“. Prüfung bestanden heißt: 10, 11, 12, ..., 19, oder 20 Antworten sind richtig. Die Anzahl der richtigen Antworten ist binomial verteilt mit $n = 20$ und $p = 0.2$. Daher

$$\begin{aligned} P(\text{bestanden}) &= P(X \geq 10) = f(10) + f(11) + \dots + f(20) \\ &\approx 2.03 \cdot 10^{-3} + 4.62 \cdot 10^{-4} + \dots + 1.05 \cdot 10^{-14} \\ &\approx 0.0026 \end{aligned}$$

Erwartungswert und Standardabweichung

$$\begin{aligned} \mu &= np = 20 \cdot 0.2 = 4 \\ \sigma &= \sqrt{np(1-p)} = \sqrt{20 \cdot 0.2 \cdot (1-0.2)} = 1.79 \end{aligned}$$

Poissonverteilung

Beschreibt die Anzahl von Ereignissen innerhalb eines Intervalls.

- Zufallsvariable heißt: Anzahl von Ereignissen je Einheit (Zeit, Länge, Fläche, Raum, ...).

Beispiele:

- Anzahl von Kunden, die innerhalb von 20 Minuten eintreffen.
- Anzahl von Streiks pro Jahr.
- Anzahl an Schlaglöchern pro Straßenkilometer.

Poissonverteilung / Modellvoraussetzungen

- Die Wahrscheinlichkeit für das Eintreten eines Ereignisses ist **konstant**.
 - Angenommen pro Stunde kommen im Durchschnitt 120 Kunden. So bedeutet dies, dass in jeder der 60 Minuten im Durchschnitt 2 Kunden eintreffen.
- Das Eintreten der einzelnen Ereignisse ist **unabhängig**.
 - Das Eintreffen eines Kunden beeinflusst nicht das Eintreffen eines anderen Kunden.
 - Das Auftreten von Schlaglöchern in einem Strassenabschnitt beeinflusst nicht andere Strassenabschnitte.
- Ein Ereignis pro Einheit.
 - Es können nicht 2 Kunden zur selben Zeit eintreffen.

Poissonverteilung / Wahrscheinlichkeitsfunktion

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

$f(x)$... Wahrscheinlichkeit für x „Erfolge“, $f(x) = P(X = x)$

λ ... Erwartete (durchschnittliche) Anzahl von „Erfolgen“

e ... Eulersche Zahl (2.71828...)

x ... Anzahl an „Erfolgen“ pro Einheit ($x = 0, 1, 2, \dots, \infty$)

$x!$... x -Faktorielle. $x! = x \cdot (x - 1) \cdot (x - 2) \cdot \dots \cdot 2 \cdot 1$,

Poissonverteilung / Erwartungswert

- Erwartungswert

$$\mu = E(X) = \lambda$$

- Standardabweichung

$$\sigma = \sqrt{\lambda}$$

Poissonverteilung / Beispiel

In einer Bankfiliale treffen pro Stunde durchschnittlich 22 Kunden pro Stunde ein.

Wie groß ist die Wahrscheinlichkeit, dass innerhalb von 15 Minuten genau 4 Kunden eintreffen?

Wie groß sind Erwartungswert und Standardabweichung für die Anzahl der Kunden in einem 15-Minuten-Intervall?

Poissonverteilung / Lösung

Anzahl der Kunden pro 15-Minuten-Intervall ist poisson verteilt mit

$$\lambda = \frac{22}{4} = 5.5.$$

(22 Kunden pro Stunde = 5.5 Kunden pro 15 Minuten).

Wahrscheinlichkeit, dass genau 4 Kunden eintreffen:

$$f(x) = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$
$$f(4) = P(X = 4) = \frac{5.5^4 e^{-5.5}}{4!} = 0.156$$

Erwartungswert und Standardabweichung

$$\mu = \lambda = 5.5$$
$$\sigma = \sqrt{\lambda} = 2.345$$

Stetige Gleichverteilung

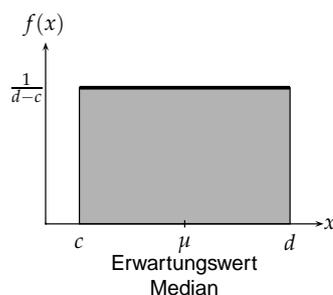
- Gleichwahrscheinliche Realisationen (Ergebnisse).

- **Dichtefunktion**, $c < x < d$:

$$f(x) = \frac{1}{d-c}$$

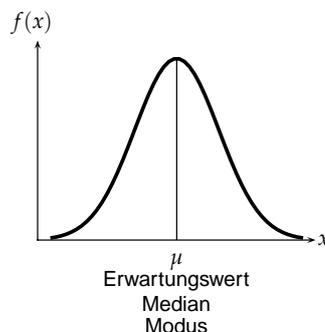
- **Erwartungswert und Standardabweichung**

$$\mu = \frac{c+d}{2}, \quad \sigma = \frac{d-c}{\sqrt{12}}$$



Normalverteilung

- Einfache mathematische Eigenschaften.
- Beschreibt gut viele stochastische (zufällige) Prozesse und stetige Phänomene.
- Approximiert gut andere (schwierig zu berechnende) Verteilungen.
- Ist Basis für klassische statistische Inferenz (statistisches Schließen).



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right]$$

$f(x)$... Wert der Dichte der Zufallsvariablen X an der Stelle x

μ ... Erwartungswert, Mittel der Grundgesamtheit

σ ... Standardabweichung in der Grundgesamtheit

π ... 3.14159...

e ... Eulersche Zahl (2.71828...), $\exp(x) = e^x$

x ... Wert der Zufallsvariablen X , ($-\infty < x < \infty$)

Die Normalverteilung hat 2 Parameter: μ und σ .

Approximation der Binomialverteilung

Die Binomialverteilung mit Parametern n (Wiederholungen) und p (Wahrscheinlichkeit für „Erfolg“) kann durch die Normalverteilung approximiert werden, falls gilt (Fausregel!)

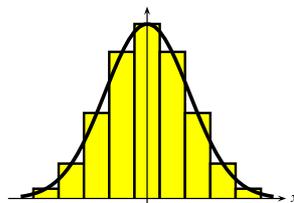
$$np(1-p) \geq 9$$

- Erwartungswert

$$\mu_p = np$$

- Standardabweichung

$$\sigma_p = \sqrt{\frac{p(1-p)}{n}}$$



Zentraler Grenzwertsatz

Seien X_i ($i = 1, \dots, n$) unabhängige stetige Zufallsvariable mit Mittelwert μ und Standardabweichung σ .

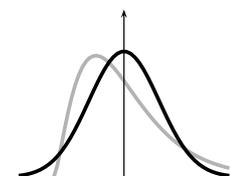
Wir betrachten die Zufallsvariable (arithmetisches Mittel)

$$\bar{X} = \frac{1}{n} \sum X_i$$

Bei hinreichend großem n (Stichprobengröße) ist \bar{X} annähernd normalverteilt mit

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

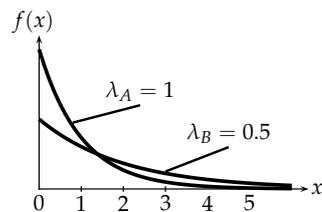
Faustregel: $n \gtrsim 30$.



Exponentialverteilung

- Parameter: λ
- Dichtefunktion, $x \geq 0$:

$$f(x) = \lambda \exp(-\lambda x)$$



- Verteilungsfunktion:

$$F(x) = P(X \leq x) = 1 - \exp(-\lambda x)$$

- Erwartungswert und Standardabweichung:

$$\mu = \frac{1}{\lambda} \quad \sigma = \frac{1}{\lambda}$$

Exponentialverteilung

Beschreibt Zeit oder Distanz zwischen Ereignissen.

- Wird in Warteschlangenmodellen verwendet.

Kapitel 4

Versicherungsmathematik

Lernziele

- Sterbetafel, Überlebenswahrscheinlichkeiten, Mittlere (Rest-)Lebensdauer
- Zeitrente
- Lebensversicherung
- Ablebensversicherung
- Erlebensversicherung
- (Sachversicherung wird nicht behandelt)

Wirtschaftliche Bedeutung

Anteil am veranlagten Barvermögen in Österreich:

	1970	1995	2003
Bargeld/Spareinlagen	77%	59%	49%
Wertpapiere	12%	23%	23%
Pension/Lebensversicherung	3%	13%	19%
sonstige	8%	5%	9%

Gesamtvolumen 2003:

Prämien: 13,2 Mrd. €

Versicherungsleistung: 9,7 Mrd. €

versicherungstechnische Rückstellung: 50,2 Mrd. €

Versicherung: Definition

Versicherung (*Assekuranz*) ist,

- die gegenseitige Deckung (im Gegensatz zur „Selbst“versicherung)
- eines im einzelnen zufälligen im ganzen aber schätzbaren (im Sinne einer ZV mit bekannter Verteilung)
- durch eine Vielzahl gleichartig bedrohter Wirtschaftseinheiten. (dadurch wird i.a. die Verteilung bekannt)

Die Versicherungsmathematik verwendet eigene Bezeichnungen.

x	Anzahl der vollendeten Jahre (Alter abgerundet).
l_x	Anzahl der Personen des Alters x (Anz. der Pers, die mindestens x Jahre alt geworden sind, bezogen auf $l_0 = 100\,000$.)
$d_x = l_x - l_{x+1}$	Anzahl der Personen, die zwischen dem Alter x und $x + 1$ sterben.
d_{x+n} $= l_{x+n} - l_{x+n+1}$	Anzahl der x -jährigen, die zwischen dem Alter $x + n$ und $x + n + 1$ sterben.

l steht für „life“, d für „death“.

Sterbetafel 1990/92 für Österreich

x	Männer		Frauen	
	l_x	d_x	l_x	d_x
0	100 000	847	100 000	671
1	99 153	54	99 329	51
2	99 099	45	99 278	41
3	99 054	37	99 237	32
4	99 017	31	99 205	24
...	
19	98 468	148	98 942	40
20	98 319	144	98 902	40
21	98 175	135	98 862	38
...	

(Q: ÖSTAT)

Lebensalter

Wir betrachten eine zufällig ausgewählte Person.

Aus der Sicht einer Versicherung ist das **Lebensalter** L , das diese Person erreichen wird, eine Zufallsvariable.

Die Wahrscheinlichkeit, dass diese Person mindestens k Jahre alt wird, beträgt

$$P(L \geq k) \approx \frac{l_k}{100\,000}$$

Das „ \approx “ Zeichen zeigt an, dass die Werte l_k empirisch in einem bestimmten Jahr erhoben wurden. Diese Werte unterliegen Veränderungen und Schwankungen. Z.B.:

- Jahre sind unterschiedlich: heißer Sommer / kühler Sommer,
- Geburtenschwache Jahrgänge,
- Steigende Lebenserwartung, etc.

Restliche Lebensdauer

Unsere zufällig ausgewählte Person ist heute x Jahre alt ist. Dann ist auch die **restliche Lebensdauer** R_x dieser Person eine Zufallsvariable.

$$R_x = L - x \quad \text{bzw.} \quad L = x + R_x$$

Die Wahrscheinlichkeit, dass diese Person noch mindestens k Jahre lebt, insgesamt also mindestens $x + k$ Jahre alt wird, ist die bedingte Wahrscheinlichkeit

$$P(R_x \geq k) = P(L \geq x + k | L \geq x) \approx \frac{l_{x+k}}{l_x}$$

Bemerkung: Es gilt immer $l_x > l_{x+k}$.

Sterbetafel: Wahrscheinlichkeit p_x

Die Wahrscheinlichkeit $P(R_x \geq k)$, dass eine x -jährige Person in k Jahren noch lebt, wird mit ${}_k p_x$ bezeichnet.

$${}_k p_x = P(L \geq x + k | L \geq x) \approx \frac{l_{x+k}}{l_x}$$

Für $k = 1$ schreiben wir $p_x = {}_1 p_x \approx \frac{l_{x+1}}{l_x}$.

$${}_k p_x = p_x \cdot p_{x+1} \cdot p_{x+2} \cdot \dots \cdot p_{x+k-1}$$

Sterbetafel: Wahrscheinlichkeit q_x

Die Wahrscheinlichkeit, dass eine x -jährige Person in k Jahren nicht mehr lebt, wird mit ${}_k q_x$ bezeichnet.

$${}_k q_x = P(R_x < k) = 1 - {}_k p_x \approx 1 - \frac{l_{x+k}}{l_x} = \frac{l_x - l_{x+k}}{l_x}$$

Für $k = 1$ schreiben wir

$$q_x = {}_1 q_x = 1 - p_x \approx \frac{l_x - l_{x+1}}{l_x} = \frac{d_x}{l_x}$$

Beispiel

Die Wahrscheinlichkeit, dass ein 60-jähriger Mann in 1, 3 oder 10 Jahren noch lebt beträgt:

- ${}_1p_{60} = \frac{l_{60+1}}{l_{60}} = \frac{82\,294}{83\,586} = 0.9845$
- ${}_3p_{60} = \frac{l_{63}}{l_{60}} = \frac{79\,400}{83\,586} = 0.9499$
- ${}_{10}p_{60} = \frac{l_{70}}{l_{60}} = \frac{65\,781}{83\,586} = 0.78698$

Die Wahrscheinlichkeit, dass ein 60-jähriger Mann in 1, 3 oder 10 Jahren nicht mehr lebt beträgt:

- ${}_1q_{60} = \frac{l_{60} - l_{60+1}}{l_{60}} = \frac{83\,586 - 82\,294}{83\,586} = 0.01545$
- ${}_3q_{60} = \frac{l_{60} - l_{63}}{l_{60}} = \frac{83\,586 - 79\,400}{83\,586} = 0.05008$
- ${}_{10}q_{60} = \frac{l_{60} - l_{70}}{l_{60}} = \frac{83\,586 - 65\,781}{83\,586} = 0.21301$

Sterbetafel / Diskussion

Die Sterbetafel geht davon aus, dass ein Neugeborener vor 50 Jahren die selbe Lebenserwartung besitzt wie ein Neugeborener, der heute auf die Welt kommt. Medizinischer Fortschritt, geänderte Ernährung, etc. werden nicht berücksichtigt.

Es zählt nur das bereits erreichte Alter, nicht aber das konkrete Geburtsjahr.

Alternativen sind Sterbetafeln für jedes Geburtsjahr. Dafür fehlen aber die Beobachtungen. Daher können sie nur unter bestimmten Annahmen hochgerechnet werden.

Sterbetafel: Wahrscheinlichkeit ${}_k p_x q_{x+k}$

Die Wahrscheinlichkeit, dass eine x -jährige Person noch k , aber nicht mehr als $k + 1$ Jahre lebt (d.h., dass sie zwischen $x + k$ und $x + k + 1$ stirbt), beträgt ${}_k p_x q_{x+k}$.

$${}_k p_x q_{x+k} \approx \frac{d_{x+k}}{l_x} = \frac{l_{x+k} - l_{x+k+1}}{l_x}$$

$$\begin{aligned} {}_k p_x q_{x+k} &= P(L = x + k | L = x) \\ &= P(L \geq x + k | L = x) \cdot P(L < x + k + 1 | L = x + k) \\ &= \frac{l_{x+k}}{l_x} \left(1 - \frac{l_{x+k+1}}{l_{x+k}} \right) = \frac{l_{x+k}}{l_x} \cdot \frac{l_{x+k} - l_{x+k+1}}{l_{x+k}} \\ &= \frac{l_{x+k} - l_{x+k+1}}{l_x} = \frac{d_{x+k}}{l_x} \end{aligned}$$

Beispiel

Wie groß ist die Wahrscheinlichkeit, dass

- ein 50jähriger Mann noch mindestens 1 Jahr lebt?

$${}_1p_{50} = \frac{l_{50+1}}{l_{50}} = \frac{91\,350}{91\,899} = 0.994026$$

- eine 70jährige Frau noch höchstens 5 Jahre lebt?

$${}_6q_{70} = \frac{l_{70} - l_{70+6}}{l_{70}} = \frac{82\,461 - 70\,668}{82\,461} = 0.14301$$

- ein 85jähriger Mann noch genau 2 Jahre lebt?

$${}_2p_{85} q_{87} = \frac{d_{85+2}}{l_{85}} = \frac{l_{85+2} - l_{85+2+1}}{l_{85}} = \frac{14\,003 - 11\,528}{19\,738} = 0.125$$

(laut Sterbetafel 1990/92 für Österreich)

Mittlere Lebensdauer

Der Erwartungswert der restlichen Lebensdauer einer x -jährigen Person, $E(R_x)$, ist

$$E(R_x) \approx E(\hat{R}_x) + 1/2 = \sum_{k=0}^{\infty} k {}_k p_x q_{x+k}$$

$1/2$ wird addiert, da die Personen, die zwischen dem k -ten und $(k+1)$ -ten Jahr sterben im Mittel (etwa) $k + 1/2$ Jahre alt werden.

$$E(R_x) \approx \frac{1}{l_x} \sum_{k=1}^{\infty} l_{x+k} + 1/2$$

$E(R_x)$ ist ein bedingter Erwartungswert. Er ist die Restlebenserwartung einer Person, die bereits x Jahre gelebt hat.

Mittlere Lebensdauer / Herleitung

$$\begin{aligned} E(\hat{R}_x) &= 1 {}_1 p_x q_{x+1} + 2 {}_2 p_x q_{x+2} + 3 {}_3 p_x q_{x+3} + \dots \\ &= 1 \frac{l_{x+1} - l_{x+2}}{l_x} + 2 \frac{l_{x+2} - l_{x+3}}{l_x} + 3 \frac{l_{x+3} - l_{x+4}}{l_x} + \dots \\ &= \frac{1}{l_x} (l_{x+1} - l_{x+2} + 2l_{x+2} - 2l_{x+3} + 3l_{x+3} - 3l_{x+4} + \dots) \\ &= \frac{1}{l_x} (l_{x+1} + l_{x+2} + l_{x+3} + \dots) \\ &= \frac{1}{l_x} \sum_{k=1}^{\infty} l_{x+k} \end{aligned}$$

Mit der Hand ist die Berechnung der mittleren Lebensdauer mühsam. Spreadsheets (wie z.B. EXCEL) hingegen scheint dafür konzipiert worden zu sein.

Zeitrente

Der Barwert einer n Jahre lang jährlich zahlbaren vorschüssigen **Zeitrente** von 1 GE und Zinssatz (*interest rate*) i ist gegeben durch (geometrische Reihe):

$$\ddot{a}_{\overline{n}|} = B_n = \sum_{k=0}^{n-1} v^k = \frac{1 - v^n}{1 - v} = \frac{1 - v^n}{d}$$

wobei $v = \frac{1}{1+i}$ der *Diskontfaktor* und $d = \frac{i}{1+i} = 1 - v$.

Ewige vorschüssige Rente: $\ddot{a}_{\infty|} = \frac{1}{d}$

Nachschüssige Rente: $a_{\overline{n}|} = vB_n = \sum_{k=1}^n v^k = \frac{1 - v^n}{v}$

Zeitrente / Beispiel

Ein 60-jähriger Mann möchte bis zu seinem 90. Geburtstag jedes Jahr 10 000 € erhalten. Berechnen Sie den Wert dieser Zeitrente bei einem Zinssatz von $i = 3\%$.

(Und wir nehmen, dass an der Mann wird wirklich so alt wird.)

$$v = \frac{1}{1+0.03} = 0.97087$$

$$\ddot{a}_{\overline{30}|} = \frac{1 - v^n}{1 - v} = 20.18845$$

Der Wert der Zeitrente beträgt daher

$$20.18845 \cdot 10\,000 \text{ €} = 201\,885 \text{ €}.$$

Leibrente

Eine **Leibrente** ist eine Zeitrente, die bis zum Tod einer bestimmten Person gezahlt wird.

Die Dauer der Zahlungen ist eine Zufallsvariable; somit auch der davon abhängende Barwert der Leibrente.

Der Wert der Leibrente (= Nettoeinmalprämie) ist gleich dem Erwartungswert dieser Zeitrente.

Leibrente / Barwert

Vorschüssige, lebenslängliche Leibrente, jährliche Zahlung 1 GE:

- Zahlungen zu den Zeitpunkten: $0, 1, \dots, K$.
- Barwert:

$$Y = 1 + v + v^2 + \dots + v^K = \ddot{a}_{\overline{K+1}|}$$

- Verteilung (für x -jährige Person):

$$P(Y = \ddot{a}_{\overline{K+1}|}) = P(K = k) = {}_k p_x q_{x+k} \quad k = 0, 1, 2, \dots$$

- Wert der Leibrente: **erwarteter Barwert**

$$\ddot{a}_x = E(Y) = \sum_{k=0}^{\infty} \ddot{a}_{\overline{k+1}|} {}_k p_x q_{x+k}$$

bzw.

$$\ddot{a}_x = \sum_{k=0}^{\infty} v^k {}_k p_x$$

Leibrente / Beispiel

Zinssatz: $i = 3\%$

$$\ddot{a}_{30} = 24.42$$

$$\ddot{a}_{50} = 17.82$$

$$\ddot{a}_{60} = 13.78$$

Wieviel „Rente“ bekommt ein 30-jähriger Mann für 100 000 €?
(Zinssatz $i = 3\%$)

$$\frac{100\,000}{\ddot{a}_{30}} = \frac{100\,000}{24.42} = 4095$$

Temporäre Leibrente

Eine **temporäre Leibrente** wird nur für eine maximale Anzahl n an Jahren (Perioden) ausbezahlt.

$$Y = \begin{cases} \ddot{a}_{\overline{K+1}|} & \text{für } K = 0, 1, \dots, n-1 \\ \ddot{a}_{\overline{n}|} & \text{für } K = n, n+1, \dots \end{cases}$$

$$\ddot{a}_{x:\overline{n}|} = \sum_{k=0}^{n-1} v^k {}_k p_x$$

Temporäre Leibrente für 50-jährigen Mann für maximal 10 Jahre,
Zinssatz $i = 3\%$:

$$\ddot{a}_{50:\overline{10}|} = \sum_{k=0}^9 v^k {}_k p_{50} = 8.50$$

Aufgeschobene Leibrente

Eine m Jahre aufgeschobene, vorschüssige Leibrente mit jährlichen Zahlungen von 1 GE wird mit ${}_m|\ddot{a}_x$ bezeichnet.

$$Y = \begin{cases} 0 & \text{für } K = 0, 1, \dots, m-1 \\ v^m + v^{m+1} + \dots + v^K & \text{für } K = m, m+1, \dots \end{cases}$$

$${}_m|\ddot{a}_x = v^m {}_m p_x \ddot{a}_{x+m} = \ddot{a}_x - \ddot{a}_{x:\overline{m}|}$$

Um 8 Jahre aufgeschobene Leibrente für 60-jährigen Mann, Zinssatz $i = 3\%$:

$${}_8|\ddot{a}_{60} = v^8 {}_8 p_{60} \ddot{a}_{68} = 6.98$$

Kommutationszahlen

Bei der Berechnung von versicherungstechnischen Größen müssen oft Summen berechnet werden. Das kann durch sogenannte **Kommutationszahlen** vereinfacht werden. Summen werden dann durch einfache Rechenoperationen ersetzt.

Beispiel: $\ddot{a}_x = \sum_{k=0}^{\infty} v^k {}_k p_x$

Wir setzen $D_x = v^x l_x$ und $N_x = \sum_{i=x}^{\infty} D_i$ und erhalten $\ddot{a}_x = \frac{N_x}{D_x}$.

Analog:

$$\ddot{a}_{x:\overline{n}|} = \sum_{k=0}^{n-1} v^k {}_k p_x = \frac{N_x - N_{x+n}}{D_x}$$

Kommutationszahlen / Übersicht

l_x	Zahl der Überlebenden des Alters x
$D_x = v^x l_x$	diskontierte Zahl der Überlebenden
$d_x = l_x - l_{x+1}$	Anzahl der im $(x+1)$ -ten Lebensjahr verstorbenen
$C_x = v^{x+1} d_x$	diskontierte Zahl der Verstorbenen
$N_x = \sum_{i=x}^{\infty} D_i$	Summe der diskontierten Zahl der Überlebenden
$S_x = \sum_{i=x}^{\infty} N_i$	Summe der diskontierten Zahl der Überlebenden 2. Ordnung
$M_x = \sum_{i=x}^{\infty} C_i$	Summe der diskontierten Zahl der Verstorbenen
$R_x = \sum_{i=x}^{\infty} M_i$	Summe der diskontierten Zahl der Verstorbenen 2. Ordnung

Ablebensversicherung

Todesfallversicherung: Stirbt der Versicherungsnehmer während der vertraglichen Laufzeit, wird ein festgelegter Betrag (an die Erben) ausbezahlt. (Versicherungsnehmer, die nicht versterben, finanzieren die Auszahlung mit.)

- Lebenslängliche Deckung. Der Betrag 1 wird am Ende des Todesjahres ausbezahlt. Erwarteter Barwert:

$$A_x = \sum_{k=0}^{\infty} v^{k+1} {}_k p_x q_{x+k} = \frac{M_x}{D_x}$$

- n -jährige temporäre Todesfallversicherung.

$${}_n A_x = \frac{M_x - M_{x+n}}{D_x}$$

Erlebensversicherung

Versicherung auf den Erlebensfall.

Falls die/der Versicherungsnehmer/in man das Alter von $x + n$ Jahren erreicht wird das Kapital von 1 ausgezahlt.

(Versicherungsnehmer, die in der Zwischenzeit versterben, finanzieren die Auszahlung mit.)

Erwarteter Barwert:

$${}_n E_x = v^n {}_n p_x = \frac{D_{x+n}}{D_x}$$

Gemischte Versicherung

Ab- und Erlebensversicherung

Das Kapital von 1 wird am Ende des Todesjahres ausbezahlt, wenn der Tod in den ersten n Jahren eintritt, andernfalls nach Ablauf der Versicherungsdauer n .

Erwarteter Barwert:

$$A_{x:\overline{n}|} = {}_n A_x + {}_n E_x = \frac{M_x - M_{x+n} + D_{x+n}}{D_x}$$

Versicherungsprämie P_N

Für die Nettoprämie muss gelten:

Barwert der Einzahlungen = Barwert der Leistungen

Zwei wichtigsten Prämienarten:

- Einmalzahlung am Beginn der Versicherung (NEP).
- Vorschüssige (jährliche) laufende Zahlungen:
(temporäre) Leibrente.

Versicherungsprämie P_N / Beispiel

Todesfallversicherung, befristet auf n Jahre für eine x -jährige Person, mit Versicherungssumme C :

$$C {}_nA_x = P_N \ddot{a}_{x:\overline{n}|}$$

Jährliche Nettoprämie P :

$$P_N = C \frac{{}_nA_x}{\ddot{a}_{x:\overline{n}|}} = C \frac{M_x - M_{x+n}}{N_x - N_{x+n}}$$

Beispiel:

Todesfallversicherung, befristet auf 10 Jahre für einen 40-jährigen Mann, mit Versicherungssumme 100 000 € (Zinssatz 3%):

$$P_N = 100\,000 \frac{M_{40} - M_{50}}{N_{40} - N_{50}} = 10^5 \frac{11\,010.30 - 10\,076.43}{627\,605.19 - 373\,763.99} = 367.90 \text{ €}$$

Bruttoprämie

Folgende weiteren Kosten müssen zur Berechnung der Bruttoprämie berücksichtigt werden:

- α Abschlusskosten bei Versicherungsbeginn
- β Verwaltungskosten während der Zahlungsdauer der Beiträge
- γ Verwaltungskosten während der gesamten Versicherungsdauer
- σ Sicherheitszuschläge

Gebräuchlichstes Verfahren:

$$\text{Bruttoprämie } P_B = P_N + P_\alpha + P_\beta + P_\gamma + P_\sigma$$

$$P_\alpha = \alpha$$

$$P_\beta = \beta P_B$$

$$P_\gamma = \gamma$$

$$P_\sigma = \sigma P_B$$

$$\Rightarrow P_B = \frac{P_N + \alpha + \gamma}{1 - \beta - \sigma}$$

Bruttoprämie / Beispiel

Gemischte Versicherung:

$$P_{x:\overline{n}} \ddot{a}_{x:\overline{n}} = A_{x:\overline{n}} + \alpha + \beta P_{x:\overline{n}} \ddot{a}_{x:\overline{n}} + \gamma \ddot{a}_{x:\overline{n}}$$

$$P_{x:\overline{n}} = \frac{A_{x:\overline{n}} + \alpha + \gamma \ddot{a}_{x:\overline{n}}}{(1 - \beta) \ddot{a}_{x:\overline{n}}}$$

(Sicherheitszuschlag durch vorsichtige Wahl der Rechnungsgrundlagen)

Deckungskapital ${}_tV$

Zum Zeitpunkt $t = 0$ sind Barwert der Leistungen und Barwert der Zahlungen gleich.

In der Regel verändert sich diese Bilanz für $t > 0$:
Es entsteht eine Differenz!

Diese Differenz ist wieder eine Zufallsgröße:

(Netto-) Deckungskapital ${}_tV$

Rückkaufwert RW

Der **Rückkaufwert** ist der Betrag, den der Versicherungsnehmer bei Kündigung des Versicherungsvertrags zurückerstattet bekommt.

Wert der Versicherung bei Kündigung (Deckungskapital) minus Stornoabzug.

(Verluste im Portfolio, zusätzliche Verwaltungskosten, reduzierte Kapitalerträge auf Grund der nötigen erhöhten Liquiditätserhaltung, etc.)

$$RW < {}_tV$$

Risikoabschätzung – Ganz einfaches Beispiel

Bei einer Versicherung gibt es 10 000 Versicherungsnehmer, die alle eine einjährige Ablebensversicherung mit einer Versicherungssumme von 100 000 € abgeschlossen haben.

Das Sterberisiko betrage bei allen versicherten Personen 0.1%.

(Es ist zu erwarten, dass circa 10 Personen sterben.)

Wie groß ist das Risiko bzw. die Wahrscheinlichkeit, dass mehr als 15, 25 bzw. 100 Personen sterben?

Mathematische Formulierung:

X ... Anzahl der verstorbenen Versicherungsnehmer

$P(X > 15) = ?$, $P(X > 25) = ?$, $P(X > 100) = ?$

⇒ Binomialverteilung

(wird mittels Normalverteilung approximiert)

Weiterführende Literatur

- Milbrodt, Hartmut, und Helbig, Manfred (1999): *Mathematische Methoden der Personenversicherung*, de Gruyter, Berlin, New York
- Gerber, Hans (1986): *Lebensversicherungsmathematik*, Springer, Berlin
- Bosch, Karl (1990): *Finanzmathematik*, Kapitel 7 (Versicherungsmathematik), Oldenbourg, München.
- Grundmann, Wolfgang (1986): *Finanz- und Versicherungsmathematik*, B.G. Teubner Verlagsgesellschaft, Stuttgart

Kapitel 5

Monotonie und Krümmung

Lernziele

Monotonie und Krümmungsverhalten von Funktionen in einer Variablen

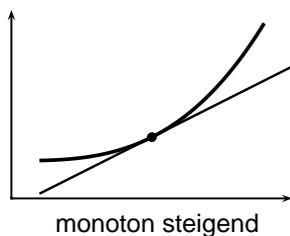
- Monotone Funktionen
- Konkave und konvexe Funktionen
- Monotonie, Krümmung und Ableitungen

Monotonie

Eine Funktion f heißt **monoton steigend** (fallend), falls

$$x_1 \leq x_2 \Leftrightarrow f(x_1) \leq f(x_2) \quad (\text{bzw. } f(x_1) \geq f(x_2))$$

$$\begin{aligned} f \text{ monoton steigend} &\Leftrightarrow f'(x) \geq 0 \quad \text{für alle } x \in \mathbb{R} \\ f \text{ monoton fallend} &\Leftrightarrow f'(x) \leq 0 \quad \text{für alle } x \in \mathbb{R} \end{aligned}$$



Krümmung (Konvexität und Konkavität)

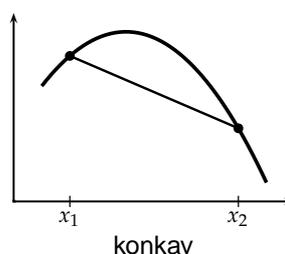
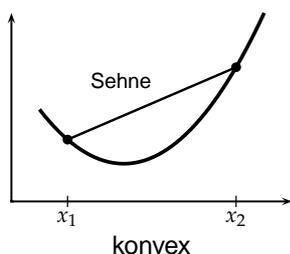
Eine Funktion f heißt **konvex**, wenn

$$f((1-h)x_1 + hx_2) \leq (1-h)f(x_1) + hf(x_2)$$

für alle x_1, x_2 und alle $h \in [0, 1]$.

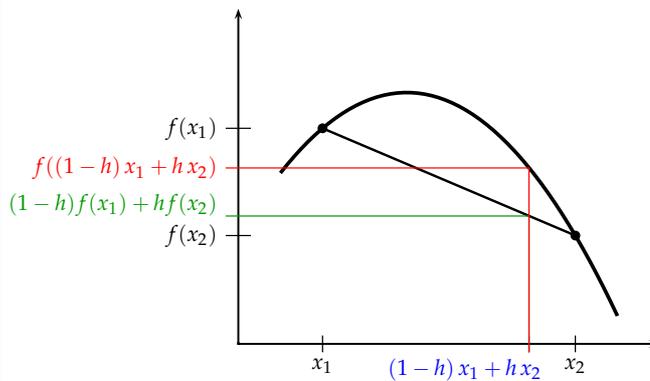
Sie heißt **konkav**, wenn

$$f((1-h)x_1 + hx_2) \geq (1-h)f(x_1) + hf(x_2)$$



Konkave Funktion

$$f((1-h)x_1 + hx_2) \geq (1-h)f(x_1) + hf(x_2)$$



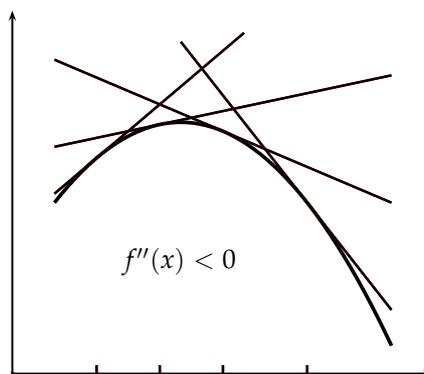
2. Ableitung einer Funktion

Die **zweite Ableitung** von $f, f''(x_0)$, ist die Ableitung der ersten Ableitungsfunktion, d.h. sie gibt die Steigung der ersten Ableitung an der Stelle x_0 an.

Ist die zweite Ableitung

- positiv, $f''(x) > 0$, so steigt dort die erste Ableitung:
eine stark negative Steigung wird weniger (wenn x größer wird) negativ oder positiv, eine positive Steigung wird stärker positiv.
- negativ, $f''(x) < 0$, so fällt dort die erste Ableitung:
eine stark positive Steigung wird weniger positiv oder negativ, eine positive Steigung wird schwächer positiv.

Ableitung einer konkaven Funktion



Konvexität und Konkavität

Einfacher für differenzierbare Funktionen:

$$\begin{aligned} f \text{ konvex} &\Leftrightarrow f''(x) \geq 0 \quad \text{für alle } x \in \mathbb{R} \\ f \text{ konkav} &\Leftrightarrow f''(x) \leq 0 \quad \text{für alle } x \in \mathbb{R} \end{aligned}$$

Bemerkung:

Eine Funktion kann auch auf einem bestimmten Intervall konkav, auf einem anderen konvex sein. Sie wird dann als **lokal** konkav, bzw. lokal konvex auf dem entsprechenden Abschnitt bezeichnet.

Streng konvex / konkav

Wir werden im Folgenden stets den Begriff der **strengen** Konvexität und Konkavität verwenden:

$$\begin{aligned} f \text{ streng konvex} &\Leftrightarrow f''(x) > 0 \quad \text{für alle } x \in \mathbb{R} \\ f \text{ streng konkav} &\Leftrightarrow f''(x) < 0 \quad \text{für alle } x \in \mathbb{R} \end{aligned}$$

Zur Erinnerung: Hinreichende Bedingung für lokale Extremwerte:

Sei x_0 ein stationärer Punkt von f (d.h. $f'(x_0) = 0$) dann gilt:

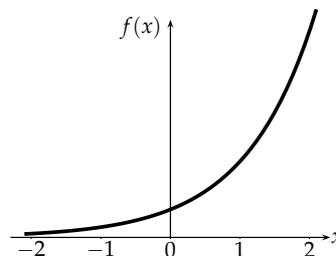
- x_0 ist ein lokales Minimum, wenn $f''(x_0) > 0$ (streng konvex).
- x_0 ist ein lokales Maximum, wenn $f''(x_0) < 0$ (streng konkav).

Konvex / Beispiel

Exponentialfunktion:

$$\begin{aligned} f(x) &= \exp(x) \\ f'(x) &= \exp(x) \\ f''(x) &= \exp(x) > 0 \end{aligned}$$

$\exp(x)$ ist (streng) konvex.



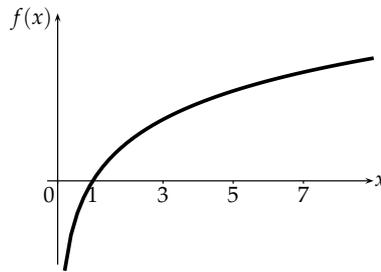
Logarithmus: $(x > 0)$

$$f(x) = \log(x)$$

$$f'(x) = \frac{1}{x}$$

$$f''(x) = -\frac{1}{x^2} < 0$$

$\log(x)$ ist (streng) konkav.



$\log(x)$ ist hier stets der *natürliche* Logarithmus.

Kapitel 6

Multivariate Analysis

Lernziele

- Funktionen in mehreren Variablen
- Graph und Niveaulinien einer Funktion in zwei Variablen
- Partielle Ableitung und Gradient
- Lokale und globale Extrema
- Lagrange-Ansatz

Funktionen in mehreren Variablen

Eine **reelle Funktion in mehreren Variablen** ist eine Abbildung, die jedem Vektor x eine reelle Zahl zuordnet.

$$f: \mathbb{R}^n \rightarrow \mathbb{R}, x \mapsto f(x) = f(x_1, x_2, \dots, x_n)$$

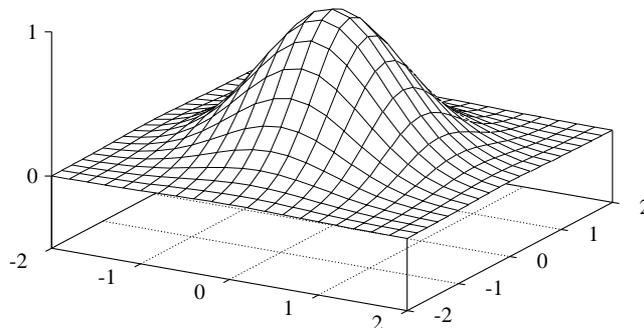
Die Komponenten x_i des Vektors x heißen die **Variablen** der Funktion f .

Funktionen in **zwei** Variablen lassen sich durch den **Graphen** (Funktionengebirge) veranschaulichen.

Dabei wird für jeden Punkt in der (x, y) -Ebene der Funktionswert $f(x, y)$ in die z -Richtung eingezeichnet.

Graph einer Funktion in zwei Variablen

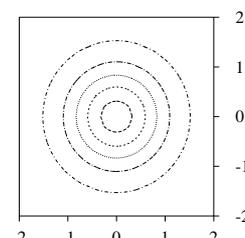
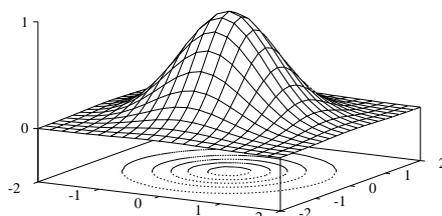
$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto f(x, y) = \exp(-x^2 - y^2)$$



Niveaulinien

Eine **Isoquante (Höhenlinie, Niveaulinie)** ist die Menge aller Punkte (x, y) mit $f(x, y) = c$ ($c \in \mathbb{R}$ konstant).

Die Funktion f hat daher auf einer Höhenlinie den gleichen Funktionswert.



Partielle Ableitung

Wir untersuchen die Änderung des Funktionswertes wenn wir eine Variable x_i variieren und alle anderen konstant lassen. Wir erhalten dadurch die (erste) **partielle Ableitung** von f nach x_i :

$$f_{x_i}(\mathbf{x}) = \frac{\partial f}{\partial x_i} = \lim_{\Delta x_i \rightarrow 0} \frac{F(\dots, x_i + \Delta x_i, \dots) - F(\dots, x_i, \dots)}{\Delta x_i}$$

Wir erhalten die partielle Ableitung nach x_i wenn wir alle anderen Variablen als Konstante auffassen und f nach den bekannten Regeln für Funktionen in einer Variable nach x_i ableiten.

Partielle Ableitung / Beispiel

Gesucht sind die ersten partiellen Ableitungen von

$$f(x_1, x_2) = \sin(2x_1) \cdot \cos(x_2)$$

$$f_{x_1} = 2 \cdot \cos(2x_1) \cdot \underbrace{\cos(x_2)}_{\text{als Konstante betrachtet}}$$

$$f_{x_2} = \underbrace{\sin(2x_1)}_{\text{als Konstante betrachtet}} \cdot (-\sin(x_2))$$

Partielle Ableitung / Beispiel

Gesucht sind die ersten partiellen Ableitungen von

$$f(x, y) = x^2 + 3xy$$

$$f_x = 2x + 3y$$

$$f_y = 0 + 3x$$

Der Gradient

Wir fassen die partiellen Ableitungen erster Ordnung zu einem Vektor, dem **Gradienten** an der Stelle x , zusammen.

$$\nabla f(x) = (f_{x_1}(x), \dots, f_{x_n}(x))$$

- ∇f heißt auch „nabla f “.
- Der Gradient wird oft als Zeilenvektor geschrieben.
- Der Gradient einer Funktion f zeigt in die Richtung des steilsten Anstieges von f . Seine Länge gibt diese Steigung an.
- Der Gradient steht immer normal auf die entsprechende Niveaulinie.
- Der Gradient „spielt“ die gleiche Rolle wie die erste Ableitung bei Funktionen in einer Variablen.

Der Gradient / Beispiel

Gesucht ist der Gradient von

$$f(x, y) = x^2 + 3xy$$

an der Stelle $x = (3, 2)$.

$$\nabla f(x) = (2x + 3y, 3x)$$

$$\nabla f(3, 2) = (12, 9)$$

Höhere partielle Ableitungen

Analog zu den Funktionen in einer Variablen können wir partielle Ableitungen nochmals ableiten und erhalten so **höhere partielle Ableitungen**.

$$f_{x_i x_k}(x) = \frac{\partial^2 f}{\partial x_k \partial x_i}(x) \quad \text{bzw.} \quad f_{x_i x_i}(x) = \frac{\partial^2 f}{\partial x_i^2}(x)$$

Es kommt auf die Reihenfolge beim Differenzieren nicht an (falls alle zweiten partiellen Ableitungen existieren und stetig sind):

$$\frac{\partial^2 f}{\partial x_k \partial x_i}(x) = \frac{\partial^2 f}{\partial x_i \partial x_k}(x)$$

Höhere partielle Ableitungen / Beispiel

Gesucht sind alle ersten und zweiten partiellen Ableitungen von

$$f(x, y) = x^2 + 3xy$$

Erste partielle Ableitungen:

$$f_x = 2x + 3y \quad f_y = 0 + 3x$$

Zweite partielle Ableitungen:

$$\begin{aligned} f_{xx} &= 2 & f_{xy} &= 3 \\ f_{yx} &= 3 & f_{yy} &= 0 \end{aligned}$$

Extrema (Optimierung)

x_0 heißt **globales Maximum** von f , falls für alle x gilt:

$$f(x_0) \geq f(x)$$

x_0 heißt **lokales Maximum** von f , falls für alle x in einer geeigneten Umgebung von x_0 (i.e. Menge aller Punkte in der Nähe von x_0) gilt:

$$f(x_0) \geq f(x)$$

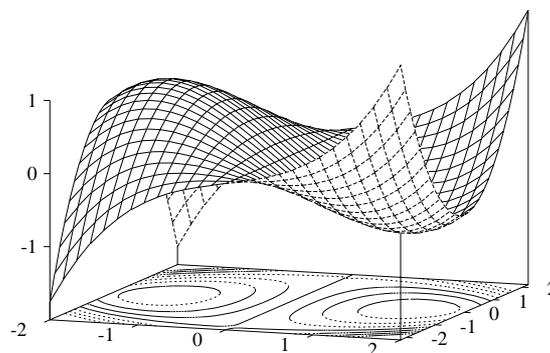
Analog werden **globales Minimum** und **lokales Minimum** definiert.

Minimum und Maximum werden als **Extremum** von f bezeichnet.

Aufgabe:

Suche Extrema einer gegebenen Funktionen.

Extrema / Beispiel



$$f(x, y) = \frac{1}{6}x^3 - x + \frac{1}{4}xy^2$$

Stationärer Punkt

Ein Punkt x_0 , in dem der Gradient verschwindet, i.e.,

$$\nabla f(x_0) = 0$$

heißt **stationärer Punkt** von f .

Es gilt (*notwendige Bedingung*):

Jedes Extremum von f ist ein stationärer Punkt von f .

Stationärer Punkt / Beispiel

Suche alle stationären Punkte von

$$f(x, y) = \frac{1}{6}x^3 - x + \frac{1}{4}xy^2$$

Bilde alle ersten partiellen Ableitungen und setze diese Null:

$$(I) \quad f_x = \frac{1}{2}x^2 - 1 + \frac{1}{4}y^2 = 0$$

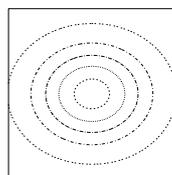
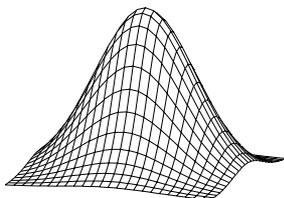
$$(II) \quad f_y = \frac{1}{2}xy = 0$$

Die stationären Punkte von f sind nun die Lösungen dieses Gleichungssystems:

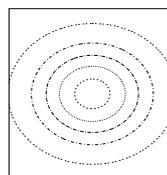
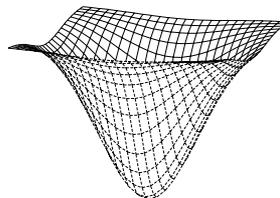
$$x_1 = (0, 2) \quad x_2 = (0, -2)$$

$$x_3 = (\sqrt{2}, 0) \quad x_4 = (-\sqrt{2}, 0)$$

Stationäre Punkte: Lokale Extrema

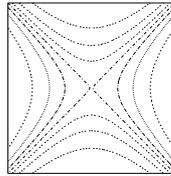
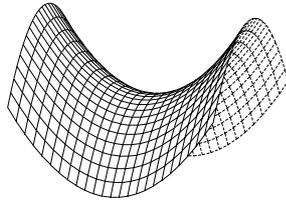


Lokales Maximum

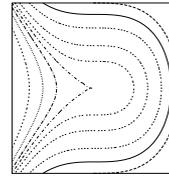
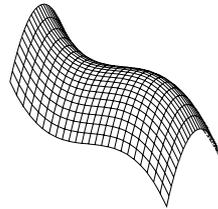


Lokales Minimum

Stationäre Punkte: Sattelpunkte und andere



Sattelpunkt



Beispiel für höhere Ordnung

Die Hesse-Matrix

Die Matrix aller zweiten partiellen Ableitungen einer Funktion f an der Stelle x wird als **Hesse-Matrix** $H_f(x)$ bezeichnet.

$$H_f(x) = \begin{pmatrix} f_{x_1x_1}(x) & f_{x_1x_2}(x) & \cdots & f_{x_1x_n}(x) \\ f_{x_2x_1}(x) & f_{x_2x_2}(x) & \cdots & f_{x_2x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1}(x) & f_{x_nx_2}(x) & \cdots & f_{x_nx_n}(x) \end{pmatrix}$$

Die Hesse-Matrix bestimmt ob die Funktion f in der Nähe von x konvex oder konkav ist (oder nicht).

Sie „spielt“ die gleiche Rolle, wie die zweite Ableitung von Funktionen in einer Variablen.

Die Hesse-Matrix / Beispiel

Berechne die Hesse-Matrix an den Stellen $x_1 = (0, 2)$ und $x_4 = (-\sqrt{2}, 0)$ von

$$f(x, y) = \frac{1}{6}x^3 - x + \frac{1}{4}xy^2$$

$$\begin{aligned} f_x &= \frac{1}{2}x^2 - 1 + \frac{1}{4}y^2 & f_{xx} &= x & f_{xy} &= \frac{1}{2}y \\ f_y &= \frac{1}{2}xy & f_{yx} &= \frac{1}{2}y & f_{yy} &= \frac{1}{2}x \end{aligned}$$

$$H_f(x) = \begin{pmatrix} f_{xx}(x) & f_{xy}(x) \\ f_{yx}(x) & f_{yy}(x) \end{pmatrix} = \begin{pmatrix} x & \frac{1}{2}y \\ \frac{1}{2}y & \frac{1}{2}x \end{pmatrix}$$

$$H_f(x_1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad H_f(x_4) = \begin{pmatrix} -\sqrt{2} & 0 \\ 0 & -\frac{\sqrt{2}}{2} \end{pmatrix}$$

Hauptminoren

Wir benötigen das „Vorzeichen“ (die **Definitheit**) der Hesse-Matrix.

Die Definitheit der Hesse-Matrix kann mit Hilfe der sogenannten **Hauptminoren** festgestellt werden.

Die Determinante der linken oberen $k \times k$ -Untermatrix heißt der **k -te (führende) Hauptminor**:

$$M_k = \begin{vmatrix} f_{x_1x_1}(\mathbf{x}) & f_{x_1x_2}(\mathbf{x}) & \cdots & f_{x_1x_k}(\mathbf{x}) \\ f_{x_2x_1}(\mathbf{x}) & f_{x_2x_2}(\mathbf{x}) & \cdots & f_{x_2x_k}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_kx_1}(\mathbf{x}) & f_{x_kx_2}(\mathbf{x}) & \cdots & f_{x_kx_k}(\mathbf{x}) \end{vmatrix}$$

Hauptminoren / Beispiel

Berechne alle Hauptminoren der Hesse-Matrix an den Stellen $\mathbf{x}_1 = (0, 2)$ und $\mathbf{x}_4 = (-\sqrt{2}, 0)$ von

$$f(x, y) = \frac{1}{6}x^3 - x + \frac{1}{4}xy^2$$

$$\mathbf{H}_f(\mathbf{x}_1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$M_1 = 0$$

$$M_2 = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1$$

$$\mathbf{H}_f(\mathbf{x}_4) = \begin{pmatrix} -\sqrt{2} & 0 \\ 0 & -\frac{\sqrt{2}}{2} \end{pmatrix}$$

$$M_1 = -\sqrt{2}$$

$$M_2 = \begin{vmatrix} -\sqrt{2} & 0 \\ 0 & -\frac{\sqrt{2}}{2} \end{vmatrix} = 1$$

Hinreichende Bedingung für Lokale Extrema

Für Funktionen in zwei Variablen erhalten wir folgende Bedingung:

Sei \mathbf{x}_0 ein stationärer Punkt von f , und M_1 und M_2 die Hauptminoren von $\mathbf{H}_f(\mathbf{x}_0)$.

- (a) $M_2 < 0$ \Rightarrow \mathbf{x}_0 ist Sattelpunkt
- (b) $M_2 > 0$ und $M_1 > 0$ \Rightarrow \mathbf{x}_0 ist lokales Minimum
- (c) $M_2 > 0$ und $M_1 < 0$ \Rightarrow \mathbf{x}_0 ist lokales Maximum
- (d) $M_2 = 0$ \Rightarrow Keine Aussage möglich

Lokale Extrema / Beispiel

Berechne alle lokalen Extrema der Funktion

$$f(x, y) = \frac{1}{6}x^3 - x + \frac{1}{4}xy^2$$

Stationäre Punkte:

$$x_1 = (0, 2): \quad M_2 = -1 < 0 \quad \Rightarrow \text{Sattelpunkt}$$

$$x_2 = (0, -2): \quad M_2 = -1 < 0 \quad \Rightarrow \text{Sattelpunkt}$$

$$x_3 = (\sqrt{2}, 0): \quad M_2 = 1 > 0, M_1 = \sqrt{2} > 0 \quad \Rightarrow \text{lokales Minimum}$$

$$x_4 = (-\sqrt{2}, 0): \quad M_2 = 1 > 0, M_1 = -\sqrt{2} < 0 \quad \Rightarrow \text{lokales Maximum}$$

Hinreichende Bedingung für Lokale Extrema / Allgemein

Sei x_0 ein stationärer Punkt von f , und M_k der k -te Hauptminoren von $H_f(x_0)$.

(a) Alle Hauptminoren $M_k > 0$

$\Rightarrow x_0$ ist ein **lokales Minimum** von f .

(b) Für alle Hauptminoren gilt $(-1)^k M_k > 0$

$\Rightarrow x_0$ ist ein **lokales Maximum** von f .

(c) $\det(H_f(x_0)) \neq 0$, aber weder (a) noch (b) sind erfüllt

$\Rightarrow x_0$ ist ein **Sattelpunkt** von f .

(d) Andernfalls ist keine Aussage möglich, d.h. x_0 kann ein lokales Extremum sein, muß aber nicht.

Lokale Extrema / Beispiel

Wir suchen die lokalen Extrema der Funktion

$$f(x_1, x_2, x_3) = (x_1 - 1)^2 + (x_2 + 2)^2 + (x_3 + 1)^2$$

$$\nabla f = (2(x_1 - 1), 2(x_2 + 2), 2(x_3 + 1)) = (0, 0, 0)$$

Der einzige stationäre Punkt ist $x_0 = (1, -2, -1)$.

$$H_f(x) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Alle Hauptminoren sind positiv: $M_1 = 2, M_2 = 4, M_3 = 8$

$\Rightarrow x_0$ ist ein lokales Minimum.

Globale Extrema

Aufgabe:

Suche den höchsten Punkt der Welt.

Die Berechnung globaler Extrema von Funktionen in mehr als einer Variablen ist im allgemeinen sehr schwierig.

Wir wollen hier nur folgenden Spezialfall behandeln:

Sei x_0 ein stationärer Punkt von f , und

M_k der k -te Hauptminoren von $H_f(x)$.

- Falls alle Hauptminoren $M_k > 0$ für **alle** x , dann ist x_0 ein **globales Minimum** von f .
- Falls für alle Hauptminor $(-1)^k M_k > 0$ für **alle** x , dann ist x_0 ist ein **globales Maximum** von f .

Globale Extrema / Beispiel

Wir suchen die globalen Extrema der Funktion

$$f(x_1, x_2, x_3) = (x_1 - 1)^2 + (x_2 + 2)^2 + (x_3 + 1)^2$$

Der einzige stationäre Punkt ist $x_0 = (1, -2, -1)$.

$$H_f(x) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Die Hesse-Matrix ist unabhängig von x und alle Hauptminoren sind (für alle x) positiv: $M_1 = 2$, $M_2 = 4$, $M_3 = 8$

$\Rightarrow x_0$ ist ein globales Minimum.

Optimierung unter Nebenbedingungen

Aufgabe:

Berechne

$$\max / \min f(x, y)$$

unter der Nebenbedingung (*constraint*)

$$g(x, y) = c$$

Beispiel:

Wir suchen die lokalen Extrema der Funktion

$$f(x, y) = x^2 + 2y^2$$

unter der Nebenbedingung

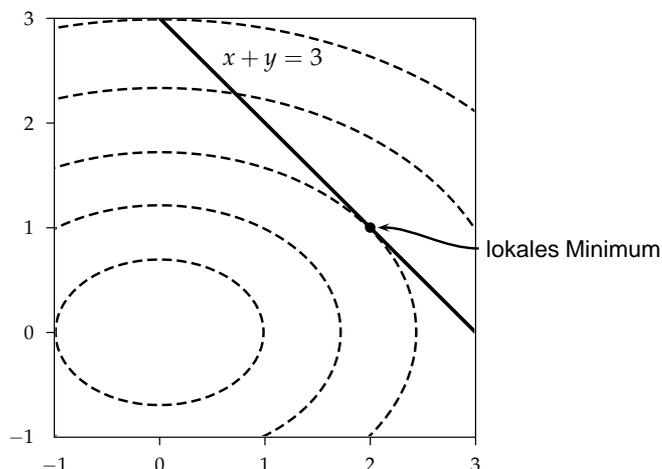
$$g(x, y) = x + y = 3$$

Graphische Lösung

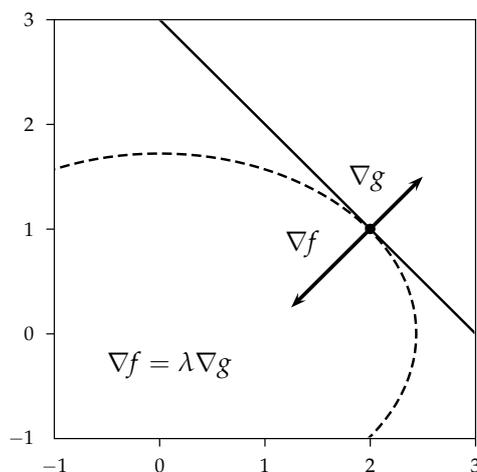
Im Falle von zwei Variablen können wir das Problem graphisch „lösen“.

- 1 Zeichne die Nebenbedingung $g(x, y) = c$ in die xy -Ebene ein. (Kurve in der Ebene)
- 2 Zeichne „geeignete“ Niveaulinien der zu optimierenden Funktion $f(x, y)$ ein.
- 3 Untersuche an hand der Zeichnung welche Niveaulinien den zulässigen Bereich schneiden und bestimme die ungefähre Lage der Extrema.

Graphische Lösung / Beispiel



Gradienten



Lagrangefunktion

Wir erzeugen uns aus f , g und einer Hilfsvariablen λ eine neue Funktion, die **Lagrange-Funktion**:

$$\mathcal{L}(x, y; \lambda) = f(x, y) - \lambda (g(x, y) - c)$$

Die Hilfsvariable λ heißt **Lagrange-Multiplikator**.

Wenn die Nebenbedingung erfüllt ist, so stimmen f und \mathcal{L} überein.

Lokale Extrema von f gegeben $g(x, y) = c$ sind stationäre Punkte der Lagrangefunktion \mathcal{L} :

$$\mathcal{L}_x = f_x - \lambda g_x = 0$$

$$\mathcal{L}_y = f_y - \lambda g_y = 0$$

$$\mathcal{L}_\lambda = -(g(x, y) - c) = 0$$

Lagrangefunktion / Beispiel

Wir suchen die lokalen Extrema der

$$f(x, y) = x^2 + 2y^2 \quad \text{gegeben} \quad g(x, y) = x + y = 3$$

Lagrangefunktion:

$$\mathcal{L}(x, y, \lambda) = (x^2 + 2y^2) - \lambda(x + y - 3)$$

Stationäre Punkte:

$$\mathcal{L}_x = 2x - \lambda = 0$$

$$\mathcal{L}_y = 4y - \lambda = 0$$

$$\mathcal{L}_\lambda = 3 - x - y = 0$$

⇒ einziger stationärer Punkt: $(x_0; \lambda_0) = (2, 1; 4)$

Geränderte Hesse-Matrix

Die Matrix

$$\bar{H}(x) = \begin{pmatrix} 0 & g_x & g_y \\ g_x & \mathcal{L}_{xx} & \mathcal{L}_{xy} \\ g_y & \mathcal{L}_{yx} & \mathcal{L}_{yy} \end{pmatrix}$$

heißt die **geränderte Hesse-Matrix**.

Hinreichende Bedingung für lokales Extremum:

Sei $(x_0; \lambda_0)$ ein stationärer Punkt von \mathcal{L} .

- $|\bar{H}(x_0)| > 0 \Rightarrow x_0$ ist lokales Maximum.
- $|\bar{H}(x_0)| < 0 \Rightarrow x_0$ ist lokales Minimum.

Geränderte Hesse-Matrix / Beispiel

Wir suchen die lokalen Extrema der

$$f(x, y) = x^2 + 2y^2 \quad \text{gegeben} \quad g(x, y) = x + y = 3$$

Lagrangefunktion: $(x_0; \lambda_0) = (2, 1; 4)$

$$\mathcal{L}(x, y, \lambda) = (x^2 + 2y^2) - \lambda(x + y - 3)$$

Stationärer Punkt: $(x_0; \lambda_0) = (2, 1; 4)$

Determinante der geränderte Hesse-Matrix:

$$|\tilde{H}(x_0)| = \begin{vmatrix} 0 & g_x & g_y \\ g_x & \mathcal{L}_{xx} & \mathcal{L}_{xy} \\ g_y & \mathcal{L}_{yx} & \mathcal{L}_{yy} \end{vmatrix} = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 4 \end{vmatrix} = -6 < 0$$

$\Rightarrow x_0 = (2, 1)$ ist ein lokales Minimum.

Lagrange-multiplikator / Eine Interpretation

Die Lage des Extremums x^* hängt von c der Nebenbedingung ab, $x^* = x^*(c)$, und somit auch der Extremalwert von f^* :

$$f^*(c) = f(x^*(c))$$

Wie ändert sich $f^*(c)$ mit c ?

Im Optimum stimmen \mathcal{L} und f überein. Daher ist

$$\frac{df^*(c)}{dc} = \frac{\partial \mathcal{L}}{\partial c} = \frac{\partial}{\partial c} (f(x) - \lambda g(x) + \lambda c) = \lambda$$

$$\frac{df^*(c)}{dc} = \lambda$$

Optimierung unter Nebenbedingungen / Allgemeiner Fall

Aufgabe:

Berechne

$$\max / \min f(x)$$

unter den Nebenbedingungen

$$g_1(x) = c_1$$

\vdots

$$(m < n)$$

$$g_m(x) = c_m$$

Lagrangefunktion

$$\mathcal{L}(x; \lambda_1, \dots, \lambda_m) = f(x) - \sum_{i=1}^m \lambda_i (g_i(x) - c_i)$$

Implizite Funktionen

Lernziele

- Explizit und implizit gegebene Funktion
- Existenz lokal expliziter Funktionen
- Differentiation impliziter Funktionen

Explizite und implizite Funktion

Der Zusammenhang zwischen zwei Variablen x und y kann gegeben werden durch eine

explizite Funktion:

$$y = f(x)$$

Beispiel:

$$y = x^2$$

existiert nicht

implizite Funktion:

$$F(x, y) = 0$$

Beispiel:

$$y - x^2 = 0$$

$$x^2 + y^2 - 1 = 0$$

Fragen:

- Wann kann man eine implizite Funktion (*lokal*) auch als explizite Funktion darstellen?
- Wie lautet die Ableitung von y nach der Variable x ?

Fall: Lineare Funktion

Im Falle einer linearen Funktion $F(x, y) = ax + by$ sind beide Fragen leicht zu beantworten:

$$ax + by = 0 \quad \Rightarrow \quad y = -\frac{a}{b}x \quad (\text{falls } b \neq 0)$$

$$\frac{dy}{dx} = -\frac{a}{b} = -\frac{F_x}{F_y}$$

Ableitung einer impliziten Funktion

Wir wollen $\frac{dy}{dx}$ ausrechnen.

Anstatt die (unbekannte) Funktion $y = y(x)$ abzuleiten, können wir $F(x, y)$ betrachten. Wir verändern x um Δx :

alte Stelle	neue Stelle	Veränderung
x	$x + \Delta x$	Δx
$F(x, y)$	$F(x + \Delta x, y)$	$F(x + \Delta x, y) - F(x, y)$

Da $F(x, y) = 0$ gilt, muss diese Änderung durch eine Veränderung von y wieder kompensiert werden:

$$F(x + \Delta x, y) - F(x, y) = -[F(x, y + \Delta y) - F(x, y)]$$

Ableitung einer impliziten Funktion I (2)

Wir erweitern auf beiden Seiten

$$\frac{F(x + \Delta x, y) - F(x, y)}{\Delta x} \Delta x = -\frac{F(x, y + \Delta y) - F(x, y)}{\Delta y} \Delta y$$

Grenzübergang ergibt partielle Ableitungen

$$\frac{\partial F(x, y)}{\partial x} dx = -\frac{\partial F(x, y)}{\partial y} dy \quad \text{bzw.} \quad F_x dx = -F_y dy$$

Umformung liefert

$$\frac{dy}{dx} = -\frac{F_x}{F_y}$$

Ableitung einer impliziten Funktion / Beispiel

Gesucht ist die implizite Ableitung $\frac{dy}{dx}$ von

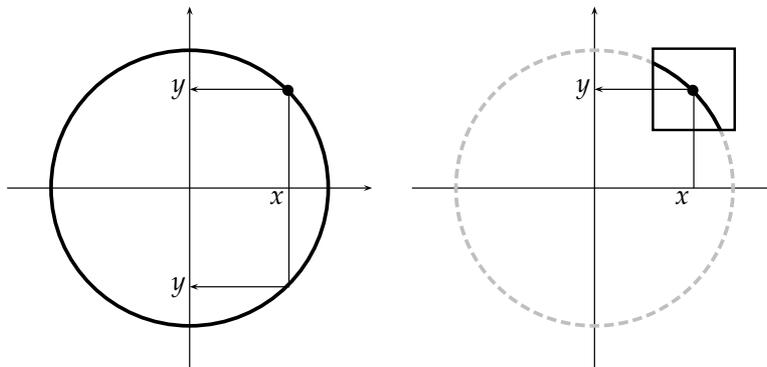
$$F(x, y) = x^2 + y^2 - 1 = 0$$

$$\frac{dy}{dx} = -\frac{F_x}{F_y} = -\frac{2x}{2y} = -\frac{x}{y}$$

Wir können auch die Ableitung von x nach der Variable y ausrechnen:

$$\frac{dx}{dy} = -\frac{F_y}{F_x} = -\frac{2y}{2x} = -\frac{y}{x}$$

Lokale Existenz einer expliziten Funktion



$y = y(x)$ existiert lokal, genau dann wenn $F_y \neq 0$.

Kapitel 8

Erwarteter Nutzen

Lernziele

- Nutzenfunktion zur Risikobewertung
- Erwarteter Nutzen
- Maße für Risikoaversion
- Indifferenzkurven in der $\sigma^2 \times \pi^e$ -Ebene
- Intertemporaler Nutzen für Mehrperioden-Entscheidungen

Überblick

Die Nutzentheorie wird hier auf Entscheidungen unter Unsicherheit angewandt. Sie dient auch zur Charakterisierung von Investoren.

- Der abnehmende Grenznutzen einer zusätzlichen Weinflasche (konkave Nutzenfunktion) wird hier einem **risikoaversen** Investor zugeschrieben.
- Ist der Beitrag der zweiten Weinflasche zum Nutzen größer als die erste (konvexe Nutzenfunktion), so werden wir sagen, der Investor ist **risikofreudig**.
- Dazwischen (die erste Flasche ist genauso gut wie die zweite) liegt **Risikoneutralität**.
- Konsumströme, die über mehrere Perioden verteilt sind, werden über einen diskontierten Nutzen bewertet.

Erwarteter Nutzen

Wir bezeichnen mit $U(W)$ den Nutzen des Vermögens W (*wealth*).

Im Gegensatz zum deterministischen Fall ist W hier stochastisch, eine Zufallsvariable. Das bedeutet, W ist keine feste Zahl, sondern kann verschiedene Werte annehmen, die wir im Voraus nicht kennen. In einem Zeitpunkt ist W klein (Pech), manchmal ist W groß (Glück).

Unsere Entscheidung wird nicht bezüglich des erwarteten (durchschnittlichen) Vermögens getroffen, sondern bezüglich des **erwarteten** (durchschnittlichen) **Nutzens** unseres Vermögens.

$$E[U(W)] = \sum_w U(w) P(w)$$

Steigende Nutzenfunktion

Wir wollen annehmen, dass eine Nutzenfunktion stets differenzierbar und **streng monoton steigend** ist.

$$U'(W) > 0$$

Risikoneutralität und Risikoaversion

Beispiel:

Wir werfen eine perfekte Münze. Die Auszahlungen sind 2 € bei Kopf bzw. 0 € bei Zahl bei einem Spieleinsatz von 1 €. Der erwartete monetäre Wert des (riskanten) Spiels ist 1 €, also der Einsatz:

$$2 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} = 1$$

Der Wert, dieses Spiel nicht zu spielen, ist ebenfalls 1 € (der Einsatz).

Der risikoaverse Investor wird an diesem Spiel nicht teilnehmen. Es stehen den sicheren 1 €, unsichere 1 € (2 € bzw. 0 € je mit Wahrscheinlichkeit $\frac{1}{2}$) gegenüber.

Risikoneutralität und Risikoaversion / (2)

Unter **Risikoaversion** gilt

$$U(1) > U(2) \cdot 1/2 + U(0) \cdot 1/2$$

d.h.

$$U(1) - U(0) > U(2) - U(1)$$

Die erste Ableitung nimmt ab. Risikoaversion impliziert daher

$$U''(W) < 0$$

Bei **Risikoneutralität** gilt das Gleichheitszeichen,

$$U''(W) = 0$$

Nutzenfunktion: Unsicherheit und Risiko

Übersicht:

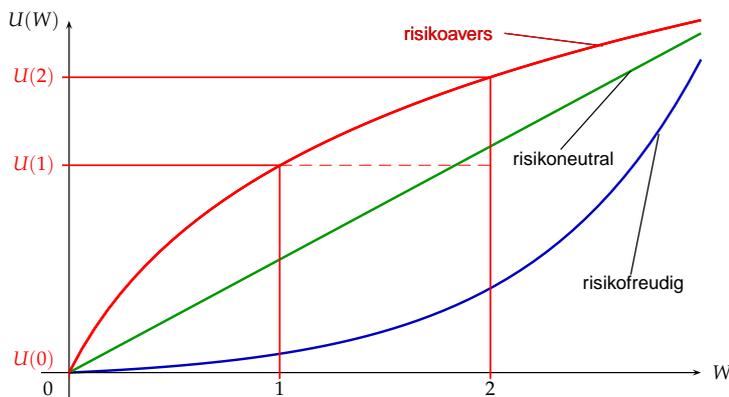
$U''(W) < 0$ **risikoavers**

$U''(W) = 0$ **risikoneutral**

Die Nutzenfunktion ist linear.

$U''(W) > 0$ **risikofreudig**

Risikoavers, Riskoneutral, Risikofreudig



Risikoavers: $U(1) - U(0) > U(2) - U(1)$

Maßzahlen für die Risikoaversion

Das **Arrow-Pratt Maß** für die **absolute Risikoaversion** ist

$$R_A(W) = -U''(W)/U'(W)$$

Das Maß für die **relative Risikoaversion** ist

$$R_R(W) = R_A(W) W$$

- Bei Risikoaversion sind $R_A(W)$ und $R_R(W) > 0$.
- Je größer R_A bzw. R_R desto größer ist die Risikoaversion.
- Sie sind ein Maß dafür, wie sich die Risikopräferenzen mit steigendem W verändern.

Abnehmende Risikoaversion

Abnehmende absolute Risikoaversion:

Angenommen ein (risikoaverser) Investor mit 10 000 € Vermögen hält 5 000 € in riskanten Anlagen. Falls sein Vermögen um weitere 10 000 € ansteigt und er dann davon mehr als 5 000 € riskant zu veranlagen bereit ist, spricht man von abnehmender Risikoaversion.

Abnehmende relative Risikoaversion:

Angenommen ein (risikoaverser) Investor hält 50% von 10 000 € in riskanten Anlagen. Falls sein Vermögen sich verdoppelt und er bereit ist den Anteil der riskanten Anlagen auf mehr als 50% zu erhöhen, spricht man von abnehmender relativer Risikoaversion.

$$U(W) = \ln(W)$$

Die Nutzenfunktion

$$U(W) = \ln(W)$$

- beschreibt Risikoaversion, $U''(W) < 0$.

$$U''(W) = [\ln(W)]'' = [1/W]' = -1/W^2 < 0$$

- hat eine abnehmende absolute Risikoaversion.

$$R_A(W) = -U''(W)/U'(W) = -(-1/W^2)/(1/W) = 1/W > 0$$

$R_A(W) = 1/W$ fällt mit W .

- hat eine konstante relative Risikoaversion.

$$R_W(W) = R_A(W) W = (1/W) W = 1.$$

$$U(W) = a - b \exp(-cW)$$

Die Nutzenfunktion

$$U(W) = a - b \exp(-cW), \quad a, b, c > 0$$

- beschreibt Risikoaversion.

$$U''(W) = [a - b \exp(-cW)]'' = [-b \exp(-cW) (-c)]' = -b c^2 \exp(-cW)$$

Da b positiv, $b > 0$, ist $U''(W) < 0$: Risikoaversion.

- hat eine konstante absolute Risikoaversion.

$$R_A(W) = -U''(W)/U'(W) = -\frac{(bc^2 \exp(-cW) (-c))}{(-bc \exp(-cW) (-c))} = c$$

Stochastisches Modell

Das Vermögen W morgen (Zeitpunkt 1) sei

$$W = (1 + \pi) W_0$$

Wobei

- W_0 das Vermögen im Zeitpunkt 0 bezeichnet, und
- π die ungewisse Rendite (stochastisch) von Zeitpunkt 0 bis 1:
$$\pi = W/W_0 - 1.$$

Die Rendite

- sei normalverteilt
- mit der Erwartung $E(\pi) = \pi^e$ und
- der Varianz $V(\pi) = \sigma_\pi^2$.

Schreibweise $\pi \sim \mathcal{N}(\pi^e, \sigma_\pi^2)$.

Stochastisches Modell / Erwarteter Nutzen

Die Maximierung des erwarteten Nutzens

$$\max_W E[U(W)] = \max_W E[a - b \exp(-cW)]$$

ist in diesem Modell äquivalent zur Maximierung von

$$U(\pi^e, \sigma_\pi^2) = \pi^e - \frac{c}{2} \sigma_\pi^2$$

Bemerkung:

- Die Herleitung verwendet, dass $\exp(-cW)$ (bedingt) lognormal-verteilt ist.
- Für π^e wird später die Bezeichnung μ verwendet werden, für σ_π^2 nur σ^2 .

Nutzenfunktion in π^e und σ_π^2

Wir beschränken uns auf Nutzenfunktionen, die sich als Funktion der erwarteten Rendite, π^e , und der Varianz der Rendite, σ_π^2 , darstellen lassen:

$$U = U(\pi^e, \sigma_\pi^2)$$

mit

$$U_1 > 0, \quad U_2 < 0, \quad U_{11}, U_{22} < 0$$

U_1 ist die partielle Ableitung nach π^e ,

U_2 die partielle Ableitung nach σ_π^2 .

U_{11} steht für $[\partial^2 U / (\partial \pi^e)^2]$,

U_{22} steht für $[\partial^2 U / (\partial \sigma_\pi^2)^2]$.

Bemerkung:

Die Funktion $U(\pi^e, \sigma_\pi^2) = \pi^e - \frac{c}{2} \sigma_\pi^2$ erfüllt für $c > 0$ nur die Bedingungen $U_1 > 0, U_2 < 0$. ($U_{11}, U_{22} = 0$)

Nutzenfunktion in π^e und σ_π^2 / (2)

Der Nutzen

- steigt mit der erwarteten Rendite,
- fällt aber mit der Varianz (Risiko) in der Rendite.

Für beide Argumente, π^e und σ_π^2 , gelten abnehmende Grenznutzen.

Indifferenzkurven für $U = U(\pi^e, \sigma_\pi^2)$

Eine **Indifferenzkurve** ist definiert als die Menge aller Punkte für die die Nutzenfunktion den gleichen Wert c aufweist.

$$U(\pi^e, \sigma_\pi^2) = c$$

- Alle Kombinationen (π^e, σ_π^2) auf einer Indifferenzkurve mit Niveau c , liefern denselben Nutzen mit dem Wert c .
- Der Investor ist indifferent zwischen diesen Kombinationen. Jede dieser Kombinationen ist für ihn gleich gut/schlecht.

Die Nutzenfunktion wird in der Portfoliooptimierung verwendet werden.

Darstellung der Indifferenzkurve

Wir stellen die Indifferenzkurven in der $\sigma_\pi^2 \times \pi^e$ -Ebene dar.
 σ_π^2 wird hier auf der x -Achse aufgetragen, π^e auf der y -Achse!

Da der Wert der (Nutzen-)Funktion konstant ist, beschreibt $U(\pi^e, \sigma_\pi^2) = c$ eine (implizite) Funktion in 2 Variablen, π^e und σ_π^2 .

Da wir die $\sigma_\pi^2 \times \pi^e$ -Ebene zugrunde legen, sehen wir π^e als Funktion von σ_π^2 an: $\pi^e = \pi^e(\sigma_\pi^2)$.

Lage:

Ein höheres π^e bei konstantem σ_π^2 erzeugt einen höheren Nutzen, da $U_1 > 0$ ist.

Daher verschieben sich die Indifferenzkurven bei steigendem Nutzen nach links oben.

Steigung und Krümmung der Indifferenzkurve

Steigung:

Für die Steigung dieser Kurve erhalten wir:

$$\frac{d\pi^e}{d\sigma_{\pi}^2} = -\frac{U_{\sigma_{\pi}^2}}{U_{\pi^e}} = -\frac{U_2}{U_1} = -\frac{(-)}{(+)} = (+)$$

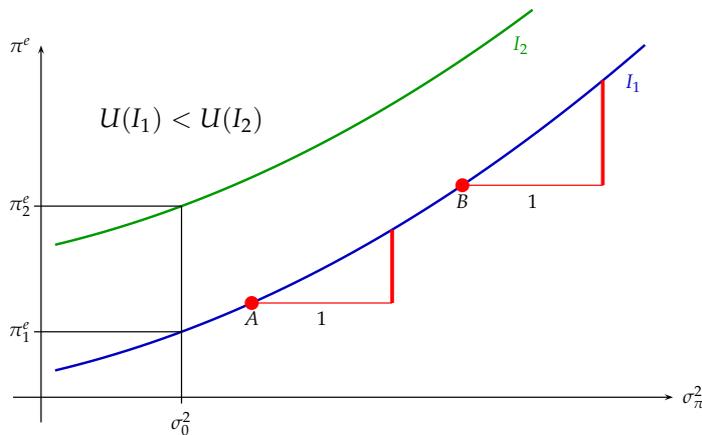
- Unsere Indifferenzkurven steigen für jede Wahl von c .
- Steigt σ_{π}^2 , das Risiko, so muss das ein höheres π^e ausgleichen.

2. Ableitung:

Die zweite Ableitung der Indifferenzkurven ist durch die Angaben der Vorzeichen von U_1, U_2, U_{11}, U_{22} noch nicht bestimmt.

Unter Risikoaversion folgt, dass sie konvex bez. der „Risiko-“ σ^2 -Achse sein müssen.

Indifferenzkurve bei Risikoaversion



Investor verlangt in B eine größere Rendite für den gleichen Zuwachs an Risiko als in A.

Intertemporaler Nutzen

Einige ökonomische Modelle gehen davon aus, dass ein Investor seinen Nutzen nur aus seinem Konsum, C , gewinnt.

$$U = U(C_t), \quad U'(C_t) > 0, \quad U''(C_t) < 0$$

t ist ein Zeitindex: $t = 0, 1, 2, 3, \dots$

Die Nutzenfunktion ist ähnlich der risikoaversen.

Die allgemeine **intertemporale** Lebensnutzenfunktion ist

$$U_N = U(C_t, C_{t+1}, \dots, C_{t+N})$$

N ist das (fixe) Lebensalter.

Separable Nutzenfunktion

Mathematisch einfacher ist die additiv separable Funktion mit einer konstanten, subjektiven Diskontrate $0 < \delta < 1$:

$$U_N = U(C_t) + \delta U(C_{t+1}) + \dots + \delta^n U(C_{t+N})$$

$$\delta = \frac{1}{1+r}$$

- r ist die subjektive **Zeitpräferenzrate**.
(Meist ähnlich dem Marktzinssatz.)
- Der Nutzen in verschiedenen Perioden ist geeignet diskontiert direkt vergleichbar.

$$U(C) = a C^{1-d}$$

Eine einfache Form (vom Cobb-Douglas-Typ) mit $a > 0$ und $0 < d < 1$ ist:

$$U(C_t) = a C_t^{1-d}$$

$$U'(C_t) = a(1-d) C_t^{-d}$$

$$U''(C_t) = -a(1-d)d C_t^{-d-1}$$

$$U(C_t) > 0, \quad U'(C_t) > 0, \quad U''(C_t) < 0$$

$$U = C_0^{\alpha_1} C_1^{\alpha_2}$$

Eine andere einfache Form (vom Cobb-Douglas-Typ) mit $0 < \alpha_1, \alpha_2 < 1$ (meist $\alpha_2 < \alpha_1$) ist:

$$U = C_0^{\alpha_1} C_1^{\alpha_2}$$

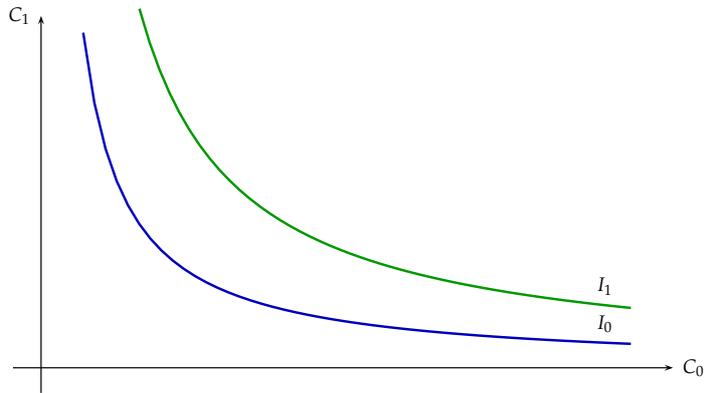
Sie wird durch Logarithmierung separabel:

$$\ln(U) = \alpha_1 \ln(C_0) + \alpha_2 \ln(C_1)$$

Die Indifferenzkurven können hier explizit angegeben werden;
z.B. in der $C_0 \times C_1$ -Ebene.

Sei \bar{U} das fixierte Nutzenniveau, $\bar{U} = C_0^{\alpha_1} C_1^{\alpha_2}$, dann kann nach C_1

aufgelöst werden: $C_1 = \bar{U}^{\frac{1}{\alpha_2}} C_0^{-\frac{\alpha_1}{\alpha_2}} = (+) \frac{1}{C_0^{\alpha_1/\alpha_2}}$



Anwendungen: Mehrperioden-Investitionsentscheidungen, Mehrperioden-Konsumentscheidungen, Portfoliowahl.

Grenzrate der Substitution

Die Grenzrate der Substitution des zukünftigen für den laufenden Konsum ist

$$\left. \frac{dC_1}{dC_0} \right|_{U=\bar{U}}$$

Der aktuelle Konsum ist C_0 , der zukünftige C_1 .

Die Grenzrate gibt den Wert einer zusätzlichen Einheit vom heutigen Konsum in Einheiten des zukünftigen an.

Sie ist die **Steigung der Indifferenzkurve** mit dem Niveau \bar{U} .

Beispiel: $U = C_0^{\alpha_1} C_1^{\alpha_2}$

$$\left. \frac{dC_1}{dC_0} \right|_{U=\bar{U}} = \left(-\frac{\alpha_1}{\alpha_2}\right) \bar{U}^{1/\alpha_2} \frac{1}{C_0^{\alpha_1/\alpha_2+1}}$$

Kapitel 9

Taylorreihen

Lernziele

- Taylorreihe und MacLaurinreihe
- Beispiele
- Alternative Schreibweisen
- Approximation der Veränderung einer Funktion
- Delta-Methode: Approximation von $E[f(X)]$

Taylorreihe

Wenn eine Funktion $f(x)$ genügend oft differenzierbar ist, kann sie durch ein Polynom n -ter Ordnung approximiert werden.

$$f(x) \approx f(x_0) + \frac{f'(x_0)}{1!} (x - x_0) + \frac{f''(x_0)}{2!} (x - x_0)^2 + \dots \\ \dots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n$$

Man sagt, die Funktion $f(x)$ wird

- an der Stelle x_0 in eine **Taylorreihe** bis zur Ordnung n entwickelt.
- durch ein Polynom n -ter Ordnung approximiert (**Taylorpolynom**).

Taylorreihe / (2)

Bezeichnung:

- $f^{(k)}(x)$ bezeichnet die k -te Ableitung von f nach x .
- $f^{(k)}(x_0)$ ist der Wert der k -ten Ableitung von f an der Stelle x_0 .
- $k!$ heißt k -Faktorielle: $k! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot k$.
 $0! = 1, \quad 1! = 1, \quad 2! = 2, \quad 3! = 1 \cdot 2 \cdot 3 = 6$

Die Approximation ist umso besser (der Fehler ist umso kleiner)

- je näher x an der Entwicklungstelle x_0 ist;
- je größer die Ordnung n ist.

Die Koeffizienten des Polynoms sind so gewählt, dass die ersten n Ableitungen mit jener von f übereinstimmen.

MacLaurinreihe

Der Spezialfall mit Entwicklungspunkt $x_0 = 0$ wird auch als **MacLaurin-Polynom** bezeichnet.

$$f(x) \approx f(0) + \frac{f'(0)}{1!} x + \frac{f''(0)}{2!} x^2 + \dots + \frac{f^{(n)}(0)}{n!} x^n$$

Taylorreihe: $n = 1$

Für $n = 1$ ergibt sich ein Polynom erster Ordnung (lineare Funktion).

$$f(x) \approx f(x_0) + f'(x_0) (x - x_0)$$

Man schreibt auch

$$f(x) \doteq f(x_0) + f'(x_0) (x - x_0)$$

\doteq bedeutet „in erster Näherung“.

Wenn wir die Näherung verwenden, rechnen wir mit der Tangente der Funktion an der Stelle x_0 anstatt mit f .

Beispiel: $\exp(x)$

Taylorentwicklung erster Ordnung von $f(x) = \exp(x)$ an der Entwicklungsstelle $x_0 = 0$.

$$\begin{aligned} f(x) &\approx f(0) + f'(0) x \\ \exp(x) &\approx 1 + 1 \cdot x = 1 + x \end{aligned}$$

$$\begin{aligned} f(x) &= \exp(x) & f(0) &= \exp(0) = 1 \\ f'(x) &= \exp(x) & f'(0) &= \exp(0) = 1 \end{aligned}$$

Das Taylorpolynom erster Ordnung von $\exp(x)$ in $x_0 = 0$ ist

$$t(x) = 1 + x$$

Beispiel: $\log(1+x)$

Taylorentwicklung erster Ordnung von $f(x) = \log(1+x)$ an der Entwicklungsstelle $x_0 = 0$.

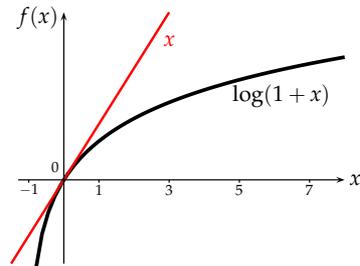
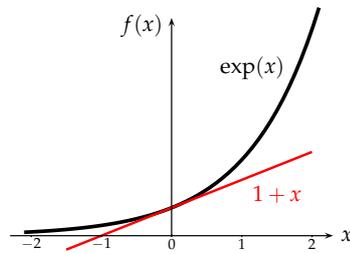
$$f(x) \approx f(0) + f'(0) x$$
$$\log(1+x) \approx 0 + 1 \cdot x = x$$

$$f(x) = \log(1+x) \quad f(0) = \log(1) = 0$$
$$f'(x) = 1/(1+x) \quad f'(0) = 1/(1+0) = 1$$

Das Taylorpolynom erster Ordnung von $\log(1+x)$ in $x_0 = 0$ ist

$$t(x) = x$$

Graph: $\exp(x)$ und $\log(1+x)$



Wir approximieren $\exp(x)$ an der Stelle 0 durch eine Gerade:

$$\exp(x) \doteq 1+x$$

Wir approximieren $\log(1+x)$ an der Stelle 0 durch eine Gerade:

$$\log(1+x) \doteq x$$

Anwendung

Für alle $x \in \mathbb{R}$ gilt:

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x = \exp(x)$$

Eine Motivation für diese Beziehung erhalten wir über

$$\left(1 + \frac{x}{n}\right)^n = e^{n \cdot \log(1+x/n)} \doteq e^{n \cdot (x/n)} = e^x$$

Die Näherung $\log(1+x) \doteq x$ ist gut, wenn x betragsmäßig klein ist.

Taylorreihe: $n = 2$

Für $n = 2$ ergibt sich ein Polynom zweiter Ordnung (quadratische Funktion).

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2$$

Die Approximation zweiter Ordnung ist in der Regel eine bessere lokale Approximation als die Erste.

Beispiel: $\exp(x)$ ($n = 2$)

Taylorentwicklung zweiter Ordnung von $f(x) = \exp(x)$ an der Entwicklungsstelle $x_0 = 0$.

$$f(x) \approx f(0) + f'(0)x + \frac{f''(x_0)}{2!}x^2$$

$$\exp(x) \approx 1 + 1x + \frac{1}{2}x^2 = 1 + x + \frac{1}{2}x^2$$

$$f(x) = \exp(x) \quad f(0) = \exp(0) = 1$$

$$f'(x) = \exp(x) \quad f'(0) = \exp(0) = 1$$

$$f''(x) = \exp(x) \quad f''(0) = \exp(0) = 1$$

Beispiel: $\log(1 + x)$ ($n = 2$)

Taylorentwicklung zweiter Ordnung von $f(x) = \log(1 + x)$ an der Entwicklungsstelle $x_0 = 0$.

$$f(x) \approx f(0) + f'(0)x + \frac{f''(x_0)}{2!}x^2$$

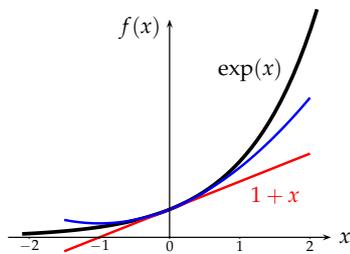
$$\log(1 + x) \approx 0 + 1x + \frac{-1}{2}x^2 = x - \frac{1}{2}x^2$$

$$f(x) = \log(1 + x) \quad f(0) = \log(1) = 0$$

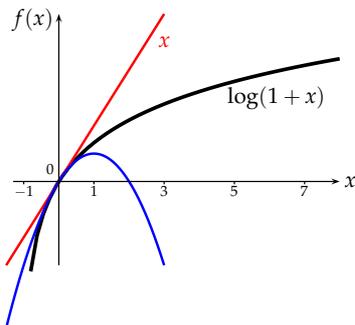
$$f'(x) = 1/(1 + x) = (1 + x)^{-1} \quad f'(0) = 1/(1 + 0) = 1$$

$$f''(x) = -(1 + x)^{-2} \quad f''(0) = -(1 + 0)^{-2} = -1$$

Graph: $\exp(x)$ und $\log(1+x)$



$$\exp(x) \approx 1 + \frac{x}{1} + \frac{x^2}{2}$$



$$\log(1+x) \approx x - \frac{x^2}{2}$$

Wichtige Taylorreihen $(x_0 = 0)$

- Exponentialfunktion

$$\exp(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

- Logarithmus

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots$$

- Sinus

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$$

- Kehrwert

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 + \dots$$

Alternative Schreibweise: $h = (x - x_0)$

Wir heben die Veränderung $(x - x_0)$ hervor, indem wir sie mit h ($= (x - x_0)$) bezeichnen.

Die Entwicklungsstelle bezeichnen wir als x .

$$f(x+h) \approx f(x) + \frac{f'(x)}{1!} h + \frac{f''(x)}{2!} h^2 + \dots + \frac{f^{(n)}(x)}{n!} h^n$$

- x_0 wird durch x ersetzt,
- $(x - x_0)$ durch h , und
- x durch $(x + h)$.

Alternative Schreibweise: $\Delta x = (x - x_0)$

Schreiben Taylorreihe mit Δx anstatt h .

$$f(x + \Delta x) \approx f(x) + \frac{f'(x)}{1!} \Delta x + \frac{f''(x)}{2!} (\Delta x)^2 + \dots + \frac{f^{(n)}(x)}{n!} (\Delta x)^n$$

- Δx bezeichnet die **Änderung** von x .

Approximation für $\Delta f = f(x + \Delta x) - f(x)$

Wir $f(x)$ auf die linke Seite. Erhalten damit eine Entwicklung für die Veränderung der Funktion.

- $\Delta f(x) = f(x + \Delta x) - f(x)$ bezeichnet die **Änderung** von $f(x)$ (wenn sich x um Δx ändert)

Wir vergleichen den Funktionswert an der Stelle x mit dem bei $(x + \Delta x)$.

$$\Delta f(x) \approx \frac{f'(x)}{1!} \Delta x + \frac{f''(x)}{2!} (\Delta x)^2 + \dots + \frac{f^{(n)}(x)}{n!} (\Delta x)^n$$

Approximation für Δf in 1. Näherung

In erster Näherung gibt sich mit $\Delta f(x) = f(x + \Delta x) - f(x)$

$$\Delta f(x) \doteq f'(x) \Delta x$$

Die Veränderung des Funktionswertes wird gut durch die Ableitung an der Stelle x mal der Schrittweite, Δx , beschrieben.
(Steigung der Tangente mal Schrittweite)

Diese „diskrete“ Approximation einer stetigen Funktion werden wir im Folgenden extensiv verwenden.

Beispiel: $\log(1 + x)$

In erster Näherung gibt sich für $\Delta f(x)$

$$\Delta f(x) \doteq \frac{1}{1+x} \Delta x$$

Änderung von $f(x) = \log(1 + x)$ an der Stelle $x = 2$:

$$\Delta f(2) = f(2 + \Delta x) - f(2) \doteq \frac{1}{1+2} \Delta x = \frac{1}{3} \Delta x$$

Wenn wir von 2 nach 2.5 gehen erhalten wir $\Delta = 1/2$ und

$$\Delta f(2) = [f(2 + 0.5) - f(2)] \doteq \frac{1}{1+2} 0.5 = \frac{1}{6}$$

Die Änderung der Funktion beträgt zwischen 2 und 2.5 ungefähr 1/6.

Delta-Methode: Approximation von $E[f(X)]$

Sei X eine ZV mit $E(X) = \mu$ und $V(X) = \sigma^2$.

$f(X)$ sei eine (nicht-lineare) Funktion in X .

Die **Delta-Methode** approximiert den Erwartungswert der nicht-linearen Funktion, $E[f(X)]$, in dem die Funktion in eine Taylorreihe bis zu einer (niedrigen) Potenz entwickelt und anschließend von dieser der Erwartungswert berechnet wird.

Approximation 1. Ordnung

Wir entwickeln f an der Stelle μ . Also $x_0 = \mu$:

$$f(X) \doteq f(\mu) + f'(\mu) (X - \mu)$$

$$E[f(X)] \approx f(\mu) \quad \text{und} \quad V[f(X)] \approx f'(\mu)^2 \sigma^2$$

Aus den Rechenregeln für Linearkombinationen von Zufallsvariablen erhalten wir:

$$\begin{aligned} E[f(X)] &\doteq E[f(\mu) + f'(\mu) (X - \mu)] \\ &= E[f(\mu)] + f'(\mu) E[(X - \mu)] = f(\mu) \end{aligned}$$

$$V[f(X)] \doteq V[f(\mu) + f'(\mu) (X - \mu)] = f'(\mu)^2 \sigma^2$$

Approximation 2. Ordnung

Wir entwickeln f an der Stelle μ . Also $x_0 = \mu$:

$$f(X) \doteq f(\mu) + f'(\mu)(X - \mu) + \frac{f''(\mu)}{2!}(X - \mu)^2$$

Dann erhalten wir für den Erwartungswert

$$E[f(X)] \approx f(\mu) + \frac{1}{2} f''(\mu) \sigma^2$$

und für die Varianz **unter der Annahme**, dass X normal verteilt ist,
 $X \sim N(\mu, \sigma^2)$,

$$V[f(X)] \approx f'(\mu)^2 \sigma^2 + \frac{1}{2} f''(\mu)^2 \sigma^4$$

Approximation 2. Ordnung / Herleitung

Erwartungswert:

$$\begin{aligned} E[f(X)] &\approx E\left[f(\mu) + f'(\mu)(X - \mu) + \frac{f''(\mu)}{2!}(X - \mu)^2\right] \\ &= E[f(\mu)] + f'(\mu) E[(X - \mu)] + \frac{1}{2} f''(\mu) E[(X - \mu)^2] \\ &= f(\mu) + f'(\mu) (E[X] - \mu) + \frac{1}{2} f''(\mu) V[X]^2 \\ &= f(\mu) + \frac{1}{2} f''(\mu) \sigma^2 \end{aligned}$$

Für die Varianz benötigen wir das 3. und 4. Moment der Normalverteilung, $E[(X - \mu)^3]$ und $E[(X - \mu)^4]$. Diese sind

$$E[(X - \mu)^3] = 0 \quad \text{und} \quad E[(X - \mu)^4] = 3 \sigma^4$$

Das 3. Moment („Schiefe“) ist Null, da die Normalverteilung symmetrisch ist.

(Wir leiten aber die Beziehung für die Varianz nicht her.)

Beispiel: $\exp(X)$

Sei $f(X) = \exp(X)$ und $X \sim N(\mu, \sigma^2)$:

Approximation 1. Ordnung:

$$E[f(X)] \approx \exp(\mu) \quad \text{und} \quad V[f(X)] \approx \exp(\mu)^2 \sigma^2$$

Approximation 2. Ordnung:

$$E[f(X)] \approx \exp(\mu) + \frac{1}{2} \exp(\mu) \sigma^2 = \exp(\mu) (1 + \sigma^2/2)$$

$$V[f(X)] \approx \exp(\mu)^2 \sigma^2 + \frac{1}{2} \exp(\mu)^2 \sigma^4 = \exp(2\mu) \sigma^2 (1 + \sigma^2/2)$$

Bemerkung: Die ZV $Y = \exp X$ ist lognormal verteilt.

Kovarianz und Korrelation

Lernziele

- Mathematische und statistische Grundlagen der Portfoliotheorie
- Kovarianz und Korrelation
- Kovarianz- und Korrelationsmatrix
- Multivariate Normalverteilung
- Erzeugen von multinormalverteilten Zufallsvektoren

Erwartungswert einer Linearkombination

Der Erwartungswert einer **Linearkombination** von ZVen ist die Linearkombination der einzelnen Erwartungswerte:

$$E(aX + bY + c) = aE(X) + bE(Y) + c$$

Allgemein

$$E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i E(X_i)$$

Voraussetzung: alle Erwartungswerte existieren.

Varianz einer Linearkombination

Die Varianz einer Linearkombination von ZVen ist **nicht** die Linearkombination der einzelnen Varianzen. Es gilt:

$$V(aX + bY + c) = a^2 V(X) + 2ab \operatorname{Cov}(X, Y) + b^2 V(Y)$$

Die Konstante c beeinflusst die Varianz nicht.

Bei der Varianz einer Summe tritt ein gemischter Term auf: die **Kovarianz** der beiden ZVen.

Nur wenn die Kovarianz der beiden ZVen Null ist, also beide unkorreliert sind, gilt:

„Die Varianz der Summe ist gleich die Summe der Varianzen“.

Voraussetzung: alle Varianzen existieren.

Kovarianz

Die Kovarianz zwischen zwei ZV X und Y ist definiert als

$$\begin{aligned} \operatorname{Cov}(X, Y) &= E[(X - E(X))(Y - E(Y))] \\ &= \sum_{x,y} (x - \mu_x)(y - \mu_y) P(X = x, Y = y) \end{aligned}$$

Hier gilt auch ein Verschiebungssatz:

$$E[(X - E(X))(Y - E(Y))] = E[XY] - E(X)E(Y)$$

Bemerkung: Ist $\operatorname{Cov}(X, Y) = 0$, folgt

$$E[XY] = E(X)E(Y)$$

Normalverteilte Zufallsvariable

Reproduktionseigenschaft der Normalverteilung:

Seien die ZVen X und Y normalverteilt mit

$$X \sim N(\mu_x, \sigma_x^2), \quad Y \sim N(\mu_y, \sigma_y^2) \quad \text{und} \quad \operatorname{Cov}(X, Y) = \sigma_{xy}$$

Dann ist

$$Z = aX + bY + c \sim N(\mu_z, \sigma_z^2)$$

$$\mu_z = a\mu_x + b\mu_y + c$$

$$\sigma_z^2 = a^2\sigma_x^2 + 2ab\sigma_{xy} + b^2\sigma_y^2$$

Korrelation

Die **Korrelation** zwischen zwei ZV X und Y ist definiert als

$$\text{Corr}(X, Y) = \rho(X, Y) = \rho = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)V(Y)}}$$

Es gilt immer

$$-1 \leq \text{Corr}(X, Y) \leq 1$$

$$\text{Cov}(X, Y) = \text{Corr}(X, Y) \sqrt{V(X)V(Y)}$$

$$\text{Corr}(X, Y) = 0 \Leftrightarrow \text{Cov}(X, Y) = 0$$

Unkorrelierte Zufallsvariable

Sind X und Y unkorreliert, so gilt

$$V(X + Y) = V(X) + V(Y)$$

aber auch

$$V(X - Y) = V(X) + V(Y)$$

Allgemein gilt für unkorrelierte ZV X_i

$$V\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 V(X_i)$$

Unabhängige Zufallsvariable

Zwei Zufallsvariable X und Y heißen (stochastisch) **unabhängig** wenn

$$P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$$

für all möglichen Merkmalsausprägungen x und y .

Unabhängige Zufallsvariable sind immer unkorreliert, i.e.

$$X, Y \text{ unabhängig} \Rightarrow \text{Corr}(X, Y) = \text{Cov}(X, Y) = 0$$

Die Umkehrung gilt jedoch **nicht!**

Beispiel 1 – 2-dimensionale Verteilung

Wir suchen für $Y = X_1 + X_2$ Erwartung und Varianz.

Die ZVen X_1 und X_2 besitzen die gemeinsame Verteilung

$P(X_1 = x_1, X_2 = x_2)$	$X_2 = -1$	$X_2 = 1$	$P(X_1 = x_1)$
$X_1 = 1$	0.12	0.48	0.60
$X_1 = 2$	0.08	0.32	0.40
$P(X_2 = x_2)$	0.20	0.80	1.00

Beispiel 1 / summe

Die Werte von $Y = X_1 + X_2$ erhält man über die Tabelle

$Y = X_1 + X_2$	$X_2 = -1$	$X_2 = 1$
$X_1 = 1$	0	2
$X_1 = 2$	1	3

mit der Verteilung

	$Y = 0$	$Y = 1$	$Y = 2$	$Y = 3$	Σ
$P(Y = y)$	0.12	0.08	0.48	0.32	1.00

Beispiel 1 / Erwartungswert $E(Y)$

Für den Erwartungswert von Y erhalten wir

$$E(Y) = \sum y P(Y = y) = 0 \cdot 0.12 + \dots + 3 \cdot 0.32 = 2.0$$

Nach unseren Regeln ergibt sich

$$E(Y) = E(X_1 + X_2) = E(X_1) + E(X_2) = 1.4 + 0.6 = 2.0$$

$$E(X_1) = \sum x_1 P(X_1 = x_1) = 1 \cdot 0.60 + 2 \cdot 0.40 = 1.4$$

$$E(X_2) = \sum x_2 P(X_2 = x_2) = (-1) \cdot 0.20 + 1 \cdot 0.80 = 0.6$$

Beispiel 1 / Varianz $V(Y)$

Für die Varianz von Y erhalten wir

$$\begin{aligned}V(Y) &= E[(Y - E(Y))^2] = E[Y^2] - [E(Y)]^2 = \\ &= 0^2 \cdot 0.12 + \dots + 3^2 \cdot 0.32 - 2.0^2 = 0.88\end{aligned}$$

Nach unseren Regeln ergibt sich

$$\begin{aligned}V(Y) &= V(X_1 + X_2) = V(X_1) + 2\text{Cov}(X_1, X_2) + V(X_2) = \\ &= 0.24 + 2 \cdot 0.00 + 0.64 = 0.88\end{aligned}$$

Beispiel 1 / Varianz $V(Y)$

$$\begin{aligned}V(X_1) &= \sum x_1^2 P(X_1 = x_1) - [E(X_1)]^2 = \\ &= 1^2 \cdot 0.60 + (2^2) \cdot 0.40 - 1.4^2 = 0.24\end{aligned}$$

$$\begin{aligned}V(X_2) &= \sum x_2^2 P(X_2 = x_2) - [E(X_2)]^2 = \\ &= (-1)^2 \cdot 0.20 + 1^2 \cdot 0.80 - 0.6^2 = 0.64\end{aligned}$$

$$\begin{aligned}\text{Cov}(X_1, X_2) &= E[X_1 - E(X_1)][X_2 - E(X_2)] = \\ &= E[X_1 X_2] - E(X_1) E(X_2) = \\ &= \sum_{x_1, x_2} x_1 x_2 P(X_1 = x_1, X_2 = x_2) - E(X_1) E(X_2) = \\ &= 1 \cdot (-1) \cdot 0.12 + 1 \cdot 1 \cdot 0.48 + \dots + 2 \cdot 1 \cdot 0.32 - 1.4 \cdot 0.6 = \\ &= 0 \quad (\text{einfacher: } X_1 \text{ und } X_2 \text{ sind unabhängig})\end{aligned}$$

Beispiel 2 – 2-dimensionale Verteilung

Die ZVen X_1 und X_2 besitzen die gleichen Randverteilungen wie in Beispiel 1. Die gemeinsame Verteilung sei hingegen

$P(X_1 = x_1, X_2 = x_2)$	$X_2 = -1$	$X_2 = 1$	$P(X_1 = x_1)$
$X_1 = 1$	0.00	0.60	0.60
$X_1 = 2$	0.20	0.20	0.40
$P(X_2 = x_2)$	0.20	0.80	1.00

Wir suchen für $Y = X_1 + X_2$ Erwartung und Varianz.

	$Y = 0$	$Y = 1$	$Y = 2$	$Y = 3$	
$P(Y = y)$	0.00	0.20	0.60	0.20	1.00

Beispiel 2 / Erwartungswert $E(Y)$

Für den Erwartungswert von Y erhalten wir

$$E(Y) = \sum y P(Y = y) = 1 \cdot 0.20 + \dots + 3 \cdot 0.20 = 2.0$$

Erwartungswert von X_1 und X_2 sind dieselben wie in Beispiel 1.

Nach unseren Regeln ergibt sich daher wie in Beispiel 1

$$E(Y) = E(X_1 + X_2) = E(X_1) + E(X_2) = 1.4 + 0.6 = 2.0$$

$$E(X_1) = \sum x_1 P(X_1 = x_1) = 1 \cdot 0.60 + 2 \cdot 0.40 = 1.4$$

$$E(X_2) = \sum x_2 P(X_2 = x_2) = (-1) \cdot 0.20 + 1 \cdot 0.80 = 0.6$$

Beispiel 2 / Varianz $V(Y)$

Für die Varianz von Y erhalten wir

$$\begin{aligned} V(Y) &= E[(Y - E(Y))^2] = E[Y^2] - [E(Y)]^2 = \\ &= 1^2 \cdot 0.20 + 2^2 \cdot 0.60 + 3^2 \cdot 0.20 - 2.0^2 = 0.48 \end{aligned}$$

Varianzen von X_1 und X_2 sind dieselben wie im Beispiel 1.

Die Kovarianz ist aber neu zu berechnen.

Nach unseren Regeln ergibt sich daher

$$\begin{aligned} V(Y) &= V(X_1 + X_2) = V(X_1) + 2 \operatorname{Cov}(X_1, X_2) + V(X_2) = \\ &= 0.24 + 2 \cdot (-0.24) + 0.64 = 0.48 \end{aligned}$$

Beispiel 2 / Varianz $V(Y)$

$$V(X_1) = 0.24 \quad (\text{Bsp. 1})$$

$$V(X_2) = 0.64 \quad (\text{Bsp. 1})$$

$$\begin{aligned} \operatorname{Cov}(X_1, X_2) &= E[X_1 - E(X_1)][X_2 - E(X_2)] = \\ &= E[X_1 X_2] - E(X_1) E(X_2) = \\ &= \sum_{x_1, x_2} x_1 x_2 P(X_1 = x_1, X_2 = x_2) - E(X_1) E(X_2) = \\ &= 1 \cdot (-1) \cdot 0.00 + 1 \cdot 1 \cdot 0.60 + \dots + 2 \cdot 1 \cdot 0.20 - 1.4 \cdot 0.6 = \\ &= -0.24 \end{aligned}$$

Beispiel 2 / Korrelation

Durch die Addition der Variablen X_1 zu X_2 wird die Varianz von $X_1 + X_2$ gegenüber der von X_2 alleine deutlich reduziert.

Die Kovarianz ist negativ, $\text{Cov}(X_1, X_2) = -0.24$, und daher auch die Korrelation:

$$\text{Corr}(X_1, X_2) = \frac{\text{Cov}(X_1, X_2)}{\sqrt{V(X_1)V(X_2)}} = \frac{-0.24}{\sqrt{0.24 \cdot 0.64}} = -0.612$$

Beispiel 3 – Unkorrelierte Zufallsvariable

Gegeben seien n unkorreliert Zufallsvariable, X_1, \dots, X_n , die die gleichen Erwartungswerte, $E(X_1) = \dots = E(X_n) = \mu$ und die gleichen Varianzen, $V(X_1) = \dots = V(X_n) = \sigma^2$ besitzen.

$$V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) = \sum_{i=1}^n \sigma^2 = n\sigma^2$$

Die Varianz steigt proportional mit der Anzahl der Summanden.

Die Standardabweichung steigt nur mit der Wurzel der Anzahl der Summanden:

$$\sqrt{V\left(\sum_{i=1}^n X_i\right)} = \sqrt{n\sigma^2} = \sqrt{n}\sigma$$

Beispiel 4 – Diversifikation

Die beiden ZVen R_1 und R_2 sind die Renditen von zwei verschiedenen Wertpapieren. Der aktuell Tageskurs sei bei beiden gleich.

Angenommen die ZVen R_1 und R_2 sind unkorreliert und haben gleichen Erwartungswert und gleiche Varianzen:

$$E(R_1) = E(R_2) = \mu, \quad V(R_1) = V(R_2) = \sigma^2$$

Wir stellen 3 Portfolios zusammen:

- A. Nur Papier 1,
- B. Nur Papier 2,
- C. Papier 1 und Papier 2 je mit einem Anteil von 1/2.

Beispiel 4 – Diversifikation

Wir berechnen von den 3 Portfolios Erwartungswert und Varianz.

$$E(A) = E(R_1) = \mu, \quad V(A) = V(R_1) = \sigma^2$$

$$E(B) = E(R_2) = \mu, \quad V(B) = V(R_2) = \sigma^2$$

$$E(C) = E\left(\frac{1}{2}R_1 + \frac{1}{2}R_2\right) = \frac{1}{2}E(R_1) + \frac{1}{2}E(R_2) = \mu,$$

$$V(C) = V\left(\frac{1}{2}R_1 + \frac{1}{2}R_2\right) = \left(\frac{1}{2}\right)^2V(R_1) + \left(\frac{1}{2}\right)^2V(R_2) = \frac{1}{2}\sigma^2$$

Wir sind indifferent zwischen den Papieren 1 und 2.

Portfolio C liefert hingegen mit derselben erwarteten Rendite nur die halbe Varianz.

Risikostreuung und Korrelation

Die Reduktion der Varianz durch Diversifikation (Risikostreuung) ist auch bei nicht zu stark positiven Korrelationen sinnvoll.

Besonders interessant wird sie bei negativen Korrelationen.

Risikostreuung und $\rho = +1$

Bei starker positiver Korrelation zwischen den Renditen ist die Varianz auch im Portfolio C hoch.

Aus $\text{Corr}(R_1, R_2) = 1$ erhalten wir

$$\text{Cov}(R_1, R_2) = \text{Corr}(R_1, R_2) \sqrt{V(R_1)V(R_2)} = \sigma^2$$

$$E(C) = E\left(\frac{1}{2}R_1 + \frac{1}{2}R_2\right) = \frac{1}{2}E(R_1) + \frac{1}{2}E(R_2) = \mu$$

$$\begin{aligned} V(C) &= V\left(\frac{1}{2}R_1 + \frac{1}{2}R_2\right) = \\ &= \left(\frac{1}{2}\right)^2V(R_1) + 2 \cdot \frac{1}{2} \cdot \frac{1}{2} \text{Cov}(R_1, R_2) + \left(\frac{1}{2}\right)^2V(R_2) \\ &= \sigma^2 \end{aligned}$$

Portfolio C ist nun gleich schlecht wie A oder B.

Risikostreuung und $\rho = -1$

Bei starker negativer Korrelation zwischen den Renditen verschwindet die Varianz im Portfolio C (fast).

Aus $\text{Corr}(R_1, R_2) = -1$ erhalten wir

$$\text{Cov}(R_1, R_2) = \text{Corr}(R_1, R_2) \sqrt{V(R_1)V(R_2)} = -\sigma^2$$

$$E(C) = E\left(\frac{1}{2}R_1 + \frac{1}{2}R_2\right) = \frac{1}{2}E(R_1) + \frac{1}{2}E(R_2) = \mu$$

$$\begin{aligned} V(C) &= V\left(\frac{1}{2}R_1 + \frac{1}{2}R_2\right) = \\ &= \left(\frac{1}{2}\right)^2 V(R_1) + 2 \cdot \frac{1}{2} \cdot \frac{1}{2} \text{Cov}(R_1, R_2) + \left(\frac{1}{2}\right)^2 V(R_2) \\ &= 0 \end{aligned}$$

Optimale Risikostreuung

Angenommen zwei Renditen R_1 und R_2 mit $E(R_1) = \mu_1$, $E(R_2) = \mu_2$ und $V(R_1) = \sigma_1^2$, $V(R_2) = \sigma_2^2$ und Kovarianz σ_{12} liegen vor. Wir suchen die Kombination $\alpha_1 R_1 + \alpha_2 R_2$ mit minimaler Varianz unter der Einschränkung, dass die Summe der Gewichte 1 sei.

$$\min_{\alpha_1, \alpha_2} V[\alpha_1 R_1 + \alpha_2 R_2]$$

$$\text{NB: } \alpha_1 + \alpha_2 = 1$$

Lagrange-Ansatz:

$$L(\alpha_1, \alpha_2; \lambda) = V[\alpha_1 R_1 + \alpha_2 R_2] + \lambda (1 - \alpha_1 - \alpha_2)$$

Optimale Risikostreuung

$$L(\alpha_1, \alpha_2; \lambda) = \alpha_1^2 \sigma_1^2 + 2 \alpha_1 \alpha_2 \sigma_{12} + \alpha_2^2 \sigma_2^2 + \lambda (1 - \alpha_1 - \alpha_2)$$

Notwendige Bedingungen (stationäre Punkte):

$$L_{\alpha_1}: 2 \alpha_1 \sigma_1^2 + 2 \alpha_2 \sigma_{12} - \lambda = 0$$

$$L_{\alpha_2}: 2 \alpha_1 \sigma_{12} + 2 \alpha_2 \sigma_2^2 - \lambda = 0$$

$$L_{\lambda}: 1 - \alpha_1 - \alpha_2 = 0$$

Mit $\alpha_2 = 1 - \alpha_1$ ergibt die Lösung des Gleichungssystems

$$\alpha_1 = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 - 2\sigma_{12} + \sigma_2^2}$$

Varianz einer Linearkombination – Allgemein

Gegeben seien n Zufallsvariable, X_1, \dots, X_n , wobei die X_i die Erwartungswerte μ_i , die Standardabweichungen σ_i und die Kovarianzen σ_{ij} besitzen,

$$E(X_i) = \mu_i, \quad V(X_i) = \sigma_i^2 = \sigma_{ii}$$

$$\text{Cov}(X_i, X_j) = E[(X_i - \mu_i)(X_j - \mu_j)] = \sigma_{ij}$$

Die Varianz einer Linearkombination dieser ZVen, $\sum_{i=1}^n a_i X_i$, ist

$$V\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \sigma_{ij}$$

Varianz einer Linearkombination: $n = 2$

Setzen wir $n = 2$, so erhalten wir

$$\begin{aligned} V\left(\sum_{i=1}^2 a_i X_i\right) &= \sum_{i=1}^2 \sum_{j=1}^2 a_i a_j \sigma_{ij} = \\ &= \sum_{j=1}^2 a_1 a_j \sigma_{1j} + \sum_{j=1}^2 a_2 a_j \sigma_{2j} = \\ &= (a_1 a_1 \sigma_{11} + a_1 a_2 \sigma_{12}) + (a_2 a_1 \sigma_{21} + a_2 a_2 \sigma_{22}) = \\ &= a_1^2 \sigma_1^2 + 2 a_1 a_2 \sigma_{12} + a_2^2 \sigma_2^2 = \\ &= a_1^2 V(X_1) + 2 a_1 a_2 \text{Cov}(X_1, X_2) + a_2^2 V(X_2) \end{aligned}$$

Varianz einer Linearkombination: $n = 3$

Setzen wir $n = 3$, so erhalten wir

$$\begin{aligned} V\left(\sum_{i=1}^3 a_i X_i\right) &= \sum_{i=1}^3 \sum_{j=1}^3 a_i a_j \sigma_{ij} = \\ &= \sum_{j=1}^3 a_1 a_j \sigma_{1j} + \sum_{j=1}^3 a_2 a_j \sigma_{2j} + \sum_{j=1}^3 a_3 a_j \sigma_{3j} = \\ &= a_1^2 \sigma_1^2 + 2 a_1 a_2 \sigma_{12} + 2 a_1 a_3 \sigma_{13} + a_2^2 \sigma_2^2 + 2 a_2 a_3 \sigma_{23} + a_3^2 \sigma_3^2 \end{aligned}$$

Kovarianzmatrix

Sei X ein Spaltenvektor der ZVen X_i und μ der zugehörige Vektor der Erwartungswerte.

$$X = (X_1, \dots, X_n)', \quad \mu = (\mu_1, \dots, \mu_n)'$$

Wir multiplizieren $(X - \mu) \cdot (X - \mu)'$ und erhalten eine $(n \times n)$ -Matrix, in Elementarschreibweise $[(X_i - \mu_i)(X_j - \mu_j)]_{n \times n}$.
Deren Erwartungswert heißt **Kovarianzmatrix** Σ von X ,

$$\Sigma = E[(X - \mu)(X - \mu)'] = [E((X_i - \mu_i)(X_j - \mu_j))]_{n \times n}$$

bzw. mittels Kovarianzen $\sigma_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)]$

$$\Sigma = [\sigma_{ij}]_{n \times n}$$

Varianz in Matrixschreibweise

Die Varianz einer Linearkombination von ZVen in Matrixschreibweise ist

$$V\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \sigma_{ij} = a' \Sigma a$$

wobei $a = (a_1, \dots, a_n)'$ der Spaltenvektor der Koeffizienten a_i ist.

Eigenschaften der Kovarianzmatrix

Die wichtigsten Eigenschaften der Kovarianzmatrix sind:

- Σ ist **symmetrisch**.

Da $\sigma_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)] = E[(X_j - \mu_j)(X_i - \mu_i)] = \sigma_{ji}$ gilt, ist

$$\Sigma = [\sigma_{ij}] = [\sigma_{ji}] = \Sigma'$$

- Σ ist positiv **semidefinit**.

$$a' \Sigma a \geq 0 \quad \text{für alle Vektoren } a.$$

Das heißt, die Varianz einer Summe von ZVen ist immer größer oder gleich Null.

Korrelationsmatrix

Die Korrelation, ρ_{ij} , zwischen X_i und X_j erhält man aus

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii} \sigma_{jj}}}$$

wobei $\sigma_{ij} = \text{Cov}(X_i, X_j)$ und $\sigma_{ii} = \text{Cov}(X_i, X_i) = \text{V}(X_i)$.

Die Korrelationsmatrix R des Vektors X ist

$$R = [\rho_{ij}]_{n \times n} = \text{diag}\{\sigma_{11}, \dots, \sigma_{nn}\}^{-1/2} \cdot \Sigma \cdot \text{diag}\{\sigma_{11}, \dots, \sigma_{nn}\}^{-1/2}$$

Die Hauptdiagonale der Korrelationsmatrix besteht nur aus Einsen.

$$\text{diag}\{\sigma_{11}, \dots, \sigma_{nn}\}^{-1/2} = \text{diag}\{\sqrt{\sigma_{11}}, \dots, \sqrt{\sigma_{nn}}\}^{-1}$$

ist Diagonalmatrix mit Standardabweichungen als Diagonalelemente.

Multivariate Normalverteilung

Sei X eine Spaltenvektor von Zufallsvariablen, $X = (X_1, \dots, X_n)'$, die gemeinsam normal verteilt sind.

$$X \sim N(\mu, \Sigma)$$

$\mu = (\mu_1, \dots, \mu_n)'$ ist der Vektor der Erwartungswerte der X_i .

$$\mu_i = E(X_i)$$

$\Sigma = [\sigma_{ij}]$ ist die Kovarianzmatrix des Vektors X .

$$\sigma_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)]$$

Man schreibt auch $X \sim MN(\mu, \Sigma)$. (multivariat normal)

Beispiel – Bivariate Normalverteilung

Sei X ein Spaltenvektor von Zufallsvariablen, $X = (X_1, X_2)'$, die gemeinsam normal verteilt sind ($n = 2$).

$$X \sim N(\mu, \Sigma)$$

Die Erwartungswerte seien $\mu = (3, 5)'$, die Kovarianzmatrix $\Sigma = [\sigma_{ij}]$ ist gegeben durch $\sigma_{11} = \text{V}(X_1) = 4$, $\sigma_{22} = \text{V}(X_2) = 9$ und $\sigma_{12} = \text{Cov}(X_1, X_2) = 3$.

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 4 & 3 \\ 3 & 9 \end{pmatrix}\right)$$

Cholesky-Zerlegung der Kovarianzmatrix

Jede symmetrische, positiv definite $(n \times n)$ -Matrix Σ kann als Produkt einer unteren Dreiecksmatrix L mit sich selbst zerlegt werden.

$$\Sigma = LL'$$

Es gibt auch die Zerlegung

$$\Sigma = \tilde{L}D\tilde{L}'$$

Hier ist \tilde{L} eine untere Dreiecksmatrix mit Einser in der Hauptdiagonale. D ist eine Diagonalmatrix. Es gilt: $\tilde{L}\sqrt{D} = L$.

(Untere Diagonalmatrix heißt, dass die Elemente oberhalb der Hauptdiagonale sind Null.)

Erzeugung von Normalverteilten Zufallsvektoren

Gesucht ist ein Vektor X mit $X \sim N(\mu, \Sigma)$ mit gegebenem μ und Σ .
Der Ansatz ist

$$X = L\epsilon + \mu$$

mit $\Sigma = LL'$ und $\epsilon \sim N(0, I)$.

- ❶ Berechne Cholesky-Faktor L der Kovarianzmatrix Σ .
- ❷ Generiere beliebig viele Vektoren ϵ der Länge n mit den üblichen Zufallszahlengeneratoren für standard-normalverteilte univariate Zufallszahlen.
- ❸ Einsetzen in $X = L\epsilon + \mu$ liefert multivariat normalverteilte Vektoren mit der gewünschten Kovarianz- bzw. Korrelationsstruktur.

Erzeugung von Normalverteilten Zufallsvektoren

Warum funktioniert die Methode?

- Der Erwartungswert von X ist der gewünschte:

$$E(X) = E(L\epsilon + \mu) = LE(\epsilon) + \mu = \mu$$

da $E(\epsilon) = 0$ gilt.

- Die Kovarianzmatrix von X ist Σ :

$$\begin{aligned} V(X) &= E[(X - \mu)(X - \mu)'] = E[(L\epsilon)(L\epsilon)'] = \\ &= E[(L\epsilon)(\epsilon'L')] = LE[\epsilon\epsilon']L' = LIL' = LL' = \Sigma \end{aligned}$$

- X ist normalverteilt, da jedes X_i aus einer Linearkombination der normalverteilten ϵ_j gebildet wird.
(Vgl. die Reproduktionseigenschaft)

Quantitative Finance

Lernziele für den Teil Quantitative Finance

- Die Welt der stetigen Zinsen (Renditen)
- Wichtige Finanzprodukte:
Aktien, Währungen, Indizes und Optionen
- Bewertung von Call- und Put-Option: Black-Scholes
- Das Binomialmodell zum Simulieren von Aktienkursen und zur Berechnung des Optionspreises.
- Das stochastische Modell für Aktienkurse

Lernziele

- Zeitwert von Geld in stetiger Zeit
- Beispiel einer deterministischen Differentialgleichung
- Finanzinstrumente:
Equities, Commodities, Währungen und Indizes
Futures und Forwards
- Arbitragefreiheit
- Hedgen und Spekulation

Zeitwert und Arbitragefreiheit

Zwei wichtige Konzepte für Quantitative Finance:

- Der **Zeitwert** des Geldes:
1 € heute ist mehr wert als 1 € in einem Jahr.
Das entspricht dem Gegenwartswert-, bzw. Barwertkonzept.
- Das Fehlen von **Arbitragemöglichkeiten**:
Mit Arbitragemöglichkeiten bezeichnen wir Portfolios, die risikolose Gewinne ermöglichen.
Diese soll es in der Theorie nicht geben.
(Sie werden durch den Handel an den Börsen eliminiert.)

Diskretes Zinsmodell

Diskrete Zinsen werden in der *fixed-income* Welt (Anleihen mit langer Laufzeit) verwendet.

$$M(n) = M(0) \cdot \left(1 + \frac{r}{m}\right)^{n \cdot m}$$

$M(n)$... *Kapital (money) nach n Jahren ($n \in \mathbb{N}$)*

$M(0)$... *Anfangskapital*

r ... *einfacher (jährlicher) Zinssatz*

m ... *Anzahl der Verzinsungsperioden*

Stetiges Zinsmodell

Stetige Zinsen dienen als Grundlage zur Herleitung der Preisbildung für Derivative (Optionen, etc.).

$$M(t) = M(0) \cdot e^{rt}$$

$M(t)$... *Kapital (money) nach n Jahren ($t \in [0, \infty)$)*

$M(0)$... *Anfangskapital*

r ... *stetiger Zinssatz*

Die stetige Verzinsung erhält man als Grenzwert für unendlich viele Verzinsungsperioden:

$$e^r = \lim_{m \rightarrow \infty} \left(1 + \frac{r}{m}\right)^m$$

Durchschnittliche Verzinsung

Angenommen wir haben die jährlichen Zinssätze r_1, r_2, \dots, r_n .
Wie lautet der durchschnittliche Zinssatz \bar{r} ?

- Diskretes Zinsmodell:
Geometrisches Mittel der Aufzinsungsfaktoren

$$(1 + \bar{r})^n = (1 + r_1) \cdot (1 + r_2) \cdot \dots \cdot (1 + r_n)$$
$$\bar{r} = \sqrt[n]{(1 + r_1) \cdot \dots \cdot (1 + r_n)} - 1$$

- Stetiges Zinsmodell:
Arithmetisches Mittel der Zinssätze

$$e^{\bar{r}n} = e^{r_1} \cdot e^{r_2} \cdot \dots \cdot e^{r_n}$$
$$\bar{r} = (r_1 + \dots + r_n) / n$$

Diskontieren in stetiger Zeit

Wir deponieren $M(0)$ Geldeinheiten bei einer Bank. Nach dem stetigen Zinseszinsmodell haben wir nach der Zeit t ein Guthaben von $M(t) = M(0) e^{rt}$:

- Der zukünftige Wert ist der *aufdiskontierte* heutige,
 $M(t) = M(0) e^{rt}$;
- der heutige Wert ist der *abdiskontierte* zukünftige,
 $M(0) = M(t) e^{-rt}$.

$M(T)$ Euro zum Zeitpunkt T in der Zukunft ist daher heute, zum Zeitpunkt t ,

$$M(t) = M(T) \cdot e^{-r(T-t)}$$

Euro wert (Gegenwartswert, Barwert, *present value*).

Veränderung des Guthabens

Um wieviel **ändert** sich unser Guthaben mit der Zeit?

- Angenommen im Zeitpunkt t besitzen wir $M(t)$.
- Wir lassen nun eine Periode der Länge dt verstreichen.
- Wir befinden uns somit im Zeitpunkt $(t + dt)$.
Der Wert unseres Guthabens beträgt $M(t + dt)$.
- Unsere Intuition sagt uns:
Die Änderung unseres Guthabens ist ungefähr gleich den einfachen Zinsen $rM(t)$ mal der Länge der (kurzen!) Periode, dt .

$$dM(t) = M(t + dt) - M(t) \approx rM(t) dt$$

Für sehr kurze Perioden dt spielen Zinseszinsen keine (große) Rolle.

Veränderung des Guthabens / (2)

Für „unendlich kleine“ Zeitschritte dt erhalten wir

$$dM(t) = rM(t) dt$$

Die Division durch dt liefert eine **gewöhnliche Differentialgleichung erster Ordnung**.

$$\frac{dM}{dt} = rM(t)$$

Die Veränderung des Guthabens ist gleich „stetiger Zinssatz mal alter Geldbestand“, wie im diskreten Modell.

Veränderung des Guthabens / (3)

Die *Lösung* der Differentialgleichung kann man intuitiv erkennen:

$$M(t) = M(0) \cdot e^{rt}$$

Beweis:

Wir setzen zur *Probe* in die Differentialgleichung ein:

$$\begin{aligned} M'(t) &= \frac{dM}{dt} = rM(t) \\ M(0) r e^{rt} &= r(M(0) e^{rt}) \end{aligned}$$

Stimmt!

Differentialgleichung

- Eine **Differentialgleichung** ist eine Gleichung, die eine Funktionen und deren Ableitung(en) enthält.
- Gesucht ist eine Funktion (die Lösung), die für alle Argumente t die DG erfüllt.
- Eine Differentialgleichung gibt **nur** die Ableitung (i.e. die Steigung der Funktion an jeder Stelle t) an, aber **nicht** deren Lage.
- Zur eindeutigen Bestimmung der Lösung ist stets ein **Anfangswert** oder **Endwert** anzugeben.

In unserem Beispiel ist der Anfangswert $M(0)$, der Wert des Bestandes im Zeitpunkt Null.

Finanzinstrumente

Basisinstrumente der Finanzmärkte sind:

Aktien (*equity, share*), Anleihen (*bonds*), ...

Das sind Eigentumsrechte an einem kleinen Teil eines Unternehmens.

Einige Wertpapiere werden auf einer regulierten Börse (*stock exchange*) notiert, oder können frei gekauft und verkauft werden. Kapital kann so relativ einfach aufgenommen werden.

Dividenden

Der Eigentümer einer Aktie besitzt theoretisch einen Teil einer Unternehmung.

Der „Wert“ einer Aktie wird i.A. für den durchschnittlichen Investor durch die Dividendenzahlungen und dem Wachstum des Wertes der Aktie bestimmt.

Dividenden sind (*lump sum*) Zahlungen, die jedes Quartal oder alle 6 Monate dem Inhaber der Aktie ausbezahlt werden.

Ob man beim Kauf einer Aktie zum Bezug der nächst fälligen Dividenden berechtigt ist, wird durch „cum“ (mit) oder „ex“ (ohne) bezeichnet.

Indizes

Zur Messung der Marktentwicklung werden **Indizes** konstruiert.

Ein typischer Index ist eine gewichtete Summe einer Auswahl oder einem Korb von repräsentativen Aktien.

Die Auswahl soll dabei einen

- gesamten Markt wiederzugeben:
Der Standard & Poor's 500 (S&P500) in den USA oder der Financial Times Stock Exchange Index (FTSE100) in den UK, etc. oder
- einen speziellen Teil des Marktes:
Branchenindizes.

Forwards und Futures

- Ein **Forward Kontrakt** ist eine Übereinkunft, bei der eine Partei verspricht ein Wertpapier von einer anderen Partei zu einem spezifizierten, zukünftigen Zeitpunkt um einen spezifizierten Preis zu kaufen. Kein Geld wechselt bis zum *Liefertag* (*delivery date*) bzw. *Fälligkeit* (*maturity*) des Kontrakts den Besitzer. Die Vertragsbedingungen machen den Kauf zu einer *Verpflichtung* (*obligation*).
- Ein **Future Kontrakt** ist einem *Forward* sehr ähnlich. *Futures* werden üblicherweise zu standardisierten Bedingungen an einer Börse gehandelt. Der Gewinn/Verlust aus der *Future* Position wird täglich errechnet und die Veränderung des Wertes bezahlt eine Partei der anderen. Bei *Futures* gibt es regelmäßig (Teil-)Zahlungen vom Beginn bis zur Fälligkeit.

Aktienpreise und Zufall

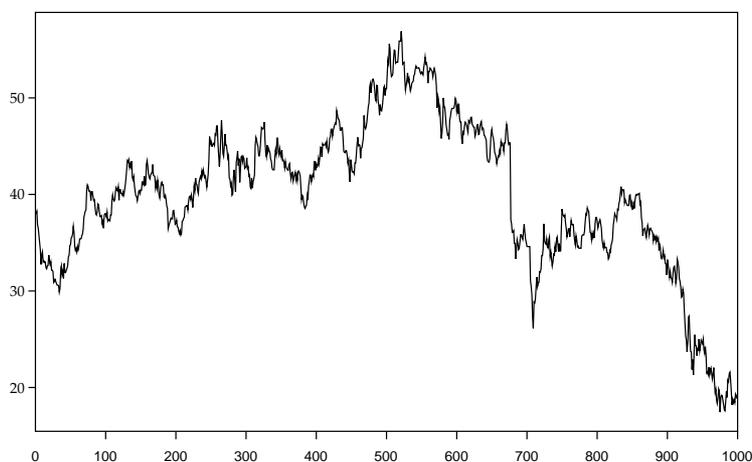
Preise von Aktien besitzen ein großes Element an Zufälligkeit.

Das heißt *nicht*, dass wir Preise nicht modellieren können.
Wir benötigen **stochastische Modelle** für

- Aktienpreis- und Renditenverlauf.
- Berechnung der durchschnittlichen Rendite und Standardabweichung.

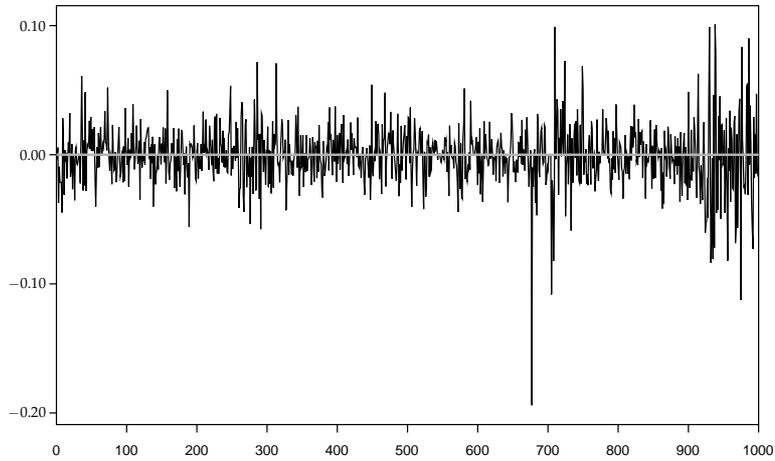
Aktienpreisverlauf

Bayer AG (7. 1. 1999 – 4. 1. 2003)



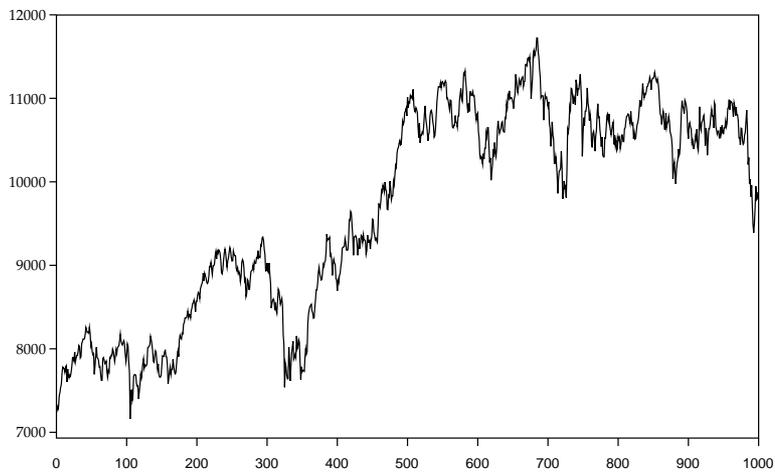
Renditenverlauf

Bayer AG (7. 1. 1999 – 4. 1. 2003)



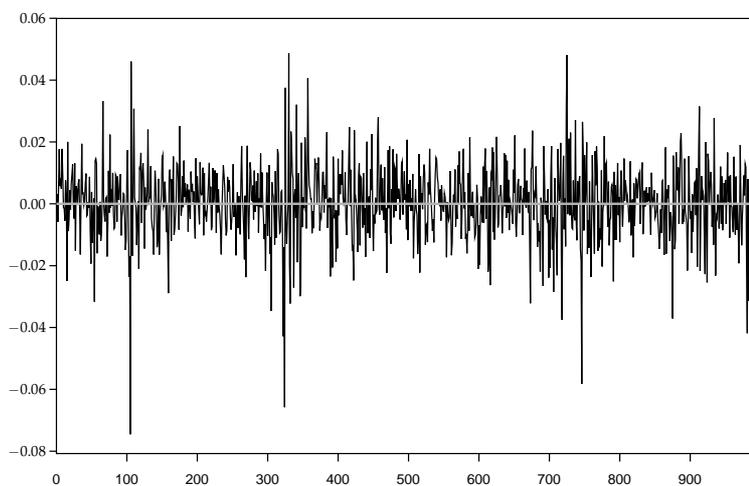
Aktienpreisverlauf

Dow Jones Industrial Index (6. 3. 1997 – 6. 3. 2001)



Renditenverlauf

Dow Jones Industrial Index (6. 3. 1997 – 6. 3. 2001)

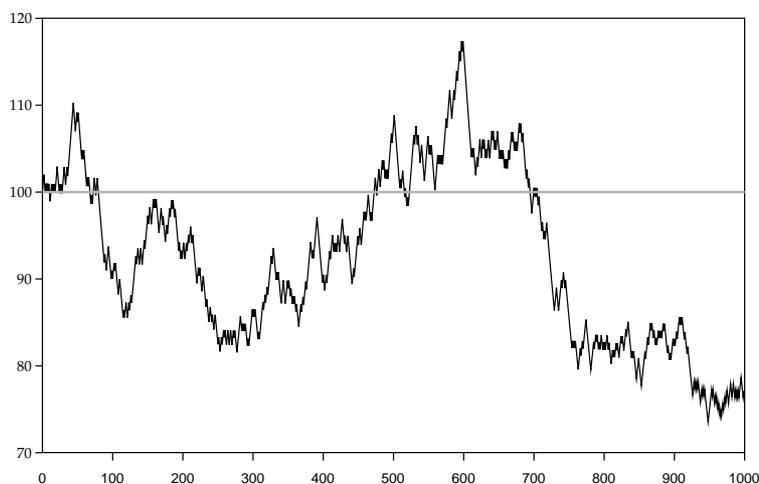


Der Binomialprozess

Der **Binomialprozess** ist ein einfacher stochastischer Prozess, der gewisse Ähnlichkeiten mit Preisprozessen von Aktien hat:

- Beginnen wir mit der Zahl 100. (Preis einer Aktie im Zeitpunkt 0)
- Nun werfen wir eine Münze.
 - Bei „Kopf“ multiplizieren wir die Ausgangszahl mit 1.01;
 - bei „Zahl“ mit 0.99.Nach einem Wurf ist unsere Zahl 99 oder 101.
- Werfen wir nochmals. Abhängig vom Resultat multiplizieren wir wieder mit 1.01 oder 0.99. Es gibt jetzt 3 Möglichkeiten, wobei die mittlere 2-mal auftritt:
 - 100×1.01^2 , ● $100 \times 1.01 \times 0.99$, oder ● 100×0.99^2 .

Ein generierter Preispfad



Spekulation und Hedgen

- **Spekulation:** Wenn jemand glaubt, dass die Preise steigen werden, so kann er möglicherweise davon profitieren, wenn er *Forwards* oder *Futures* kauft. Spekulation ist sehr riskant.
- **Hedging:** *Hedgen* ist das Gegenteil. Es ist die Vermeidung von Risiko.

Hedgen / Beispiel

Angenommen wir erwarten in 6 Monaten eine größere Zahlung in Yen, leben aber in den US und haben unsere Ausgaben in US Dollar zu bestreiten. Wir wollen uns absichern, so dass wir in 6 Monaten tatsächlich den vereinbarten Dollarbetrag erhalten. Dazu können wir einen *Future* Kontrakt eingehen.

Ist einmal der Wechselkurs fixiert, sind wir keinen Schwankungen des Wechselkurses Dollar/Yen mehr ausgesetzt.
(Allerdings können wir auch nicht mehr durch eine mögliche Aufwertung des Yen profitieren.)

Arbitragefreiheit

Wir betrachten einen *Forward*-Kontrakt, der uns verpflichtet $F \text{ €}$ zum Zeitpunkt T zu bezahlen, um ein zugrunde liegendes (*underlying*) Wertpapier (Basisobjekt, *asset*) zu erhalten. Heute ist der Zeitpunkt t und der Preis des Papiers ist gegenwärtig $S(t) \text{ €}$.

Wenn der Fälligkeitszeitpunkt (*maturity*) erreicht ist, bezahlen wir $F \text{ €}$ und erhalten im Gegenzug das Basisobjekt mit dem Wert $S(T) \text{ €}$.

Wieviel Gewinn wir machen, können wir nicht vor dem Zeitpunkt T wissen.

Gibt es eine Beziehung zwischen den Werten F , $S(t)$, t und T ?

Portfolio: Forward, Basisobjekt, Geld

Wir legen ein geeignetes Portfolio an:

- *Forward-Kontrakt abschließen.*
Dieses kostet uns nichts, setzt uns aber der Unsicherheit über den Wert des Basisobjekts zum Zeitpunkt T aus.
- *Basisobjekt verkaufen.*
Es wird als Leerverkauf (*going short*) bezeichnet, wenn wir etwas verkaufen, was wir nicht besitzen. Das ist in vielen Märkten möglich.
- *Bargeld auf die Bank legen.* (Zinssatz r)

Portfolio zur Fälligkeit

Wenn wir den Fälligkeitszeitpunkt erreichen,

- bezahlen wir den Preis F und erhalten das betreffende Basisobjekt.
- Dies schließt die short Position, unabhängig vom Preis $S(T)$.
- Zur Fälligkeit bleibt uns ein garantierte Minusposition, $(-F)$.
- Dem gegenüber steht unser garantiertes Guthaben bei der Bank, dass die anfängliche Investition von $S(t)$ mit den zusätzlichen Zinsen umfasst:

$$S(t) e^{r(T-t)}$$

Das Wort „garantiert“ ist wichtig, da es betont, dass es *unabhängig vom Wert des Basisobjekts* ist.

Spotpreis und Forwardpreis

Wir haben nun einen Betrag $S(t)$ in bar (Verkauf des Basisobjekts), einen *Forward*-Kontrakt, und eine *short* Position mit Wert $S(t)$. Unsere Nettoposition (heute) ist daher Null.

Unsere Nettoposition zur Fälligkeit T ist daher

$$S(t) e^{r(T-t)} - F$$

Unser Portfolio begann mit dem Wert Null in t , und wir enden mit einem vorhersagbarem (sicheren) Ergebnis. Wegen unserer Annahme der Arbitragefreiheit muß dieser Wert ebenfalls Null sein.

Daher gilt die Beziehung zwischen *Spotpreis* und *Forwardpreis*

$$F = S(t) e^{r(T-t)}$$

Cashflow

Cashflows in einem gehedgten Portfolio einer Aktie und zugehörigem *Forward*.

Position	Wert heute (t)	Wert zur Fälligkeit (T)
Forward	0	$S(T) - F$
minus Basisobjekt	$-S(t)$	$-S(T)$
Bargeld	$S(t)$	$S(t) e^{r(T-t)}$
Gesamt	0	$S(t) e^{r(T-t)} - F$

Portfolio Management

Lernziele

- Konzept der modernen Portfolio-Theorie
- Capital Asset Pricing Model
- Optimieren eines Portfolios
- Analyse der Portfolioperformance

Einleitung

Einer der Bausteine der Optionspreistheorie ist Hedging des Basisobjekts mit der Option um ein risikofreies Portfolio zu erhalten.

Nicht jedes Portfolio ist bzw. soll risikofrei sein, nicht jeder Anleger hedgt.

In diesem Abschnitt sehen wir, wie ein riskantes Portfolio mit dem Ziel einen Ertrag zu garantieren und dabei das Risiko zu kontrollieren erstellt werden kann.

Annahmen

Wir treffen folgende Annahmen:

- Wir halten ein Portfolio eine *Periode* gegebener Länge, und untersuchen das Ergebnis am Ende dieser Periode.
- Während dieser Periode sind die Renditen *normalverteilt*.
- Daher genügt es die Renditen durch Erwartungswert und Standardabweichung (*Volatilität*) zu charakterisieren.

Notation

- Wir ziehen N Wertpapiere in die engere Wahl aus denen wir ein Portfolio zusammenstellen wollen. Im Folgenden beziehen wir uns nur auf Aktien.
- Der heutige Wert des Papiers i sei S_i und seine Rendite über unseren **Zeithorizont** T sei R_i .
- R_i ist die zukünftige Rendite und daher nicht bekannt.
- Die Renditen werden als **normalverteilt** angenommen mit Erwartungswert ($\mu_i T$) und Standardabweichung ($\sigma_i \sqrt{T}$).
- Die Korrelationen zwischen den Renditen von Papieren i und Papieren j sind ρ_{ij} .
- Die Parameter μ_i , σ_i und ρ_{ij} beziehen sich auf *Drift*, *Volatilität* und Korrelation, jeweils auf die Zeiteinheit von 1 Jahr bezogen.

Ein Portfolio über den Zeithorizont T

- Ein nomineller Anlagebetrag Π wird auf N Papiere aufgeteilt.
- Wir halten von jedem Papier w_i Stück.
- Der Wert des gesamten Portfolios ist zu Beginn der Periode

$$\Pi = \sum_{i=1}^N w_i S_i$$

- Am Ende der Halteperiode ist der Wert des Portfolios

$$\Pi + \delta\Pi = \sum_{i=1}^N w_i S_i (1 + R_i)$$

Portfoliorendite und Gewichte

Für die relative Veränderung des Portfolios, die Portfoliorendite, $\frac{(\Pi + \delta\Pi) - \Pi}{\Pi} = \frac{\delta\Pi}{\Pi}$, ergibt sich

$$\frac{\delta\Pi}{\Pi} = \sum_{i=1}^N W_i R_i$$

wobei

$$W_i = \frac{w_i S_i}{\sum_{i=1}^N w_i S_i} = \frac{w_i S_i}{\Pi}$$

Die W_i sind Gewichte und geben den *wertmäßigen Anteil* für Papier i an der Gesamtinvestition Π an. Sie summieren sich zu Eins.

$$\sum_{i=1}^N W_i = 1$$

Drift und Volatilität der Portfoliorendite

Die erwartete Rendite unseres Portfolios bezogen auf die Halteperiode der Länge T , $(\mu_{\Pi} T)$, ist dann einfach

$$\mu_{\Pi} T = E\left[\frac{\delta\Pi}{\Pi}\right] = \sum_{i=1}^N W_i E[R_i] = \sum_{i=1}^N W_i (\mu_i T)$$

Der Drift des Portfolios ist

$$\mu_{\Pi} = \sum_{i=1}^N W_i \mu_i$$

Die Volatilität ist analog

$$\sigma_{\Pi} = \sqrt{\sum_{i=1}^N \sum_{j=1}^N W_i W_j \sigma_{ij}} = \sqrt{\sum_{i=1}^N \sum_{j=1}^N W_i W_j \rho_{ij} \sigma_i \sigma_j}$$

Stochastisches Modell der Portfoliorendite

Die Portfoliorendite hat die Form

$$\frac{\delta\Pi}{\Pi} = \mu_{\Pi} T + \sigma_{\Pi} \sqrt{T} Z$$

wobei Z eine standard-normalverteilte Zufallsvariable ist.

$\frac{\delta\Pi}{\Pi}$ ist normalverteilt.

(Reproduktionseigenschaft der Normalverteilung)

Mittel und Standardabweichung der Portfoliorendite hängen nur von den Mittel, Standardabweichungen und Korrelationen (Kovarianzen) der Renditen der Papiere ab.

Portfoliobildung

Einfaches Beispiel

- Angenommen wir haben nur Papiere in unserem Portfolio, die unkorreliert sind, $\rho_{ij} = 0, i \neq j$.
- Jedes Papier soll mit demselben Gewicht eingehen:
 $W_i = 1/N$.
- Der Drift der Portfoliorendite ist dann

$$\mu_{\Pi} = \frac{1}{N} \sum_{i=1}^N \mu_i$$

- Die Volatilität wird zu

$$\sigma_{\Pi} = \sqrt{\sum_{i=1}^N \frac{1}{N^2} \sigma_i^2}$$

Großes Portfolio ($N \rightarrow \infty$)

- Angenommen, alle Papiere haben dieselbe Varianz σ^2 .

$$\sigma_{\Pi} = \sqrt{\sum_{i=1}^N \frac{1}{N^2} \sigma^2} = \sqrt{N \frac{1}{N^2} \sigma^2} = \sqrt{\frac{1}{N} \sigma^2} = \frac{\sigma}{\sqrt{N}}$$

- σ/\sqrt{N} wird auch als $O(1/\sqrt{N})$, „groß O von $1/\sqrt{N}$ “, bezeichnet.
Wir sagen dazu: Die Volatilität ist $O(1/\sqrt{N})$.
Sie konvergiert mit $N \rightarrow \infty$ wie $1/\sqrt{N}$ gegen Null.
- Wenn wir die Anzahl der Papiere in unserem Portfolio erhöhen, geht die Standardabweichung unseres Portfolios gegen Null: Diversifikation reduziert die Volatilität ohne die erwartete Rendite zu reduzieren.

Ziel der Portfoliohaltung

Wir wünschen uns

- eine **hohe** Rendite, und
- eine **geringe** Volatilität.

Volatilität und Risiko:

- Volatilität wird hier als Risiko interpretiert. Große Volatilität bringt ein großes Risiko, dass kleine Renditen erzielt werden, mit. (Natürlich kommen umgekehrt auch hohe Renditen vor.)
- In Hinblick auf unsere Diskussion über die Nutzenfunktion, bedeute kleine Volatilität Risikoaversion. Wir unterstellen also risikoaverses Verhalten.
- Warum nur dieses hier sinnvoll ist, werden wir gleich sehen.

Moderne Portfolio-Theorie

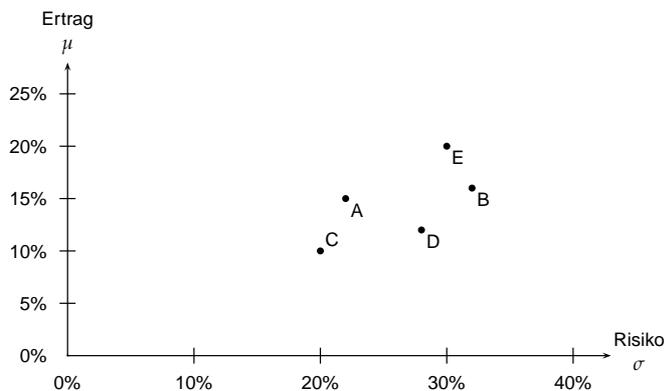
Wir suchen das „beste“ Portfolio, die geeignetste Zusammensetzung einer fix vorgegebenen Auswahl von Wertpapieren.

Markowitz definiert ein **effizientes** Portfolio als eines

- mit dem höchsten Ertrag für ein gegebenes Niveau von Unsicherheit, oder
- das Portfolio mit dem geringsten Risiko bei gegebenem Ertrag.

Zur Illustration legen wir ein $\sigma \times \mu$ -Diagramm an.

$\mu \times \sigma$ -Diagramm



Risiko und Ertrag von fünf Wertpapieren

Diversifikation

Setzen Portfolio aus den Papieren C und E zusammen.

Welchen Effekt haben unterschiedliche Kombinationen auf das Risiko und den Ertrag?

Aus der Formel für den Portfoliodrift und -volatilität erhalten wir

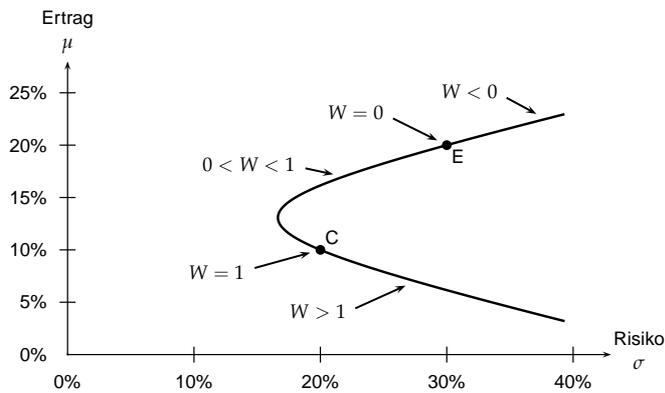
$$\mu_{\Pi} = W \mu_C + (1 - W) \mu_E$$

$$\sigma_{\Pi}^2 = W^2 \sigma_C^2 + 2W(1 - W) \rho_{CE} \sigma_C \sigma_E + (1 - W)^2 \sigma_E^2$$

W ist das Gewicht für Papier C, $(1 - W)$ das Gewicht von E.

Durch einsetzen verschiedener Werte für W erhalten wir eine Hyperbel. (Annahme: $\rho_{CE} = 0$)

Efficient Frontier / Hyperbel



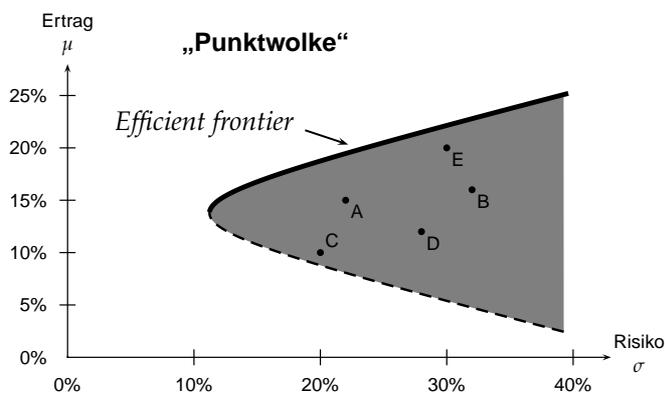
Kombination der Wertpapiere C und E

Efficient Frontier / Punkte

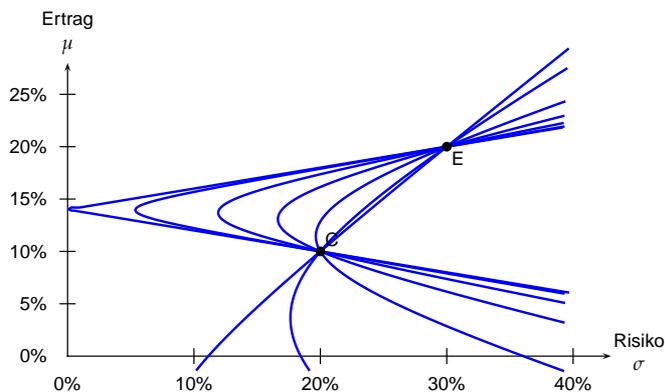
$$R_{\Pi} = W R_C + (1 - W) R_E$$

- $W = 1$: Nur Papier C.
- $W = 0$: Nur Papier E.
- $0 < W < 1$: Punkt auf der Hyperbel zwischen C und E.
- $W > 1$: Punkt unterhalb von C;
long in Papier C, *short* in Papier E.
- $W < 0$: Punkt oberhalb von E;
short in Papier C, *long* in Papier E.

Efficient Frontier



Kombination aller Wertpapiere



Kombination der Wertpapiere C und E bei verschiedenen Korrelationen

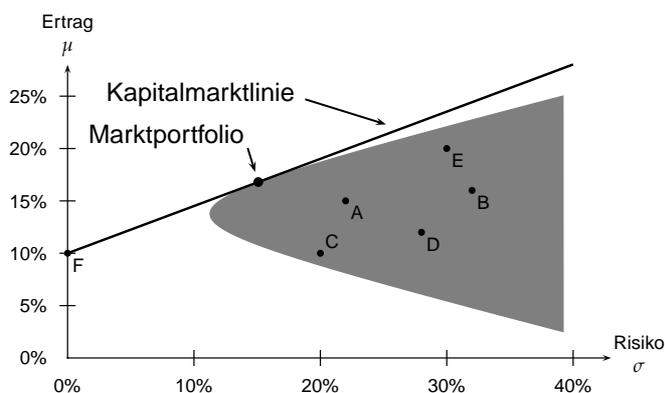
Kapitalmarktklinie und Marktportfolio

Wir betrachten Portfolios, die aus **allen** am Markt gehandelten riskanten Papieren zusammengesetzt werden, und ermitteln deren *Efficient Frontier*. Zusätzlich nehmen wir zusätzlich risikofreie Papiere auf.

Eine risikofreie Investition garantiert eine fixe Rendite von r . Die **Kapitalmarktklinie** beschreibt die *Efficient Frontier* von Portfolios, die risikofreie Papiere im Portfolio aufgenommen haben.

Das **Marktportfolio**, das Portfolio aller Wertpapiere, das sich automatisch am Markt einstellt, liegt am Berührungspunkt der Kapitalmarktklinie und der *Efficient Frontier* der riskanten Papiere.

Portfolio mit risikofreiem Wertpapier



Wahl des individuellen Portfolios

Wo wollen wir auf unserer *Efficient Frontier* sein?

Die „beste“ Position auf der *Efficient Frontier* eines individuellen Portfolios ist subjektiv.

Wir bieten dazu zwei Entscheidungsregeln an:

- Mittels der Steigung der Verbindungslinie zum risikofreien Papier. Das entspricht dem **maximalen Sharpe ratio**.
- Mittels **Nutzenfunktion** des Anlegers.

Maximaler Sharpe ratio

Die Steigung der Verbindungslinie des Portfolios II zum risikofreien Papier – Punkt $(0, r)$ zu Punkt (σ_{II}, μ_{II}) – ist

$$s = \frac{\mu_{II} - r}{\sigma_{II}}$$

heißt *Sharpe ratio* einer Anlage.

Maximierung des *Sharpe ratio* ist hier gleichbedeutend mit der Maximierung der Wahrscheinlichkeit, dass die Anlage II eine größere Rendite als die des risikofreien Zinssatzes r bringt.

Maximaler Sharpe ratio / Eigenschaften

Die Rendite R_{II} des Portfolios II ist normalverteilt,

$$R_{II} \sim \mathcal{N}(\mu_{II}, \sigma_{II}^2)$$

Wahrscheinlichkeit, dass R_{II} größer als der risikofreie Zinssatz r ist,

$$P(R_{II} > r) = P\left(Z > \frac{r - \mu_{II}}{\sigma_{II}}\right) = P(Z > -s) = P(Z \leq s) = \Phi(s)$$

($\Phi(z)$ ist die Verteilungsfunktion der Standardnormalverteilung.)

Je größer s ist, desto größer ist die Wahrscheinlichkeit, dass die Rendite des gewählten Portfolios die einer risikofreien Anlage übertrifft.

Maximaler *Sharpe ratio* / Eigenschaften

Allgemein gilt

$$\Phi\left(\frac{\mu_{\Pi} - r^*}{\sigma_{\Pi}}\right)$$

ist die Wahrscheinlichkeit, dass die Rendite des Portfolios Π eine Rendite r^* übertrifft.

Wenn wir die Wahrscheinlichkeit maximieren wollen, dass wir besser als r^* abschließen, suchen wir auf der *Efficient Frontier* das Portfolio, das den steilsten Anstieg liefert.

$$\max_{\Pi} \frac{\mu_{\Pi} - r^*}{\sigma_{\Pi}}$$

Wahl des Portfolios durch Nutzenmaximierung

Wir zeichnen Indifferenzkurven der gegebenen Nutzenfunktion in das $\sigma \times \mu$ -Diagramm.

Die letzte nach oben verschobene Indifferenzkurve, die noch einen Punkt mit der *Efficient Frontier* gemeinsam hat, bestimmt das Portfolio.

Optimierungsproblem und *Efficient Frontier*

Die *Efficient Frontier* erhält man durch Minimierung der Standardabweichung bei gegebener erwarteter Portfoliorendite, μ_0 .

$$\min_{W_i} V(R_{\Pi})$$

$$\text{NB: } \begin{aligned} E(R_{\Pi}) &= \mu_0 \\ \sum_{i=1}^n W_i &= 1 \end{aligned}$$

Lösung sind die optimalen Gewichte, die wertmäßigen Anteile der einzelnen Aktien.

Die Minimierung wiederholt man für Werte aus einem Intervall von interessanten Renditen, μ_0 : $\mu_0 > r$.

(Minimierung mittels Lagrange-Ansatz)

Diskussion, Numerik

Es sind quadratische Optimierungsprobleme zu lösen für die es Lösungsalgorithmen gibt. Alternativ kann man iterative Minimierungsverfahren herabziehen, die allerdings numerisch suboptimal sind.

Diskussion um die Zuverlässigkeit des Ansatzes:

- Probleme treten i.A. bei der Verwendung der durchschnittlichen beobachteten Renditen auf. Dabei dient die Entwicklung in der Vergangenheit als Prognose für die zukünftige Rendite. Bekanntlich ist es schwierig zukünftige Kurse, wie Renditen vorherzusagen.
(„Vorhersagen ist schwierig, vor allem für die Zukunft.“)
- Die Varianzen der Renditen kennt man i.A. besser.

Diskussion, Numerik

- In Krisensituationen und Crashes brechen plötzlich die Korrelationsstrukturen zusammen, auf die das Verfahren aufbaut. Alle Papiere korrelieren plötzlich hoch positiv. (Z.B. die Kurse fallen alle gleichzeitig.)
Diversifikation hilft nicht, es sei denn sie streuen stärker.
(... , Immobilien, Realitäten, Gold)
- Numerische Problem durch singuläre Korrelationsmatrizen können durch eine aufmerksam erstellte Auswahl von Papieren umgangen werden.
So sollten nicht zwei (oder mehrere) Papiere mit fast identischem Verhalten in das Portfolio genommen werden. (etwa Stamm- und Vorzugsaktie eines Unternehmens)
Für diese Probleme kann die numerische Mathematik Hilfestellungen anbieten.

Capital Asset Pricing Model (CAPM)

Das *Capital Asset Pricing Model*, **CAPM**, dient zur Bewertung einzelner Papiere bezüglich des Marktes, der durch einen Index repräsentiert wird.

Ein Index ist mit einem Portfolio vergleichbar: DAX30, S&P500, etc.

Dazu wird ein Beta-Koeffizient berechnet. Das Beta, β , eines Papiers **relativ zum Marktportfolio** M ist der Quotient aus der Kovarianz zwischen der Rendite des Papiers und der Rendite des Portfolios, und der Varianz der Rendite des Portfolios.

$$\beta_i = \frac{\text{Cov}(R_i, R_M)}{V(R_M)}$$

Die Interpretation erfolgt im Rahmen des *Single-Index Modells*.

Single-Index Modell

Wir setzen die Rendite jedes einzelnen Papiers zur Rendite **eines repräsentativen Index** M in Beziehung

$$R_i = \alpha_i + \beta_i R_M + \epsilon_i$$

Das lineare Modell für R_i besteht aus drei Teilen:

- Einem konstanten Drift, α_i ,
- Einer gemeinsamen Komponente mit dem Marktindex R_M , $\beta_i R_M$, und
- einem zufälligen Teil ϵ_i , der unkorreliert mit dem Index ist.

Die Koeffizienten werden durch Schätzen des linearen Regressionsmodells ermittelt.

Single-Index Modell / Interpretation

- α_i ist die Konstante in der Regressionsgleichung.
 $R_i = \alpha_i$ (im Durchschnitt), wenn $R_M = 0$.
Also, wenn $\alpha_i \neq 0$, besitzt Papier i eine autonome Rendite, auch wenn der Markt nur Null erwirtschaftet.
- β_i ist der Steigungskoeffizient.
Steigt die Marktrendite um 1, so steigt R_i um β_i .
 - Papiere, die sich mit dem Markt bewegen, haben positive Koeffizienten; Papiere, die sich gegen den Markt bewegen, negative (Gold?).
 - Papiere mit einem $\beta_i > 1$ reagieren stärker als der Markt; Papiere mit einem $\beta_i < 1$ schwächer.
- Der Fehler der Regression, ϵ_i , ist mit R_M unkorreliert.

Single-Index Modell und Marktindex

Sei μ_M die erwartete Rendite des Index, und σ_M deren Standardabweichung.

- Erwartete Rendite des i -ten Papiers:

$$\mu_i = \alpha_i + \beta_i \mu_M$$

- Standardabweichung:

$$\sigma_i = \sqrt{\beta_i^2 \sigma_M^2 + e_i^2}$$

- e_i^2 ist die Varianz von ϵ_i .

Single-Index Modell und Portfoliorendite

Die Rendite eines Portfolio Π ist

$$\frac{\delta\Pi}{\Pi} = \sum_{i=1}^N W_i R_i = \left(\sum_{i=1}^N W_i \alpha_i \right) + R_M \left(\sum_{i=1}^N W_i \beta_i \right) + \sum_{i=1}^N W_i \epsilon_i$$

Die erwartete Rendite ist ($E(\epsilon_i) = 0$)

$$\mu_{\Pi} = \left(\sum_{i=1}^N W_i \alpha_i \right) + E(R_M) \left(\sum_{i=1}^N W_i \beta_i \right)$$

Wir erhalten daher für die erwartete Portfoliorendite

$$\mu_{\Pi} = \alpha_{\Pi} + \beta_{\Pi} \mu_M$$

mit

$$\alpha_{\Pi} = \sum_{i=1}^N W_i \alpha_i, \quad \beta_{\Pi} = \sum_{i=1}^N W_i \beta_i, \quad \mu_M = E(R_M)$$

Single-Index Modell und Portfoliorendite

Die Volatilität in Π ergibt sich als

$$\sigma_{\Pi} = \sqrt{\sum_{i=1}^N \sum_{j=1}^N W_i W_j \beta_i \beta_j \sigma_M^2 + \sum_{i=1}^N W_i^2 e_i^2}$$

Dies geht so einfach, weil alle ϵ_i unkorreliert mit R_M sind.

Unter der vereinfachten Annahme, dass alle Gewichte gleich ($1/N$) sind, alle $e_i^2 = e^2$ und alle $\beta_i = \beta$ ergibt sich für die Volatilität der Portfoliorendite

$$\begin{aligned} \sigma_{\Pi}^2 &= \sum_{i=1}^N \sum_{j=1}^N (1/N) (1/N) \beta \beta \sigma_M^2 + \sum_{i=1}^N (1/N)^2 e^2 \\ &= N^2 (1/N)^2 \beta^2 \sigma_M^2 + N (1/N)^2 e^2 \\ &= \beta^2 \sigma_M^2 + e^2 / N \end{aligned}$$

Single-Index Modell und Diversifikation

Die Volatilität der Portfoliorendite

$$\sigma_{\Pi}^2 = \beta^2 \sigma_M^2 + e^2 / N$$

besteht aus 2 Teilen:

- Das **diversifizierbare Risiko** ist mit den ϵ_i verbunden: e^2/N . Dessen Beitrag zur Varianz verschwindet mit steigendem N , ($O(N^{-1/2})$).
- Das **systematische Risiko** korreliert mit dem Index: $\beta^2 \sigma_M^2$. Es kann durch Portfoliobildung nicht reduziert werden.

Single-Index Modell und optimales Portfolio

Die Optimierung unter Verwendung des *Single-Index* Modells sieht folgende Schritte vor:

1. Berechne α_i , β_i und e_i^2 zu allen Papieren.
2. Wähle einen Wert für die erwartete Portfoliorendite μ_{Π} .
3. Minimiere σ_{Π} unter dieser Nebenbedingung.
4. Wiederhole die Minimierung für verschiedene Portfoliorenditen um die *Efficient Frontier* zu erhalten.
5. Die Position auf der Kurve ist subjektiv zu entscheiden, oder nach dem maximalen *Sharpe ratio* Kriterium.

Multi-Index Modell

Die Idee des *Single-Index* Modells kann um andere repräsentative (i.A. korrelierte) Indizes R_{M_j} erweitert werden.

Zum Beispiel kann zusätzlich zu einem Aktienmarktindex

- ein repräsentativer Bondmarktindex,
- ein Index für Währungsmärkte, oder
- ein volkswirtschaftlicher Index

in die Regression miteinbezogen werden, wenn man glaubt, dass er für die Papiere von Bedeutung ist.

$$R_i = \alpha_i + \sum_{j=1}^k \beta_{ij} R_{M_j} + \epsilon_i$$

Performance-Messung

Wie misst man die Performance von Anlagestrategien?

Ein Kriterium basiert auf den Vergleich mit der risikofreien Rate. Ideal wäre die risikofreien Rate **konsistent** zu überbieten.

Ein anderes Kriterium hat zum Ziel nicht nur eine hohe Rendite zu erzielen, sondern dies auch mit so wenig Varianz wie möglich zu bewältigen.

Sharpe ratio und Treynor ratio

Zwei der häufigsten Maßzahlen von „Rendite pro Risikoeinheit“

- *Sharpe ratio* setzt „Ertrag zur Variabilität“ in Beziehung:

$$\text{Sharpe ratio} = \frac{\mu_{\Pi} - r}{\sigma_{\Pi}}$$

- *Treynor ratio* setzt „Ertrag zur Volatilität“ in Beziehung:

$$\text{Treynor ratio} = \frac{\mu_{\Pi} - r}{\beta_{\Pi}}$$

μ_{Π} und σ_{Π} bzw. β_{Π} sind die realisierten Werte in der Beobachtungsperiode.

σ_{Π} ist die beobachtete Standardabweichung,

β_{Π} ist ein Maß für die beobachtete „Volatilität“ des Portfolios.

Sharpe ratio und Treynor ratio

- Der *Sharpe ratio* wird üblicherweise verwendet, wenn man das Portfolio aller Investitionen betrachtet.
- Der *Treynor ratio* wird verwendet, wenn einzelne Komponenten eines gesamten Unternehmensportfolios bewertet werden.