§ 1 Elementare statistische Grundlagen

Literatur:

Bleymüller, J. u. a. (1985): Statistik für Wirtschaftswissenschaftler. 4. Aufl., München 1985, Kapitel 7 und 8.

DeFusco, R. A. u. a. (2015): Quantitative Investment Analysis. 3rd ed., Hoboken, N.J., S. 151 – 183, 199 – 232, 327 – 347.

Schäfer, H. (1999): Unternehmensinvestitionen. Grundzüge in Theorie und Management. Heidelberg, S. 280 – 290.

1.1 Eindimensionale Zufallsvariable

(1) Begriff der Zufallsvariablen

(a) Zufallsvariable X

- Variable, deren Wert vom Zufall abhängt
- Beispiel: Rendite eines Wertpapiers für einen bestimmten zukünftigen Zeitraum

(b) Realisation (Ausprägung) x

der einzelne Wert, den die Zufallsvariable annimmt

(c) Diskrete Zufallsvariable:

- besitzt nur endlich viele Ausprägungen
- Beispiel: Zufallsexperiment „Zweimaliges Werfen einer Münze“

- Zufallsvariable X: „Anzahl Wappen“

o Ausprägungen:	$x_1 = 0$,	$x_2 = 1$,	$x_3 = 2$
	↑	↑	↑
o Elementarereignisse: des Zufallsexperiments “Zweimaliges Werfen einer Münze“	zz	wz, zw	ww

(d) Stetige Zufallsvariable:

- besitzt unendlich viele Ausprägungen (in einem bestimmten Bereich der reellen Zahlen)

- Beispiel: Länge des aus einer Produktionsmenge zufällig ausgewählten Werkstückes

- beachte: aus Gründen der nicht beliebig erhöhbaren Maßgenauigkeit lassen sich in der Praxis Zufallsvariable nur diskret erfassen

(2) Wahrscheinlichkeitsfunktion, -verteilung

(a) Diskrete Zufallsvariable

- Funktion $f(x_j)$, die für jede Ausprägung der Zufallsvariablen die Wahrscheinlichkeit ihres Auftretens angibt:

  $f(x_j) = W(X = x_j) = w_j, \qquad j = 1, \ldots, J$

  Es gilt: $f(x_j) \geq 0, \quad j = 1, \ldots, J$ und $\sum_j f(x_j) = 1$

➔ Wahrscheinlichkeiten $f(x_j)$ haben formal die gleichen Eigenschaften wie die relativen Häufigkeiten von empirischen Verteilungen in der deskriptiven Statistik

- Beispiel: dreimaliges Werfen einer Münze
    - acht gleichwahrscheinliche Elementarereignisse e_1, e_2, …, e_8
    - erste Spalte = Definitionsbereich des Zufallsexperiments
    - zweite Spalte: Wahrscheinlichkeiten der Elementarereignisse

Elementarereignis e	Wahrscheinlichkeit W(e)	„Anzahl Wappen" X	Wahrscheinlichkeit W(X=x) = f(x)
e_1 = ZZZ	$W(e_1) = 0{,}125$	$x_1 = 0$	$f(x_1) = 0{,}125$
e_2 = ZZW e_3 = ZWZ e_4 = WZZ	$W(e_2) = 0{,}125$ $W(e_3) = 0{,}125$ $W(e_4) = 0{,}125$	$x_2 = 1$	$f(x_2) = 0{,}375$
e_5 = ZWW e_6 = WZW e_7 = WWZ	$W(e_5) = 0{,}125$ $W(e_6) = 0{,}125$ $W(e_7) = 0{,}125$	$x_3 = 2$	$f(x_3) = 0{,}375$
e_8 = WWW	$W(e_8) = 0{,}125$	$x_4 = 3$	$f(x_4) = 0{,}125$

Ableitung der Wahrscheinlichkeitsfunktion der Zufallsvariablen X: „Anzahl der Wappen" bei dreimaligem Werfen einer Münze

- dritte Spalte: Realisation der Zufallsvariablen

  X = "Anzahl Wappen pro dreimaligen Wurf" (d.h. pro Versuch)

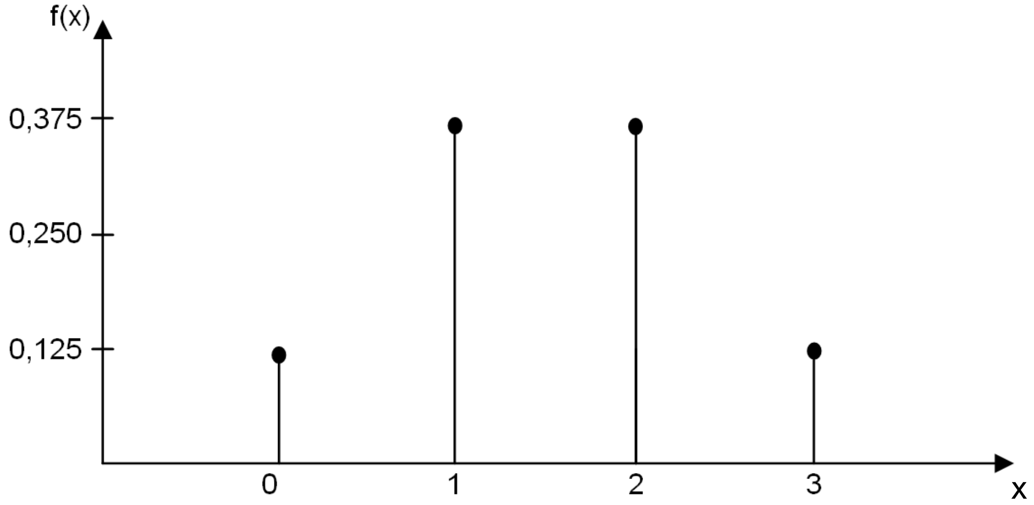
  → Zufallsvariable lässt sich als Funktion auffassen, die jedem Elementarereignis e_k eine reelle Zahl zuordnet:
  $X(e_k) = x_j$

- vierte Spalte: Wahrscheinlichkeitsfunktion

- man erkennt: Wahrscheinlichkeit, dass X die spezielle Ausprägung x_j annimmt, ergibt sich als Summe der Wahrscheinlichkeiten derjenigen Elementarereignisse e_k, denen die Ausprägung x_j zugeordnet ist:

$$W(X=x_j) \quad = \sum_{X(e_k) = x_j} W(e_k) \quad = \quad w_j$$

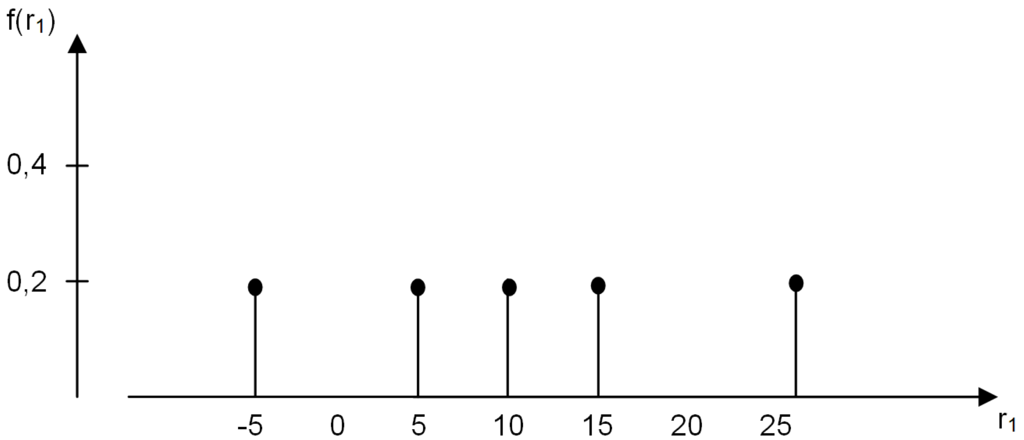
- grafische Darstellung: Stabdiagramm



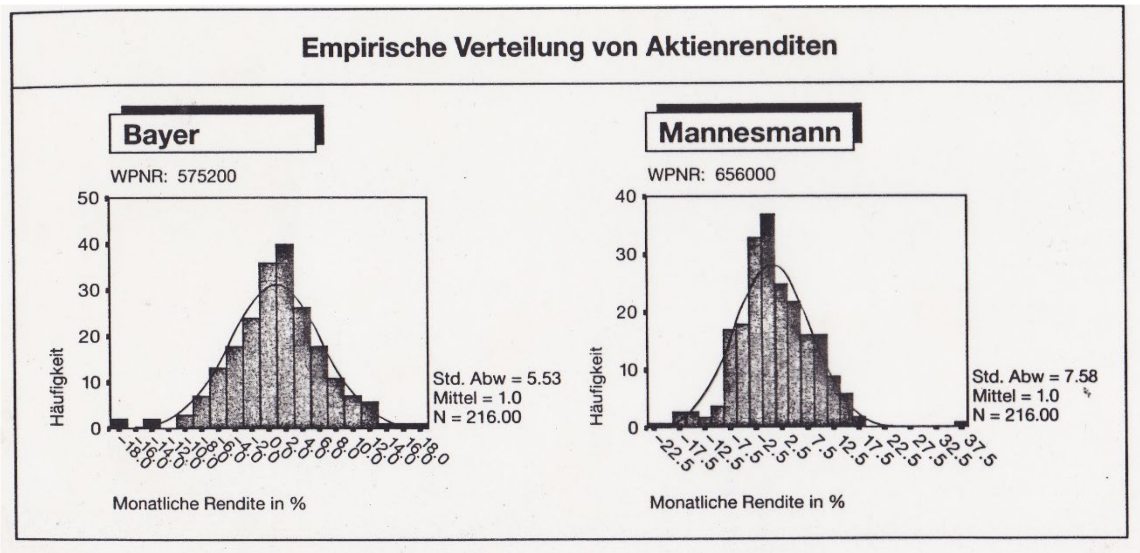
Wahrscheinlichkeitsfunktion einer diskreten Zufallsvariable
Quelle: Bleymüller u.a. (1985), S. 40

- Beispiel: risikobehaftete Anlage → Rendite ist Zufallsvariable

	s_1	s_2	s_3	s_4	s_5
R_1	$r_{11} = 15$	$r_{12} = 5$	$r_{13} = -5$	$r_{14} = 25$	$r_{15} = 10$
$w_j = f(r_{1j})$	$f(r_{11}) = 0{,}2$	$f(r_{12}) = 0{,}2$	$f(r_{13}) = 0{,}2$	$f(r_{14}) = 0{,}2$	$f(r_{15}) = 0{,}2$



Die folgende Abbildung zeigt mittels Flächen von Säulen ("Histogramm") die Häufigkeiten der in der Vergangenheit tatsächlich gemessenen monatlichen Aktienrenditen von Bayer und Mannesmann. Wie bereits erwähnt, werden Vergangenheitswerte oft als Indikatoren für die Prognose zukünftiger Renditen herangezogen. Damit ließen sich die Histogramme als Wahrscheinlichkeitsfunktion der Zufallsvariablen "zukünftige Aktienrendite" auffassen. Man erkennt, dass die Renditen – anders als im vorherigen Beispiel – nicht gleich-, sondern glockenförmig verteilt sind.



Quelle: H. Graz u. a., Portfolio-Management. Frankfurt/M. 1997, S. 31

(b) Stetige Zufallsvariable

- Dichtefunktion f(x) ist stetige Funktion
- es gilt: $f(x) \geq 0$ und $\int_{-\infty}^{+\infty} f(x)\,dx = 1$
- da die Zufallsvariable unendlich viele Werte annehmen kann, ist die Wahrscheinlichkeit, dass irgendein spezieller Wert angenommen wird, immer 0:

  $W(X=x) = 0$

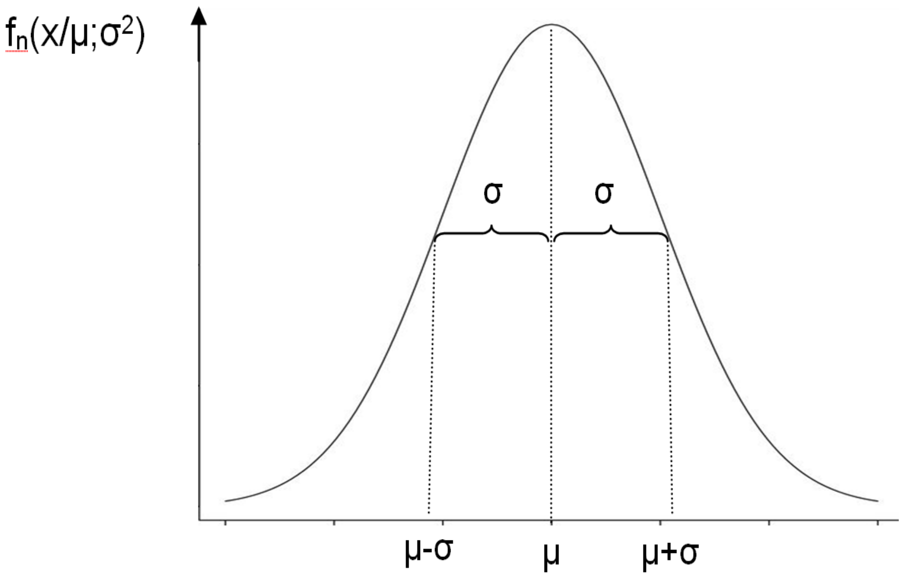
  → nur Wahrscheinlichkeit sinnvoll auszugehen, dass X einen Wert im Intervall [a,b] annimmt:

$$W\ (a \leq X \leq b) = \int_a^b f(x)\, dx$$

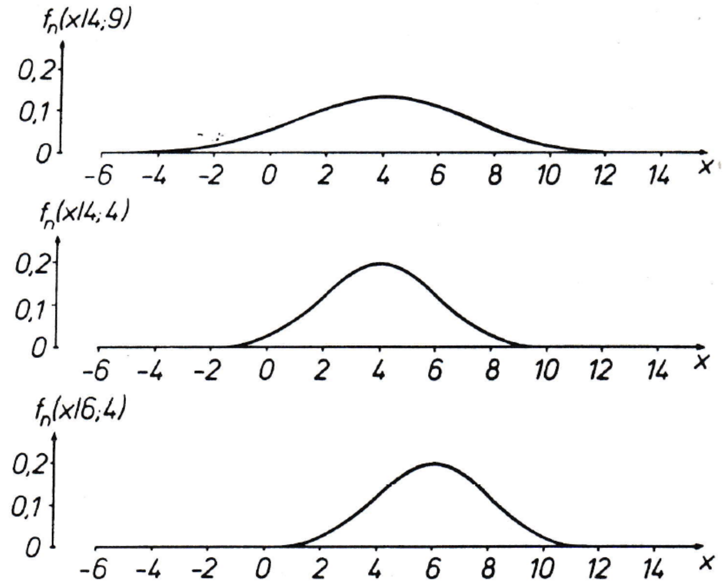
- wichtigste stetige Verteilung: Normalverteilung:
  - symmetrische Dichtefunktion:

$$f_n(x/\mu;\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} * e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

  - Maximum bei $x = \mu$
  - Wendepunkte bei $x = \mu + \sigma$ und $x = \mu - \sigma$



- in der folgenden Abbildung sind Dichtefunktionen von Normalverteilungen für alternative μ-σ-Kombinationen dargestellt:



Vergleich von Normalverteilungsdichten mit verschiedenen Parametern μ und σ²

Quelle: Bleymüller u. a. (1985), S. 60

- Portfoliotheorie unterstellt aus zwei Gründen normalverteilte Renditen:

  - obige Abbildung von Aktienrenditen verdeutlicht, dass Normalverteilung eine gute Approximation darstellt

  - die Dichtefunktion lässt sich durch nur zwei Parameter hinreichend beschreiben; dies bedeutet wiederum zweierlei:

    (i) Dateninput kann komprimiert werden für die Anlageentscheidung, ohne dass es zu unzulässigem Informationsverlust und damit zu einer unangemessenen Entscheidung kommt

    (ii) es lässt sich eine eindeutige Beziehung zwischen einer auf die Rendite bezogenen Risikonutzenfunktion und μ-σ-Nutzenfunktionen feststellen; m. a. W.: jeder Risikonutzenfunktion entspricht ein–eindeutig ein Indifferenzkurvensystem im μ-σ-Diagramm. So entspricht dann der für Risikonutzenfunktionen definierten Eigenschaften der Risikoaversion (Konkavität von μ, d.h. $\mu^{II} < 0$) in der μ-σ-Analyse die Eigenschaft $\delta v / \delta\sigma < 0$; diese Eigenschaft ist unter der Annahme $\delta v / \delta\mu > 0$ äquivalent mit positiver Steigung der μ-σ-Indifferenzkurven. Analog entsprechen der Risikoneutralität ein horizontaler Verlauf und der Risikofreude eine negative Steigung der Indifferenzkurven.

(3) Erwartungswert und Varianz von Zufallsvariablen

Wie die Häufigkeitsverteilung der deskriptiven Statistik lassen sich auch die Wahrscheinlichkeitsverteilungen von Zufallsvariablen durch Maßzahlen charakterisieren:

(a) Erwartungswert

Dem arithmetischen Mittel als Lageparameter der Häufigkeitsverteilung entspricht hier der Erwartungswert E(X).

- diskrete Zufallsvariable:

$$E(X) = \sum x_j * f(x_j) = \sum x_j * w_j$$

Beispiel: “Dreimaliges Werfen einer Münze“ → s. o. (2a)

- E(X) = “Anzahl Wappen“, die man bei einer größeren Zahl von Versuchswiederholungen im Durchschnitt pro Versuch erwarten kann

- $E(X) = \sum_{j=1}^{4} x_j * f(x_j)$

$$= 0 * 0{,}125 + 1 * 0{,}375 + 2 * 0{,}375 + 3 * 0{,}175$$

$$= 1{,}5$$

Beispiel: Wertpapierrendite aus (2a)

$$E(R_1) = 10$$

- stetige Zufallsvariable:

  $$E(X) = \int_{-\infty}^{+\infty} x * f(x)\, dx$$

  falls x nur in einem Intervall $x_n \leq x \leq x_0$ positive Werte annimmt, gilt:

  $$E(X) = \int_{x_n}^{x_0} x * f(x)\, dx$$

- für eine normalverteilte stetige Rendite R gilt:

  $E(R) = \mu$

  - der Parameter μ der Verteilung stimmt gerade mit dem Erwartungswert der geschätzten Rendite überein, charakterisiert also die Lage der Normalverteilung

  - μ lässt sich also als eine Art Mittelwert der Renditenschätzungen anschaulich deuten
    → weiterer Grund für Beliebtheit der Normalverteilung in der Portfoliotheorie

(b) Varianz

Wie bei Häufigkeitsverteilungen findet auch bei Wahrscheinlichkeitsverteilungen die Varianz als Streuungsparameter Anwendung.

- diskrete Zufallsvariable

$$\mathrm{Var}(X) = \sum_j [x_j - E(X)]^2 \bullet f(x_j)$$

$$= \sum_j [x_j - E(X)]^2 \bullet w_j$$

$$= \sum x^2_j \bullet w_j - [E(X)]^2$$

Beispiel: dreimaliges Werfen einer Münze → s. o. (2a)

$$\mathrm{Var}(X) = \sum_{j=1}^{4} x^2_j \bullet w_j\ [E(X)]^2$$

$$= 0 * 0{,}125 + 1 * 0{,}375 + 4 * 0{,}375 + 9 * 0{,}125 - 1{,}5^2$$

$$= 0{,}75$$

Beispiel Wertpapier aus (2a): $\mathrm{Var}(R_1) = 100$

- stetige Zufallsvariable

$$\mathrm{Var}(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 * f(x)\, dx$$

$$= \int_{-\infty}^{+\infty} x^2 * f(x)\, dx - [E(X)]^2$$

für eine normalverteilte stetige Rendite R gilt: $\mathrm{Var}(R) = \sigma^2$

→ zwei weitere Gründe für die Beliebtheit der Normalverteilung:

- Parameter σ^2 der Verteilung stimmt gerade mit Varianz der geschätzten Rendite überein, charakterisiert also deren Streuung
- er lässt sich als Risiko des Wertpapieres interpretieren; wegen ihrer Symmetrie vermag die Normalverteilung dabei insbesondere zu verdeutlichen, dass der Begriff Risiko …
  - … sowohl die Chance dass $r_j > E(R)$
  - … als auch die Gefahr dass $r_j < E(R)$ umfasst

(4) Rechnen mit Erwartungswert und Varianz einer eindimensionalen Zufallsvariablen

(a) Erwartungswertmethode von Funktionen der Zufallsvariablen X

$Y = g(X)$

- diskreter Fall

$$E(Y) = E[g(X)] = \sum_j g(x_j) * f(x_j)$$

- stetiger Fall:

$$E(Y) = E[g(X)] = \int_{-\infty}^{+\infty} g(x) * f(x)\,dx$$

(b) Beispiel: $Y = [X - E(X)]^2$

Man erhält:

$$E(Y) = E\{[X - E(X)]^2\}$$

$$= \sum_j [x_j - E(X)]^2 * f(x_j)$$

$$= Var(X)$$

Also:

$$Var(X) = E\{[X - E(X)]^2\}$$

(c) Formelsammlung für lineare Funktionen einer einzelnen Variablen

Y	**E(Y)**	**Var(Y)**
A	a	0
bX	$b * E(X)$	$b^2 * Var(X)$
a + x	$a + E(X)$	$Var(X)$
a + bX	$a + b * E(X)$	$b^2 * Var(X)$

Quelle: Bleymüller u. a. (1985), S. 43

1.2 Zweidimensionale Zufallsvariable

(1) Gemeinsame Wahrscheinlichkeitsfunktion

(a) Ausgangspunkt:

Zufallsexperiment, als dessen Ergebnis man zwei Zufallsvariable betrachtet

(b) Gemeinsame Wahrscheinlichkeitsfunktion

Gibt die Wahrscheinlichkeit dafür an, daß R_1 den Wert r_{1j} und R_2 den Wert r_{2l} annimmt.

$W(R_1 = r_{1j} \wedge R_2 = r_{2l}) = f(r_{1j}, r_{2l})$, $j = 1, ..., J; l = 1, ..., L$

Wobei $f(r_{1j}, r_{2l}) \geq 0$

$$\sum_j \sum_l f(r_{1j}, r_{2l}) = 1$$

(2) Kovarianz und Korrelationskoeffizient

(a) Ausgangspunkt

- Positive Korrelation:
  bei großen Werten von R1 nimmt tendenziell auch R2 große Werte an

- Negative Korrelation:
  bei großen Werten von R1 nimmt tendenziell R2 kleine Werte an

(b) Kovarianz

$$Cov(R_1, R_2) = E\{[r_1 - E(R_1)][r_2 - E(R_2)]\}$$

$$= \sum_j \sum_l \left[r_{1j} - E(R_1)\right]\left[r_{2l} - E(R_2)\right] w_{lj}$$

$$= \underbrace{\sum_j \sum_l r_{1j} r_{2l} w_{lj}}_{} - E(R_1)\, E(R_2)$$

$$= \quad E(R_1 R_2) \qquad - E(R_1)\, E(R_2)$$

(c) Korrelationskoeffizient

$$\rho(R_1 R_2) = E\left\{ \frac{\left[r_1 - E(R_1)\right]}{\sigma_{R1}} \frac{\left[r_2 - E(R_2)\right]}{\sigma_{R2}} \right\}$$

$$wobei \quad \sigma_{R1} = \sqrt{Var(R_1)}, \quad \sigma_{R2} = \sqrt{Var(R_2)}$$

- Es gilt:

$$\rho\,(R_1 R_2) = \frac{E\,\left\{\left[r_1 - E(R_1)\right]\left[r_2 - E(R_2)\right]\right\}}{\sigma_{R1}\sigma_{R2}}$$

$$= \frac{Cov(R_1, R_2)}{\sigma_{R1}\sigma_{R2}}$$

- es lässt sich zeigen:

$$-1 \leq \rho \leq 1$$

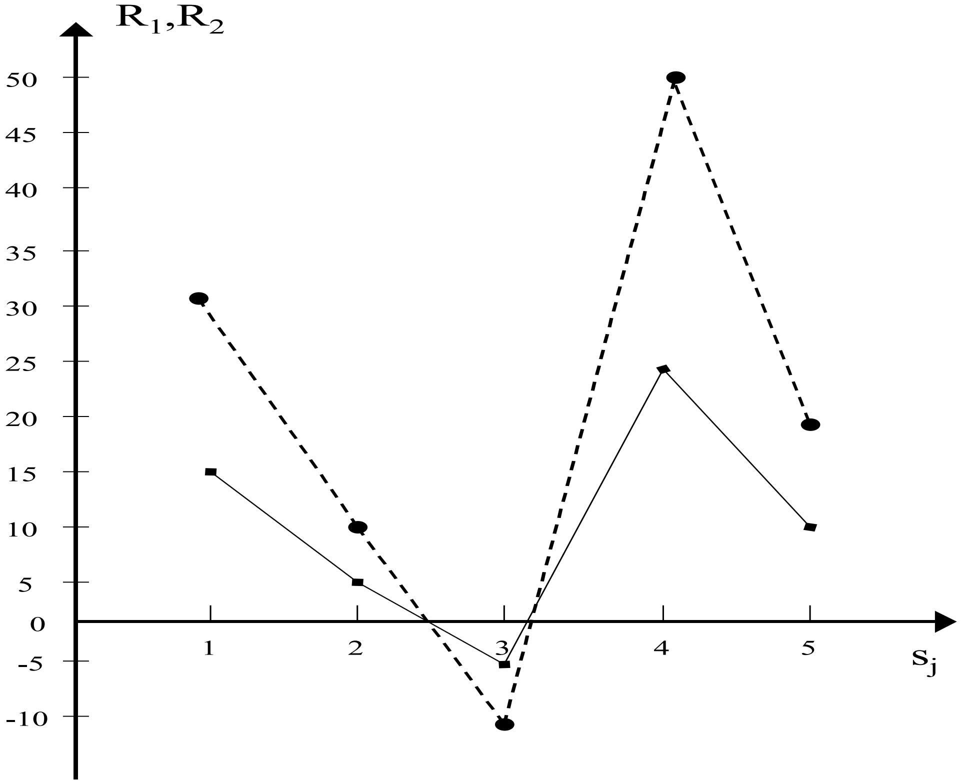
- bei zwei unabhängigen Zufallsvariablen gilt:

$Cov\ (R_1,R_2) = 0$

$\Rightarrow \rho\ (R_1,R_2) = 0$

(3) Zur Bedeutung der Korrelation

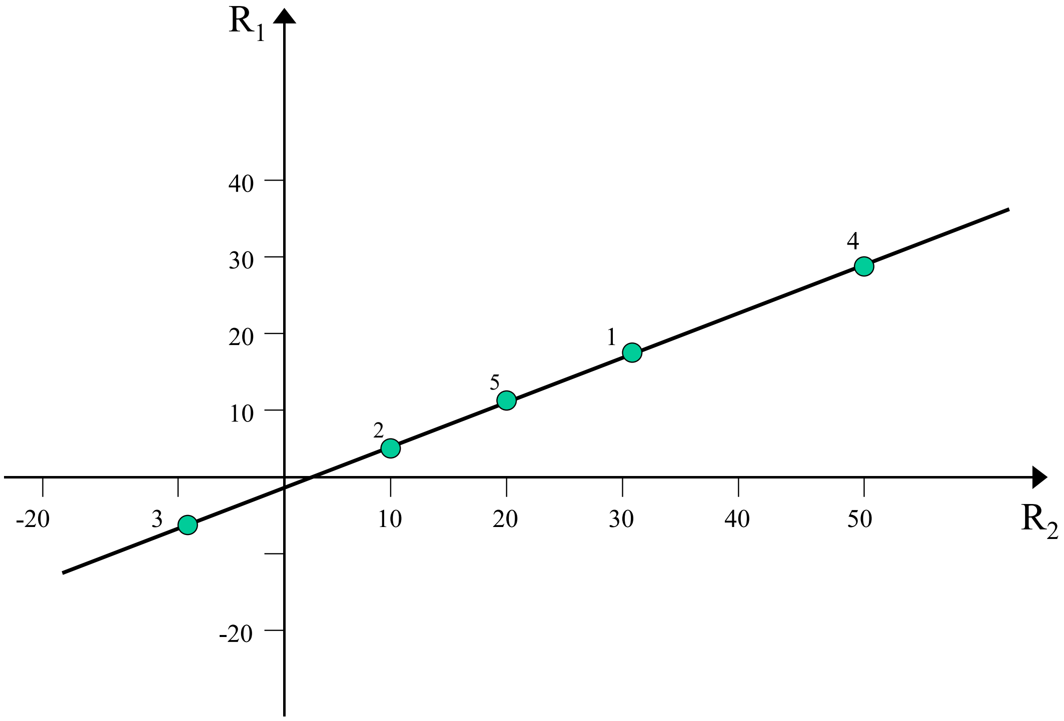
(a) Vollständig positive Korrelation



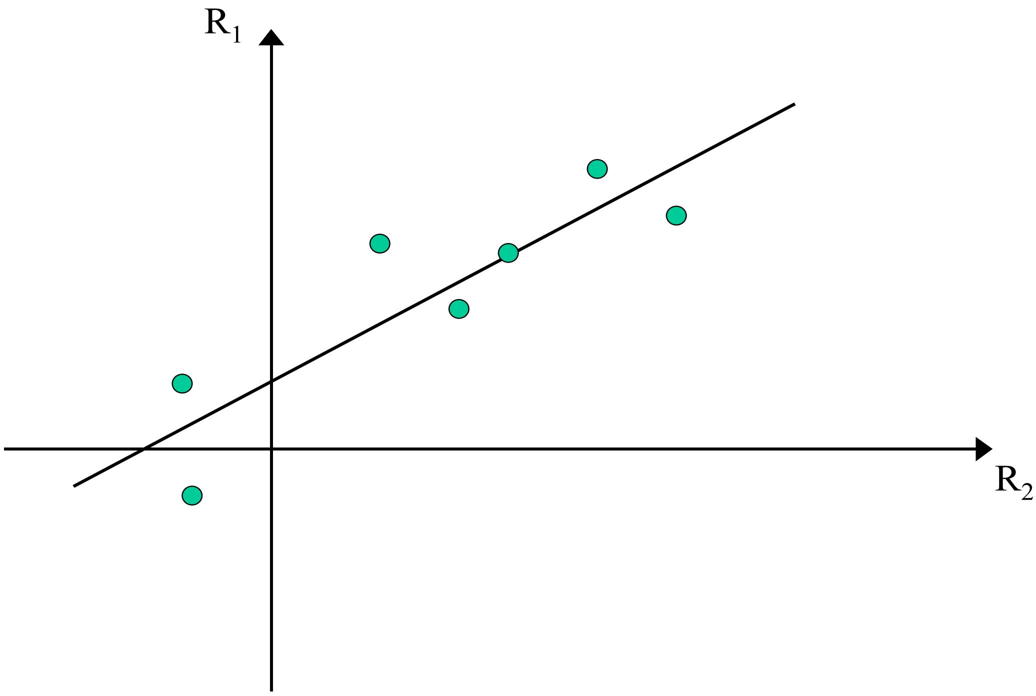
R_1	15	5	-5	25	10
R_2	30	10	-10	50	20

Cov (R_1R_2) =

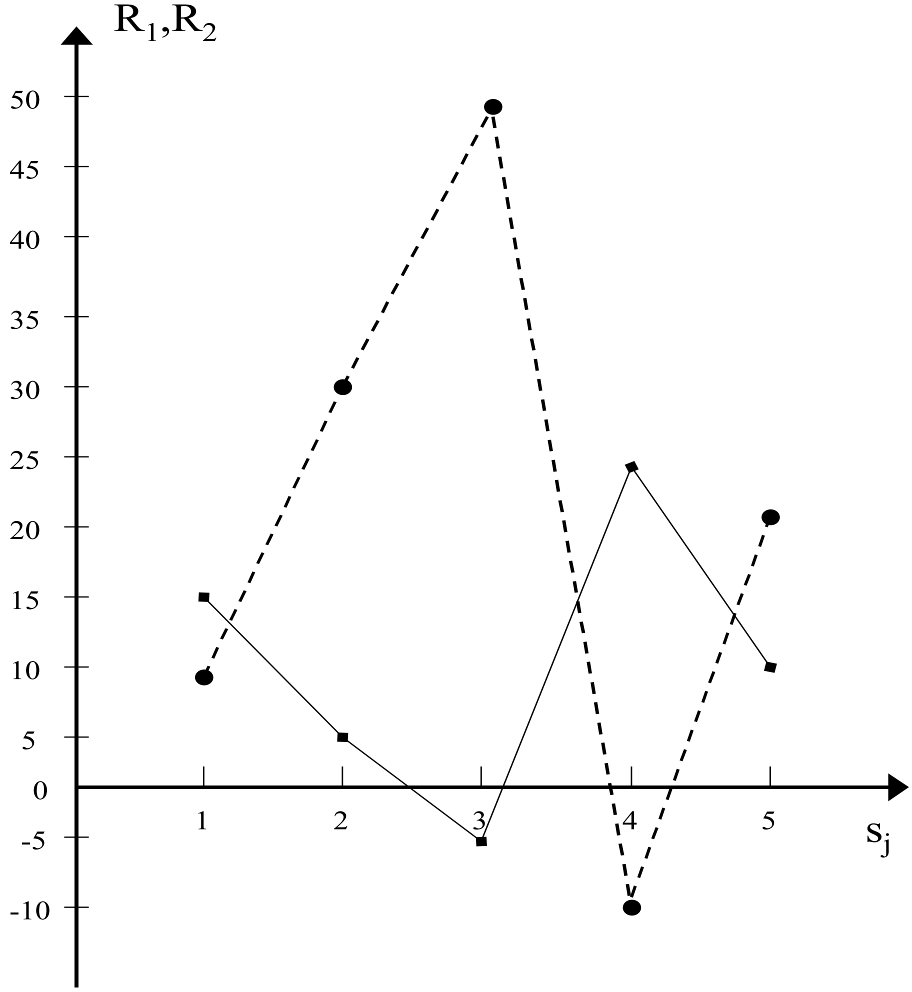
ρ ($R_{1,}R_2$) =



Nicht-perfekte positive Korrelation: $0 < \rho < 1$



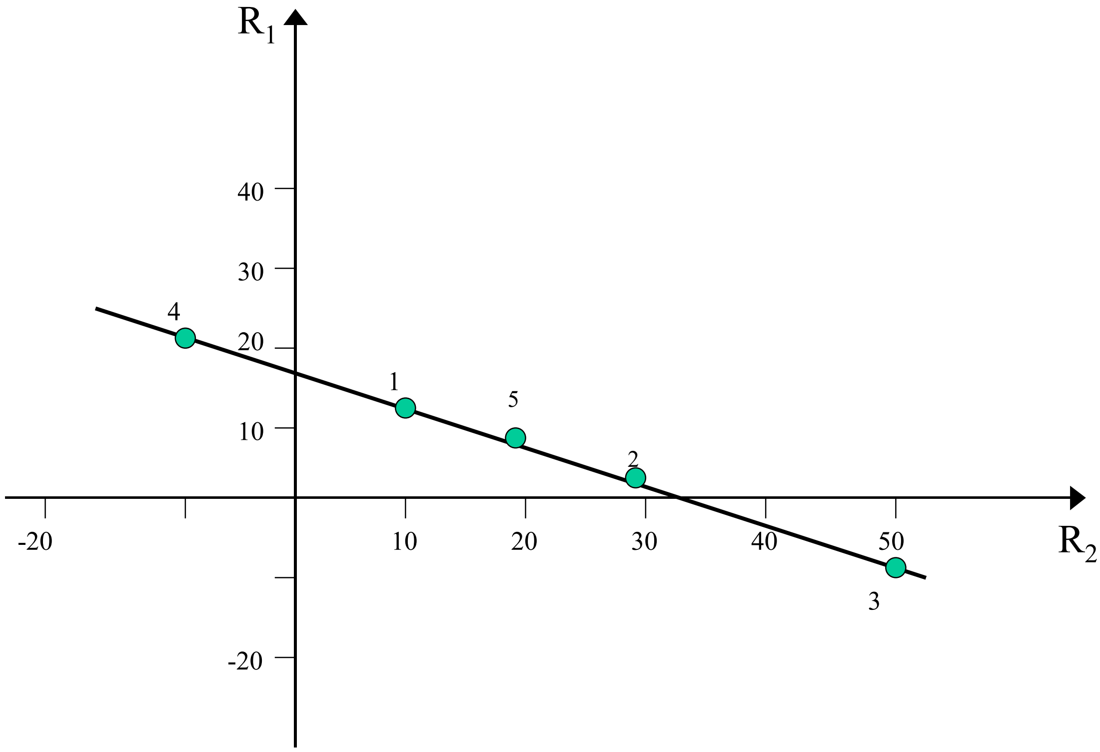
(b) Vollständig negative Korrelation



R_1	15	5	-5	25	10
R_2	10	30	50	-10	20

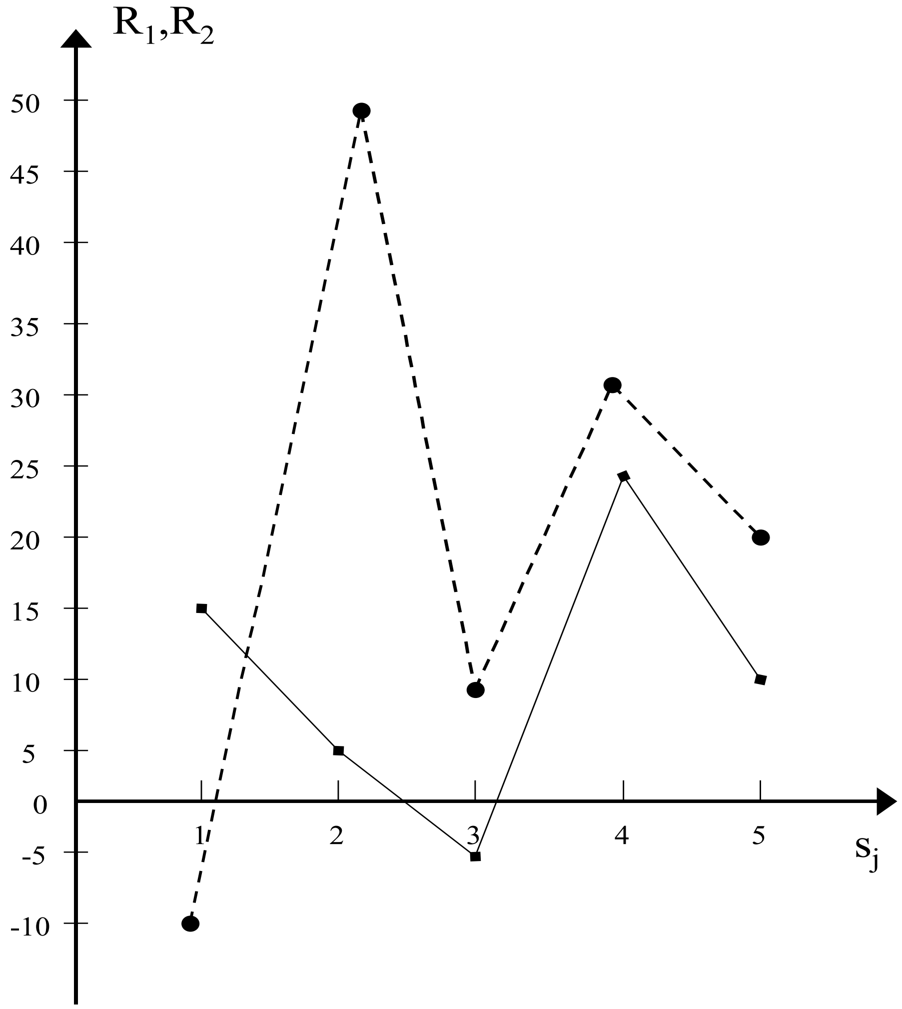
Cov (R_1R_2) =

ρ (R_1,R_2) =



R1
40
30
20
10
-20
-20
10
20
30
40
50
R2
4
1
5
2
3

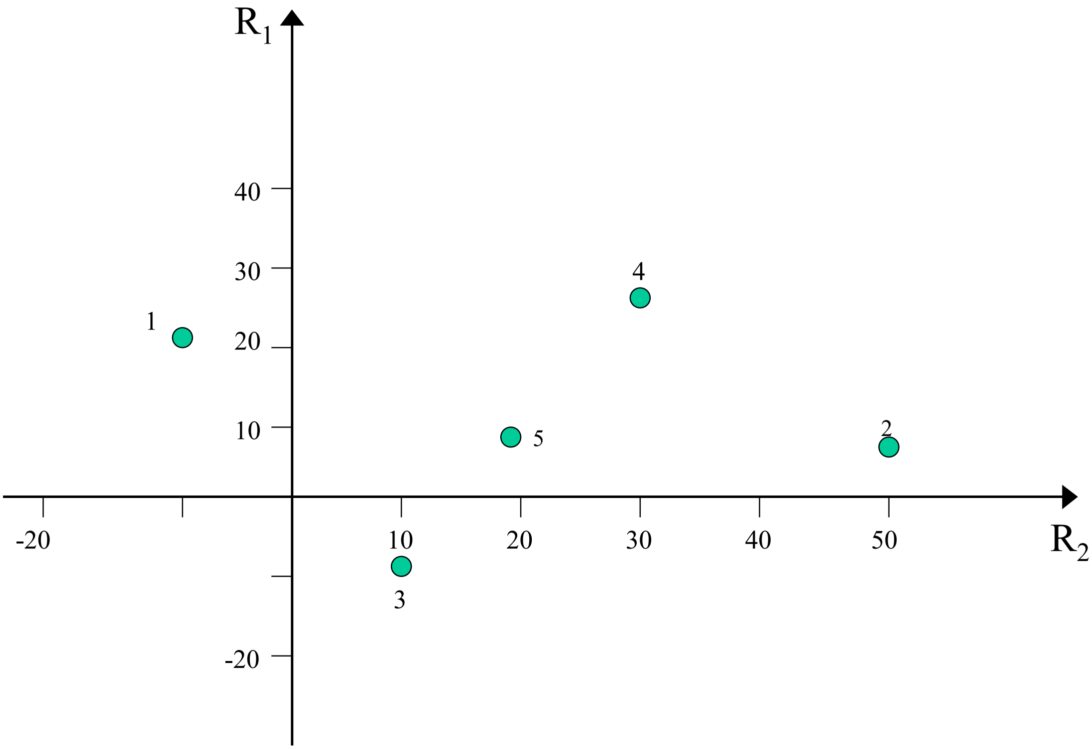
(c) Keine Korrelation



R_1	15	5	-5	25	10
R_2	-10	50	10	30	20

Cov (R_1R_2) =

ρ (R_1,R_2) =



(d) Nur Indikator für lineare Abhängigkeit

(4) Erwartungswerte und Varianzen von Linearkombinationen zweier Variablen

Z	**E(Z)**	**Var(Z)**
$aX \pm bY$	$aE(X) \pm bE(Y)$	$a^2Var\ (X) + b^2Var\ (Y) \pm 2abCov\ (X,Y)$
$X + Y$ $(a = 1, b = 1)$	$E(X) + E(Y)$	$Var\ (X) + Var\ (Y) + 2Cov(X,Y)$
$X - Y$ $(a = 1, b = -1)$	$E(X) - E(Y)$	$Var\ (X) + Var - 2Cov(X,Y)$
$0{,}5(X + Y)$ $(a = 0{,}5, b = 0{,}5)$	$0{,}5[E(X) + E(Y)]$	$0{,}25Var\ (X) + 0{,}25Var\ (Y) + 0{,}5Cov(X,Y)$

Quelle: Bleymüller u. a. (1985), S. 49