Inhaltsverzeichnis

1 Wahrscheinlichkeitsrechnung **4**

1.1 Grundlagen 5

1.2 Rechnen mit Wahrscheinlichkeiten 5

1.2.1 Laplace Wahrscheinlichkeit 5

1.2.2 Rechenregeln, bedingte Wahrscheinlichkeiten und der Satz von BAYES . 10

1.2.3 Das Urnenmodell und Baumdiagramme 14

2 Zufallsvariablen und Wahrscheinlichkeitsverteilungen **19**

2.1 Die Begriffe Zufallsvariable und Verteilung 19

2.2 Diskrete Verteilungen 20

2.2.1 Wahrscheinlichkeitsfunktion und Verteilungsfunktion 20

2.2.2 Erwartungswert und Varianz 25

2.2.2.1 Der Erwartungswert 25

2.2.2.2 Varianz, Standardabweichung und Kovarianz 26

2.2.3 Die Binomialverteilung 30

2.2.4 Die Hypergeometrische Verteilung 35

2.2.5 Die Poisson-Verteilung 39

2.3 Stetige Verteilungen 43

2.3.1 Dichtefunktion und Verteilungsfunktion 43

2.3.2 Erwartungswert, Varianz und Standardabweichung 47

2.3.2.1 Der Erwartungswert 47

2.3.2.2 Die Varianz 48

2.3.3 Die Gleichverteilung 49

2.3.4 Die Dreiecksverteilung 51

2.3.5 Die Exponentialverteilung 52

2.3.6 Die Normalverteilung 55

3 Risikomaße **61**

3.1 Quantile und der Value at Risk 61

3.1.1 Quantile 61

3.1.2 Value at Risk 64

3.2 Bedingte Erwartungswerte und Expected Shortfall 68

3.2.1 Bedingter Erwartungswert 68

3.2.2 Expected Shortfall und Tail Value at Risk 69

4 Methoden der Risikoaggregation **77**

4.1 Der Varianz-Kovarianz-Ansatz 77

4.2 Monte-Carlo-Simulation 80
4.2.1 Integration mit Monte-Carlo-Simulation 81
4.2.2 Risikoaggregation mit Monte-Carlo-Simulation 82

A Anhang **88**

1 Wahrscheinlichkeitsrechnung

Motivation

Wahrscheinlichkeit ist ein in der Alltagssprache inflationär gebrauchter Begriff. Oft ist die Wahrscheinlichkeit für das Eintreten eines Ereignisses intuitiv klar. Zum Beispiel muss nicht weiter erklärt werden, dass das Würfeln eines „Einsers“ mit einem fairen Würfel, eine Eintrittswahrscheinlichkeit von $1/6$ aufweist. Anders verhält es sich bereits bei folgendem bekannten Beispiel:
Monty-Hall Dilemma, oder das Ziegenproblem: Bei einer TV-Spielshow kann ein Teilnehmer, oder eine Teilnehmerin zwischen drei Türen auswählen. Hinter einer Tür verbirgt sich der Gewinn, hinter den beiden anderen Türen eine Ziege[1]. Teilnehmer und Teilnehmerinnen wissen zu Beginn des Spiels nicht wo sich der Gewinn, oder die Ziegen befinden, nur dem Moderator ist die Lage des Gewinns bekannt. Der Moderator bittet zuerst darum, sich auf eine Türe festzulegen. Danach wird von den verbleibenden beiden Türen eine geöffnet, hinter der sich eine Ziege befindet. Nun hat die Teilnehmerin, der Teilnehmer, die Möglichkeit bei ihrer ursprünglichen Wahl zu bleiben, oder sich für die noch verbleibende, verschlossene Türe zu entscheiden. In der Spielshow[2] kann nun beobachtet werden, dass die Teilnehmerinnen und Teilnehmer mit sich selbst ringen, wie sie wählen sollen. Diese Entscheidungsfindung ist es auch, welche der Show ihre Spannung verleiht. Aus mathematischer Sicht wäre die Wahl einfach. Bleibt man bei der ursprünglichen Wahl, so ist die Gewinnwahrscheinlichkeit $1/3$, wie zu erwarten war. Entscheidet man sich hingegen für einen Wechsel, so beträgt die Gewinnwahrscheinlichkeit $2/3$. Dieses Ergebnis, wie es sehr oft beim Bestimmen von Wahrscheinlichkeiten der Fall ist, *widerspricht unserer Intuition.* Die erstveröffentlichte Lösung des Ziegenproblems hat viel Aufsehen erregt. Die Kolumnistin Marylin vos Savant erhielt als Reaktion auf ihre Veröffentlichung an die zehntausend Briefe von Lesern und Leserinnen, die ihre Lösung anzweifelten[3].
Die Lösung des Ziegenproblems wird im Abschnitt *Bedingte Wahrscheinlichkeit* erarbeitet.

Studierhinweise

Grundlegende Begriffe der Wahrscheinlichkeitsrechnung, sowie die mathematische Definition der Wahrscheinlichkeit selbst, sollen erlernt werden. Wahrscheinlichkeiten für unterschiedliche *Zufallsexperimente* werden berechnet und eine „neue Art“ Probleme zu verstehen, nämlich in wahrscheinlichkeitstheoretischem Sinne, soll erarbeitet werden.

[1] Die Ziege wird in diesem Fall nicht als Gewinn gesehen.
[2] Let's make a Deal, mit Monty Hall.
[3] https://de.wikipedia.org/wiki/Marilyn_vos_Savant

1.1 Grundlagen

Der Zufall unterliegt Bedingungen, welche im Rahmen der Wahrscheinlichkeitsrechnung untersucht werden. Die Vorgänge, welche besagten Bedingungen unterliegen, werden *Zufallsexperimente* genannt. Beispiele für Zufallsexperimente sind:

- Die Anzahl der fehlerhaften Produkte einer Tagesproduktion.
- Die Augenzahl beim Werfen eines Würfels.
- Die Wartezeit in der Hotline eines Mobilfunkanbieters.
- usw.

Die Folgenden Bedingungen müssen von einem Experiment erfüllt werden, damit es als Zufallsexperiment bezeichnet werden kann:

- Beliebig häufige Wiederholbarkeit unter den gleichen Bedingungen.
- Mehrere, sich ausschließende Ereignisse sind möglich.
- Das Ergebnis ist zufällig, das heißt es ist nicht eindeutig vorhersagbar.

Ein bestimmtes Ergebnis eines Zufallsexperimentes wird *Elementarereignis* ω_i genannt. Die Gesamtheit aller möglichen Elementarereignisse wird als *Ereignisraum* $\Omega = \{\omega_1, ..., \omega_m\}$ bezeichnet. Jede Teilmenge des Ereignisraumes, das heißt jeder mögliche Versuchsausgang, wird ein *Ereignis* genannt[4].

Beispiel 1.1.1 *Ereignisse und Elementarereignisse.*

Betrachten wir das einmalige Werfen eines fairen Würfels, so können wir sechs Elementarereignisse unterscheiden und zwar das Werfen der Augenzahlen 1, 2, 3, 4, 5, oder 6.
Neben den Elementarereignissen gibt es noch weitere mögliche Ereignisse, wie zum Beispiel das Werfen einer geraden Zahl das der Teilmenge $G = \{2, 4, 6\}$ entsprechen würde.

Übung 1.1.2 *Würfeln.*

Betrachten Sie das zweimalige Werfen eines fairen Würfels. Listen Sie die Elementarereignisse auf und formulieren Sie zwei weitere Ereignisse.

1.2 Rechnen mit Wahrscheinlichkeiten

1.2.1 Laplace Wahrscheinlichkeit

Ein Zufallsexperiment mit endlicher Anzahl an Elementarereignissen $\Omega = \{\omega_1, \ \omega_2, \ \ldots, \ \omega_i\}$, das beliebig oft wiederholt werden kann und bei dem alle Elementarereignisse gleich wahrscheinlich auftreten können, wird *Laplace Experiment* genannt. Einfachste Beispiele dafür sind eine *faire Münze*, oder ein *fairer Würfel*.
Für die Wahrscheinlichkeit[5] m verschiedener, gleich wahrscheinlicher Elementarereignisse gilt:

[4] Üblicherweise werden Ereignisse mit Großbuchstaben bezeichnet.
[5] Als Symbol für die Wahrscheinlichkeit wird in Anlehnung an das englische Wort *Probability* der Buchstabe P verwendet.

Formel 1.2.1 *Wahrscheinlichkeit für Elementarereignisse.*

$$P(\omega_i) = \frac{1}{m} \tag{1.1}$$

Mit $|M|$ bezeichnen wir allgemein die Anzahl der Elemente einer Menge M und nennen das deren Mächtigkeit. Ein beliebiges Ereignis E, das aus einem oder mehreren Elementarereignissen zusammengesetzt ist, kann als Teilmenge des Ereignisraums Ω betrachtet werden. $|E|$ ist also die Anzahl an günstigen Versuchsausgängen die zum Ereignis E führen, so gilt für die *Laplace Wahrscheinlichkeit* des Ereignisses E:

Formel 1.2.2 *Laplace Wahrscheinlichkeit.*

$$P(E) = \frac{|E|}{|\Omega|} = \frac{Guenstige}{Moegliche} \tag{1.2}$$

Die Wahrscheinlichkeit für den gesamten Ereignisraum ist immer 100 %, weil „irgendetwas das möglich ist, passiert mit *Sicherheit.*“, bzw. wenn die Anzahl der Günstigen gleich der Anzahl der Möglichen ist, beträgt ihr Quotient 1.

Formel 1.2.3 *Wahrscheinlichkeit des Ereignisraums.*

$$P(\Omega) = 1 \tag{1.3}$$

Andererseits ist die Wahrscheinlichkeit für ein *unmögliches Ereignis*[6], ein Ereignis dass sich nicht aus den Elementarereignissen des Ereignisraumes zusammensetzen lässt, immer 0.

Formel 1.2.4 *Unmögliches Ereignis.*

$$P(\omega_k) = 0 \quad wenn \quad \omega_k \notin \Omega \tag{1.4}$$

Zu jedem Ereignis E existiert stets ein sog. *Gegenereignis* $\neg E$ das als Teilmenge des Ereignisraums betrachtet alle Elemente aus Ω enthält die E nicht enthält. Für die sog. *Gegenwahrscheinlichkeit* gilt:

Formel 1.2.5 *Gegenwahrscheinlichkeit*

$$P(\neg E) = 1 - P(E) \tag{1.5}$$

[6] Man kann zum Beispiel mit einem Würfel keine 7 würfeln.

Wir wollen die Formeln für die Wahrscheinlichkeit nun in einigen Beispielen zur Anwendung bringen.

Beispiel 1.2.6 *Einmaliges Würfeln.*

Die Wahrscheinlichkeit für jede einzelne Augenzahl beim Werfen eines fairen Würfels beträgt nach Laplace $\frac{1}{6}$. Wir fragen uns nun, wie groß die Wahrscheinlichkeit bei einem Wurf ist,

i) eine gerade Augenzahl

ii) eine Zahl ≥ 2

iii) 1, 2, oder 3 zu werfen.

Lösung:

i) Es gibt 3 günstige Fälle weil das Ereignis der Teilmenge $\{2, 4, 6\}$ entspricht. Für die Wahrscheinlichkeit des Ereignisses *gerade Augenzahl* ergibt sich daher:

$$P(\text{gerade Zahl}) = \frac{3}{6} = \frac{1}{2} = 50\,\%$$

ii) Auf einem Würfel gibt es 5 Zahlen die größer oder gleich 2 sind, das ist also die Anzahl der Günstigen für dieses Ereignis. Die Anzahl der möglichen ist beim Würfeln stets 6. Wir berechnen:

$$P(\text{mindestens } 2) = \frac{5}{6} \approx 83.3\,\%$$

iii) 1, 2, oder 3 zu werfen, das sind drei Günstige. Die Wahrscheinlichkeit ist also:

$$P(\text{höchstens } 3) = \frac{3}{6} = 50\,\%$$

Beispiel 1.2.7 *Zwei Würfel.*

Zwei faire Würfel werden geworfen. Wir wollen berechnen wie groß die Wahrscheinlichkeit ist,

i) eine 4 und eine 5,

ii) mindestens einmal die Augenzahl 5,

iii) eine Augensumme von 7 zu werfen

Lösung:

i) Wir können entweder $(4;5)$, oder $(5;4)$ würfeln um das gewünschte Ereignis zu erhalten. Die Anzahl der Günstigen ist also 2. Insgesamt gibt es sechs Möglichkeiten durch den ersten Würfel **und** weitere sechs Möglichkeiten durch den zweiten Würfel[7]. Für jede Möglichkeit des ersten Würfels gibt es sechs Möglichkeiten des zweiten Würfels. Die Wahrscheinlichkeit ergibt:

$$P(\text{1x4, 1x5}) = \frac{2}{36} = \frac{1}{18} \approx 5.6\,\%$$

ii) Um die Anzahl an Günstigen zu bestimmen wollen wir zunächst alle möglichen Augenzahlen der zwei Würfel betrachten, siehe Abbildung **??**.

11	21	31	41	51	61
12	22	32	42	52	62
13	23	33	43	53	63
14	24	34	44	54	64
15	25	35	45	55	65
16	26	36	46	56	66

Abbildung 1.1: Augenzahlen zweier Würfel, die 5 enthalten

Die Anzahl der Günstigen beträgt demnach 11 und für die Wahrscheinlichkeit erhalten wir:

$$P(\text{mindestens 1x5}) = \frac{11}{36} \approx 30.6\,\%$$

iii) Wir betrachten wieder die möglichen Augenzahlen zweier Würfel und lesen die Anzahl der Günstigen ab, siehe Abbildung **??**.

11	21	31	41	51	61
12	22	32	42	52	62
13	23	33	43	53	63
14	24	34	44	54	64
15	25	35	45	55	65
16	26	36	46	56	66

Abbildung 1.2: Augenzahlen zweier Würfel, deren Summe 7 ergibt.

Die Anzahl der Günstigen, deren Augensumme 7 ergibt ist 6, siehe Abbildung **??**, und wir erhalten wiederum die Wahrscheinlichkeit:

$$P(\text{Augensumme 7}) = \frac{6}{36} \approx 16.7\,\%$$

[7] Eine logische *und* Verknüpfung von sich ausschließenden Ereignissen führt zu einer Multiplikation der Möglichkeiten.

Beispiel 1.2.8 *Ampelschaltung.*

Laut § 38 Abs. 6 Satz 1 der Straßenverkehrsordnung (StVO) gilt:
„Das grüne Licht ist jeweils mit viermal grün blinkendem Licht zu beenden, wobei die Leucht- und die Dunkelphase abwechselnd je eine halbe Sekunde zu betragen haben.“ Tabelle **??** zeigt die Dauer der einzelnen Ampelphasen an.

Tabelle 1.1: Ampelphasen

Ampelphase	Dauer in s
Rot	30
Gelb	3
Grün	20
Grün blinkend	4

i) Wie groß ist die Wahrscheinlichkeit, die Ampel in einer *Gelbphase* anzutreffen?

ii) Wie groß ist die Wahrscheinlichkeit, die Ampel nicht bei Rot zu erreichen?

Lösung: Wir erkennen durch Addition, dass ein gesamter Ampelzyklus 57 s dauert. Die Anzahl an Sekunden in einem gesamten Zyklus wird unsere Anzahl an Möglichen sein.

i) Eine Gelbphase dauert 3 s, es gibt also je Zyklus 3 günstige Sekunden für unser Ereignis. Die gesuchte Wahrscheinlichkeit beträgt demnach:

$$P(\text{gelb}) = \frac{3}{57} = \frac{1}{19} \approx 5.3\,\%$$

ii) Wir könnten nun alle nicht roten Ampelphasen zusammenfassen und wie gewohnt fortfahren. Es bietet sich aber an mit der sogenannten *Gegenwahrscheinlichkeit* zu arbeiten. Wir bestimmen die Wahrscheinlichkeit des logischen Gegenteils, des von uns gewünschten Ereignisses[8] und ziehen diese von 1 ab.

$$P(\text{nicht rot}) = 1 - P(\text{rot}) = 1 - \frac{30}{57} = \frac{27}{57} \approx 47.4\,\%$$

Übung 1.2.9 *Alphabet*
Ein Buchstabe aus dem Alphabet wird zufällig ausgewählt. Mit welcher Wahrscheinlichkeit handelt es sich um einen Vokal? Was ist die zu diesem Ereignis gehörende Teilmenge?

Übung 1.2.10 *Roulette*
Mit welcher Wahrscheinlichkeit gewinnen Sie wenn Sie beim Roulette auf ”rot“ setzen? Was wäre das Gegenereignis?

[8] Das logische Gegenteil von „nicht Rot“ ist „Rot“

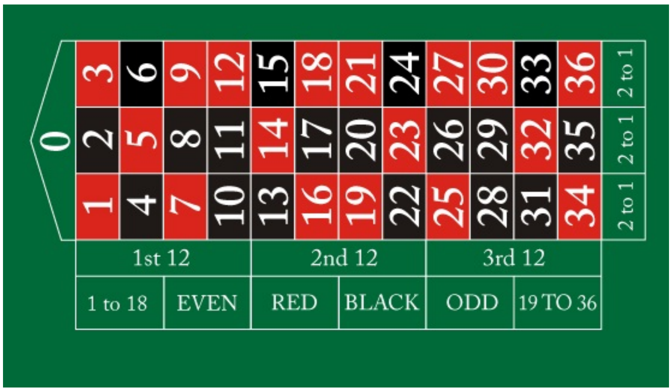


Abbildung 1.3: Spielfeld beim Roulette

1.2.2 Rechenregeln, bedingte Wahrscheinlichkeiten und der Satz von BAYES

Im Folgenden geht es um die Verknüpfungen von Ereignissen. Dabei werden wir folgende Schreibweisen zur Verknüpfung zweier Ereignisse benutzen:

- $P(A \vee B)$ bezeichnet die Wahrscheinlichkeit, dass mindestens eines der beiden Ereignisse A oder B eintritt, man spricht *A oder B* (nicht zu verwechseln mit "entweder oder").

- $P(A \wedge B)$ bezeichnet die Wahrscheinlichkeit, dass sowohl Ereignis A als auch Ereignis B zugleich eintritt. Man spricht *A und B*.

Bsp.: Beim Würfeln bezeichne
A das Ereignis, dass eine gerade Zahl geworfen wird und
B das Ereignis, dass mindestens eine 4 geworfen wird. Dann ist also:

$P(A \vee B) = \frac{4}{6}$

$P(A \wedge B) = \frac{2}{6}$

Übung 1.2.11 *Roulette*
Sei *G: Gerade Zahl* und *R: Rote Zahl*. Berechnen Sie die Wahrscheinlichkeiten $P(G \vee R)$ und $P(\neg G \wedge \neg R)$. (Hinweis: Die "0" wird beim Roulette weder als gerade noch als ungerade betrachtet.)

10x Rot,Ungerade + 18 Rot = 28/37
Bei Gegenbeispiel beachten, dass Null nicht mitgezählt wird!

Die Rechenregeln für Wahrscheinlichkeiten verknüpfter Ereignisse lassen sich anhand von sog. Vierfeldertafeln herleiten.

Beispiel 1.2.12 *Vierfeldertafel 1*

Ein Gruppe von 100 Personen wird hinsichtlich des Geschlechts sowie ob eine Person raucht oder nicht betrachtet:

Tabelle 1.2: RaucherInnen 1

	männlich (M)	weiblich (W)	Gesamt
Raucher (R)	40	10	50
Nichtraucher (N)	30	20	50
Gesamt	70	30	Σ100

Wird nun eine Person zufällig ausgewählt, so erhält man anhand der Formel für die Laplace-Wahrscheinlichkeit folgende Wahrscheinlichkeiten:

$$P(R) = \frac{50}{100} = \frac{1}{2} = 50\%$$
$$P(M) = \frac{70}{100} = \frac{7}{10} = 70\%$$
$$P(M \wedge R) = \frac{40}{100} = \frac{2}{5} = 40\%$$

Möchte man mit Hilfe dieser Wahrscheinlichkeiten nun die Wahrscheinlichkeit dafür berechnen, dass eine zufällig ausgewählte Person männlich ist *oder* raucht, so muss man beachten, dass die 40 männlichen Raucher sowohl in den 50 Rauchern sowie in den 70 Männern enthalten sind. Diese dürfen natürlich nicht doppelt gezählt werden und man erhält:

$$P(M \vee R) = P(M) + P(R) - P(R \wedge M) = \frac{50}{100} + \frac{70}{100} - \frac{40}{100} = \frac{80}{100} = 80\%$$

Formel 1.2.13 *Die ODER-Verknüpfung*

$$P(A \vee B) = P(A) + P(B) - P(A \wedge B) \tag{1.6}$$

Gilt $P(A \wedge B) = 0$ *(wenn es also z.B. keine rauchenden Männer gibt), so nennt man die Ereignisse unvereinbar und es gilt:*

$$P(A \vee B) = P(A) + P(B) \tag{1.7}$$

Nun gehen wir davon aus, dass eine Person zufällig ausgewählt wurde und man weiß es handelt sich um eine Frau. Es soll ermittelt werden wie groß unter dieser Bedingung die Wahrscheinlichkeit ist, dass es sich um eine Raucherin handelt. Als Grundgesamtheit werden nun also nur die 30 Frauen betrachtet. Für diese sogenannte *bedingte Wahrscheinlichkeit* $P(R|W)$ ("R unter der Bedingung B“) gilt dann:

$$P(R|W) = \frac{10}{30} = \frac{\frac{10}{100}}{\frac{30}{100}} = \frac{P(R \wedge W)}{P(W)} \quad \Leftrightarrow \quad P(R \wedge W) = P(R|W) \cdot P(W)$$

Im betrachteten Beispiel ist der Anteil der RaucherInnen bei den Männern höher als bei den Frauen. Im folgenden Beispiel ist der Anteil identisch.

Beispiel 1.2.14 *Vierfeldertafel 2*

Tabelle 1.3: RaucherInnen 2

	männlich (M)	weiblich (W)	Gesamt
Raucher (R)	10	30	40
Nichtraucher (N)	20	60	80
Gesamt	30	90	Σ120

In diesem Fall gilt

$$\begin{aligned} P(R|W) &= \frac{30}{90} = \frac{1}{3} \\ P(R|M) &= \frac{10}{30} = \frac{1}{3} \\ P(R) &= \frac{40}{120} = \frac{1}{3} \end{aligned}$$

Die Wahrscheinlichkeit ob eine Person raucht oder nicht ist also *unabhängig* vom Geschlecht und es gilt $P(R|W) = P(R|M) = P(R)$.

Formel 1.2.15 *Die UND-Verknüpfung*

$$P(A \wedge B) = P(A|B) \cdot P(B) \quad \textit{(allgemeiner Fall)} \tag{1.8}$$

bzw.

$$P(A \wedge B) = P(A) \cdot P(B) \quad \textit{(wenn A und B unabhängig)} \tag{1.9}$$

Weil außerdem gilt $P(A \wedge B) = P(B \wedge A)$ folgt:

$$P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

und damit:

Formel 1.2.16 *Satz von BAYES*

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} = \frac{P(A \wedge B)}{P(B)} \tag{1.10}$$

Übung 1.2.17 *Glühbirnen*
In einer Kiste befinden sich 150 Glühbirnen mit 25W (90 Stück) bzw. 40W Leistung. 65 Glühbirnen haben eine Fassung aus Milchglas, die restlichen sind durchsichtig. 15 der durchsichtigen Glühbirnen haben einen Leistung von 40W.
i) Stellen sie diese Situation mit einer Vierfeldertafel dar.
ii) Mit welcher Wahrscheinlichkeit ist eine zufällig ausgewählte Glühbirne durchsichtig?
iii) Mit welcher Wahrscheinlichkeit wird eine 25W-Glühbirne aus Milchglas gezogen?
iv) Mit welcher Wahrscheinlichkeit hat eine durchsichtige Glühbirne eine Leistung von 25W?

Übung 1.2.18 *Rohlinge.*

Für eine Fräsmaschine werden 18 Rohlinge (das sind Werkstücke, die noch weiterbearbeitet werden müssen) geliefert. Die Rohlinge werden in 3 Behältern geliefert. Im ersten Behälter befinden sich 6 Rohlinge, im zweiten Behälter 5 Rohlinge und im dritten Behälter 7 Rohlinge. Aufgrund von Transportschädigung befindet sich in jedem Behälter je 1 defekter Rohling. Jedem Behälter wird genau 1 Rohling entnommen. Von diesen 3 Rohlingen ist keiner defekt.
Berechnen Sie die Wahrscheinlichkeit für dieses Ereignis.

Übung 1.2.19 *Ölbohrung.*

Eine Ölgesellschaft führt Probebohrungen in Texas und in Alaska durch. Erfahrungsgemäß findet man bei einer Bohrung in Texas mit einer Wahrscheinlichkeit von 85 % und in Alaska mit einer Wahrscheinlichkeit von 65 % Öl. Berechnen Sie die Wahrscheinlichkeit, dass nur in Texas oder nur in Alaska Öl (also genau 1 mal) gefunden wird, wenn die beiden Bohrungen unabhängig voneinander sind.

Übung 1.2.20 *Alarmanlage.*

Ein bestimmtes Alarmanlagensystem löst jeweils mit der Wahrscheinlichkeit 0,85 im Einbruchsfall Alarm aus. Eine Familie lässt zwei dieser Anlagen in ihr Haus so einbauen, dass sie unabhängig voneinander Alarm auslösen. Berechnen Sie die Wahrscheinlichkeit, dass im Einbruchsfall mindestens eine der beiden Anlagen Alarm auslöst.

1.2.3 Das Urnenmodell und Baumdiagramme

Viele Experimente lassen sich mit dem dem sog. *Urnenmodell* veranschaulichen. Dabei werden mehrere Kugeln verschiedener Farben hintereinander aus einer Urne gezogen. Wird jede Kugel nach dem sie gezogen wurde wieder zurückgelegt, so sind die einzelnen Züge voneinander unabhängig. Andernfalls hängen die Wahrscheinlichkeiten der einzelnen Züge voneinander ab und wir erhalten bedingte Wahrscheinlichkeiten. Wir beschäftigen uns also mit sog. *mehrstufigen* Experimenten bei denen verschiedene Experimente oder das selbe Experiment mehrfach hintereinander ausgeführt werden. Eine sehr gute grafische Veranschaulichung stellen die sog. *Baumdiagramme* dar. Beides werden wir anhand von Beispielen genauer betrachten.

Beispiel 1.2.21 *Ziehen mit Zurücklegen*

Wir stellen uns eine Urne mit 7 roten und 3 blauen Kugeln vor. Es wird zwei Mal eine Kugel gezogen, wobei jede Kugel vor dem nächsten Zug wieder zurückgelegt wird. An einem Baumdiagramm veranschaulicht sieht das folgendermaßen aus:

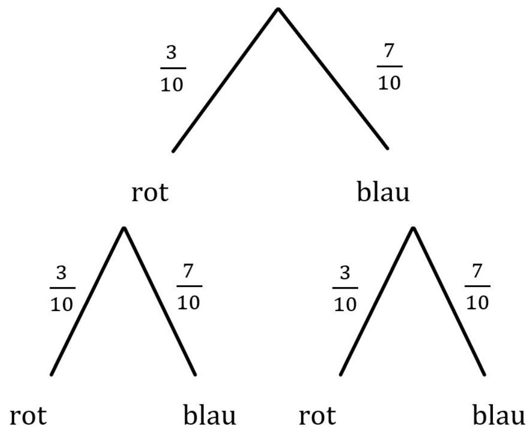


Abbildung 1.4: Ziehen mit Zuruecklegen

Wir bezeichnen einen Weg von ganz oben nach ganz unten als *Pfad*, dabei repräsentiert jeder Pfad ein gewisses Ereignis. Da die einzelnen Züge voneinander unabhängig sind dürfen wir die Wahrscheinlichkeiten entlang eines Pfades multiplizieren. Die einzelnen Pfade sind wiederum unvereinbar, somit dürfen die Wahrscheinlichkeiten die sich aus verschiedenen Pfaden ergeben zueinander addiert werden.

Bsp.:
A: Es werden zwei rote Kugeln gezogen
B: Es wird zuerst eine rote und dann eine blaue Kugel gezogen
C: Es werden zwei Kugeln unterschiedlicher Farbe gezogen

$$P(A) = \frac{3}{10} \cdot \frac{3}{10} = \left(\frac{3}{10}\right)^2 = \frac{9}{100} = 9\%$$
$$P(B) = \frac{3}{10} \cdot \frac{7}{10} = \frac{21}{100} = 21\%$$
$$P(C) = \frac{3}{10} \cdot \frac{7}{10} + \frac{7}{10} \cdot \frac{3}{10} = \frac{42}{100} = 42\%$$

Bei den Ereignissen B und C ist zu beachten, dass bei B die Reihenfolge genannt ist in der die Kugeln gezogen werden sollen, B entspricht also nur der Pfad rot-rot. Bei C hingegen ist die Reihenfolge egal und somit entspricht C den beiden Pfade rot-blau und blau-rot.

Im nächsten Fall wird eine Kugel nach dem sie gezogen wurde nicht wieder zurückgelegt. Dadurch ändern sich die Wahrscheinlichkeiten in der 2. Stufe, je nachdem welche Farbe beim 1. Zug gezogen wurde. Das Baumdiagramm sieht dann folgendermaßen aus:

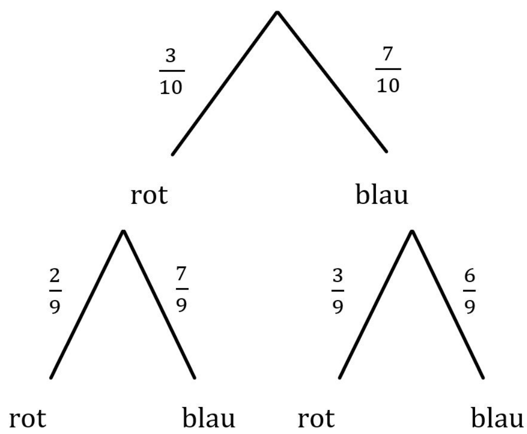


Abbildung 1.5: Ziehen ohne Zurücklegen

Die Züge sind jetzt zwar nicht mehr unabhängig voneinander, allerdings handelt es sich in der 2. Stufe um bedingte Wahrscheinlichkeiten: z.B. ist $P(2.\text{ Kugel blau}|1.\text{ Kugel rot}) = \frac{7}{9}$ oder $P(2.\text{ Kugel blau}|1.\text{ Kugel blau}) = \frac{6}{9}$. Die Wahrscheinlichkeiten entlang eines Pfades dürfen also wieder miteinander multipliziert werden. Außerdem sind die Ereignisse der einzelnen Pfade nach wie vor unvereinbar, sodass für die Wahrscheinlichkeiten der Ereignisse A, B und C des vorigen Beispiels gilt:

$$P(A) = \frac{3}{10} \cdot \frac{2}{9} = \frac{6}{90} = \frac{1}{15} = 6,7\%$$
$$P(B) = \frac{3}{10} \cdot \frac{7}{9} = \frac{21}{90} = \frac{7}{30} = 23,3\%$$

$$P(C) = \frac{3}{10} \cdot \frac{7}{9} + \frac{7}{10} \cdot \frac{3}{9} = \frac{42}{90} = \frac{7}{15} = 46,7\%$$

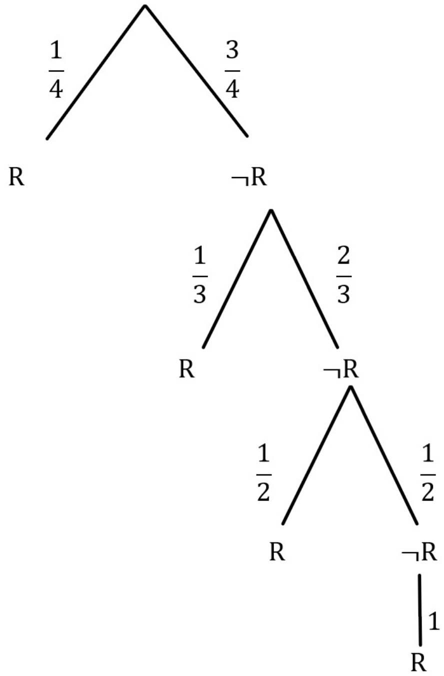
Pfadregeln: Zusammengefasst gelten also beim Ziehen sowohl mit als auch ohne Zurücklegen folgende sog. *Pfadregeln*:

- 1. Pfadregel (Multiplikationsregel): "Entlang eines Pfades werden die Wahrscheinlichkeiten multipliziert." (Merkregel: Und wird zu Mal)
- 2. Pfadregel (Additionsregel): "Mehrere Pfade werden ggf. addiert." (Merkregel: Oder wird zu Plus)

Beispiel 1.2.22 *Schlüsselbund*

Sie haben 4 Schlüssel an Ihrem Schlüsselbund und kommen nachts nach Hause. Weil das Licht im Stiegenhaus kaputt ist bleibt Ihnen nichts anderes übrig als einen Schlüssel nach dem anderen auszuprobieren. Mit welcher Wahrscheinlichkeit finden Sie den richtigen Schlüssel nach 1,...,4 Versuchen?

Lösung:
Es wird ohne Zurücklegen gezogen, sonst wäre es möglich, dass es nie zu einem Ende kommt (was theoretisch natürlich auch denkbar wäre). Dabei bezeichnet das Ereignis R das Ziehen des richtigen Schlüssels. Nur wenn der richtige Schlüssel nicht gezogen wird geht es weiter, ansonsten ist die Suche natürlich beendet. Spätestens nach vier Mal Ziehen muss der richtige Schlüssel gefunden worden sein.



Für die Wahrscheinlichkeiten ergibt sich (evtl. etwas überraschend):

$$P(1 \text{ Versuch}) = \frac{1}{4}$$
$$P(2 \text{ Versuche}) = \frac{3}{4} \cdot \frac{1}{3} = \frac{1}{4}$$
$$P(3 \text{ Versuche}) = \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{4}$$
$$P(4 \text{ Versuche}) = \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} \cdot 1 = \frac{1}{4}$$

Beispiel 1.2.23 *Alarmanlage*

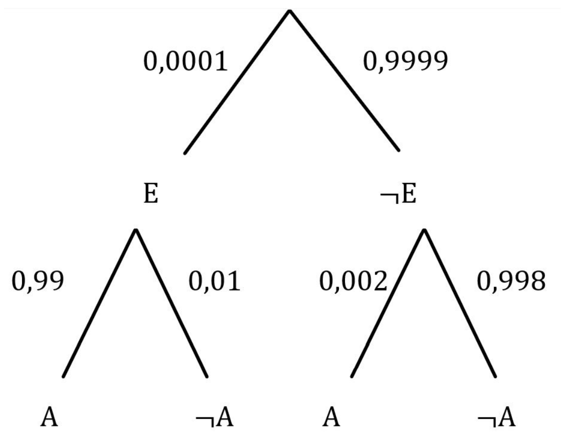
Die Wahrscheinlichkeit, dass in einer Nacht in ihr Auto eingebrochen wird beträgt 0,01%. Das Auto ist mit einer Alarmanlage ausgerüstet die im Einbruchsfall mit einer Wahrscheinlichkeit von 99% auslöst. Mit einer Wahrscheinlichkeit von 0,2% kommt es allerdings auch zu einem Fehlalarm.
Stellen sie den Sachverhalt mit einem Baumdiagramm dar und berechnen Sie folgende Wahrscheinlichkeiten:

i) Mit welcher Wahrscheinlichkeit löst die Alarmanlage Alarm aus?
ii) Sie hören Alarm. Mit welcher Wahrscheinlichkeit wurde tatsächlich eingebrochen?

Lösung:

Wir benennen folgende Ereignisse:

E: Einbruch
A: Alarm



$i)\quad P(A) = P(E) \cdot P(A|E) + P(\neg E) \cdot P(A|\neg E) = 0,0001 \cdot 0,99 + 0,9999 \cdot 0,002 = 0,0020988 \approx 0,21\%$

$$ii)\quad P(E|A) = \frac{P(E \wedge A)}{P(A)} = \frac{P(E) \cdot P(A|E)}{P(A)} = \frac{0,0001 \cdot 0,99}{0,0001 \cdot 0,99 + 0,9999 \cdot 0,002} \approx 4,72\%$$

Übung 1.2.24 *Prüfung*
Bei einer Prüfung muss eine Aufgabe gezogen werden. 6 von 10 Aufgaben sind Rechenaufgaben, der Rest Multiple-Choice-Aufgaben. Eine Rechenaufgabe wird mit einer Wahrscheinlichkeit von 75% gelöst, Multiple-Choice-Aufgaben werden mit 35% Wahrscheinlichkeit falsch beantwortet.
i) Mit welcher Wahrscheinlichkeit wird eine zufällig gezogene Aufgabe gelöst?
ii) Die Aufgabe wurde richtig gelöst. Mit welcher Wahrscheinlichkeit war es eine Multiple-Choice-Aufgabe?

Übung 1.2.25 *Sockenlade*
In einer Lade befinden sich 12 rote und 17 schwarze Socken.
i) Mit welcher Wahrscheinlichkeit hat man nach 2-maligen Ziehen zwei gleichfarbige Socken gezogen?
ii) Mit welcher Wahrscheinlichkeit bei 3-maligem Ziehen?
iii) Jemand hat 2 Socken zufällig gezogen und behauptet sie seien gleichfarbig. Mit welcher Wahrscheinlichkeit sind die Socken rot?

2 Zufallsvariablen und Wahrscheinlichkeitsverteilungen

Motivation

Zur Abschätzung von Risiken ist es notwendig mögliche Ereignisse in Ihrer Gesamtheit und hinsichtlich Ihrer Wahrscheinlichkeiten zu kennen und mathematisch beschreiben zu können. Neben der Kenntnis von Wahrscheinlichkeiten einzelner Ereignisse (mit welcher Wahrscheinlichkeit tritt ein Ereignis ein das zu einem gewissen Verlust führt?) gehört auch die Abschätzung langfristiger Gesamteffekte dazu (mit welchem Gewinn kann langfristig gerechnet werden?).
Die Gesamtheit aller möglichen Ereignisse eines Experiments und Ihrer Wahrscheinlichkeiten nennen wir eine *(Wahrscheinlichkeits-)Verteilung*. Auf dieser Grundlage werden dann die entsprechenden Kenngrößen und Prognosen berechnet.

Studierhinweise

In diesem Kapitel werden Sie den Unterschied zwischen *diskreten* und *stetigen* Zufallsvariablen und ihren Wahrscheinlichkeitsverteilungen kennenlernen. Die Begriffe *Erwartungswert* und *Standardabweichung* der Grundgesamtheit werden eingeführt und angewandt. Sie bilden die Grundlage für viele Methoden des Risikomanagements.

2.1 Die Begriffe Zufallsvariable und Verteilung

Eine Zufallsvariable ist eine Größe die verschiedene Zahlenwerte annehmen kann, worüber aber der Zufall entscheidet. Wir haben schon intuitiv Zufallsvariablen verwendet, so ist eine Zufallsvariable z.B. die Augenzahl beim einfachen Würfelwurf, die Augensumme beim Wurf zweier Würfel, oder aber auch der Inhalt einer Flasche die von einer Abfüllanlage abgefüllt wird da hier unvermeidbar zufällige Schwankungen auftreten. Natürlich gibt es auch im Bereich der Finanzwirtschaft Ereignisse und Entwicklungen die vom Zufall abhängen, wie z.B. die Entwicklung von Aktienkursen oder Absatzmärkten und den damit verbundenen Gewinnen und Verlusten.
Für Zufallsvariablen werden in der Regel Großbuchstaben verwendet, ihre Beschreibung geschieht in Textform.

Bsp.: Beim einfachen Würfeln benennen wir die Zufallsvariable mit *X: Augenzahl*

Mit Hilfe der Zufallsvariablen können wir Ereignisse knapp und übersichtlich formulieren. Für das Ereignis: *E: Würfeln einer 3* schreiben wir kurz $X = 3$

und für die zugehörige Wahrscheinlichkeit: $P(X = 3) = \frac{1}{6}$

Wir unterteilen Zufallsvariablen nach ihrer Art in diskrete- und stetige Zufallsvariablen bzw. Verteilungen.

- **Diskrete Zufallsvariable/Verteilung:** Wenn nur einzelne, unterscheidbare und abzählbare (diskrete) Zahlenwerte angenommen werden können - oft eine Teilmenge der natürlichen Zahlen - so spricht man von einer diskreten Zufallsvariable bzw. Verteilung. Dazu würde beispielsweise die Augenzahl beim Würfeln zählen.

- **Stetige Zufallsvariable/Verteilung:** Wenn (theoretisch) jeder[1] beliebige Wert in einem Intervall $[a, b] \in \mathbb{R}$ angenommen werde kann, so spricht man von einer stetigen Zufallsvariable bzw. Verteilung.. Hierfür wären der erwähnte Flascheninhalt oder die Rendite eines Aktienportfolios geeignete Beispiele.

2.2 Diskrete Verteilungen

2.2.1 Wahrscheinlichkeitsfunktion und Verteilungsfunktion

Wird jedem möglichen Ergebnis einer Zufallsvariable eine Wahrscheinlichkeit zugeordnet, so wird quasi die Gesamtwahrscheinlichkeit von $1 = 100\%$ auf die einzelnen Elementarereignisse *verteilt*, daher *Verteilung*.Die Summe aller Wahrscheinlichkeiten einer Verteilung muss also immer 1 ergeben.
Die Funktion f, die jedem Elementarereignis x_i eine Wahrscheinlichkeit $P(X = x_i)$ zuordnet nennt man *Wahrscheinlichkeitsfunktion* mit $f(x_i) = P(X = x_i)$.

Bei einer diskreten Zufallsvariable wird die Verteilung oft in Tabellenform dargestellt. Wir betrachten nochmal das Beispiel des einfachen Würfelwurfs.

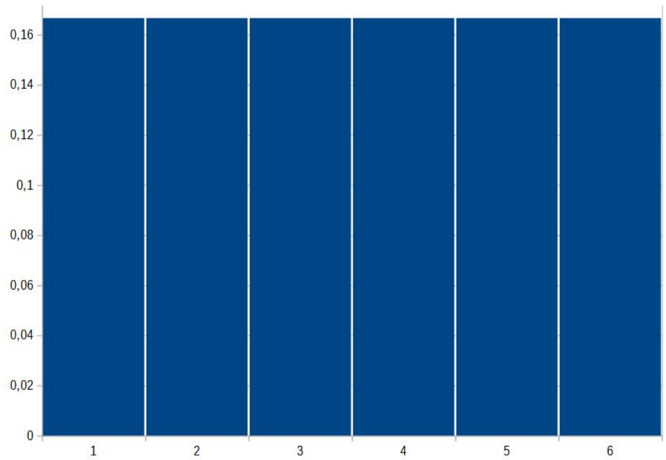
[1] In der Praxis ist man immer durch die Messgenauigkeit auf (nur) endlich viele Zahlen in einem Intervall eingeschränkt.

Beispiel 2.2.1 *Einmaliges Würfeln*

Die Wahrscheinlichkeitsverteilung beim einfachen Würfeln in Tabellenform sieht folgendermaßen aus:

x_i	1	2	3	4	5	6
$f(x_i) = P(X = x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Außerdem üblich ist die Darstellung als Säulendiagramm:



Beachte, dass die Breite eines jeden Balken gleich 1 ist, damit entspricht nicht nur die Höhe sondern auch die Fläche der einzelnen Balken der zugehörigen Wahrscheinlichkeit. Das wird bei der Betrachtung stetiger Verteilungen noch eine wichtige Rolle spielen.

Neben dem Begriff der Wahrscheinlichkeitsfunktion gibt es auch noch den Begriff der *Verteilungsfunktion* F mit $F(x_i) = P(X \leq x_i)$.
Die Verteilungsfunktion ordnet also jedem Wert x_i die Wahrscheinlichkeit zu, dass dieser Wert oder ein kleinerer herauskommt.

Damit ergibt sich folgende Wertetabelle für die Verteilungsfunktion beim einfachen Würfeln:

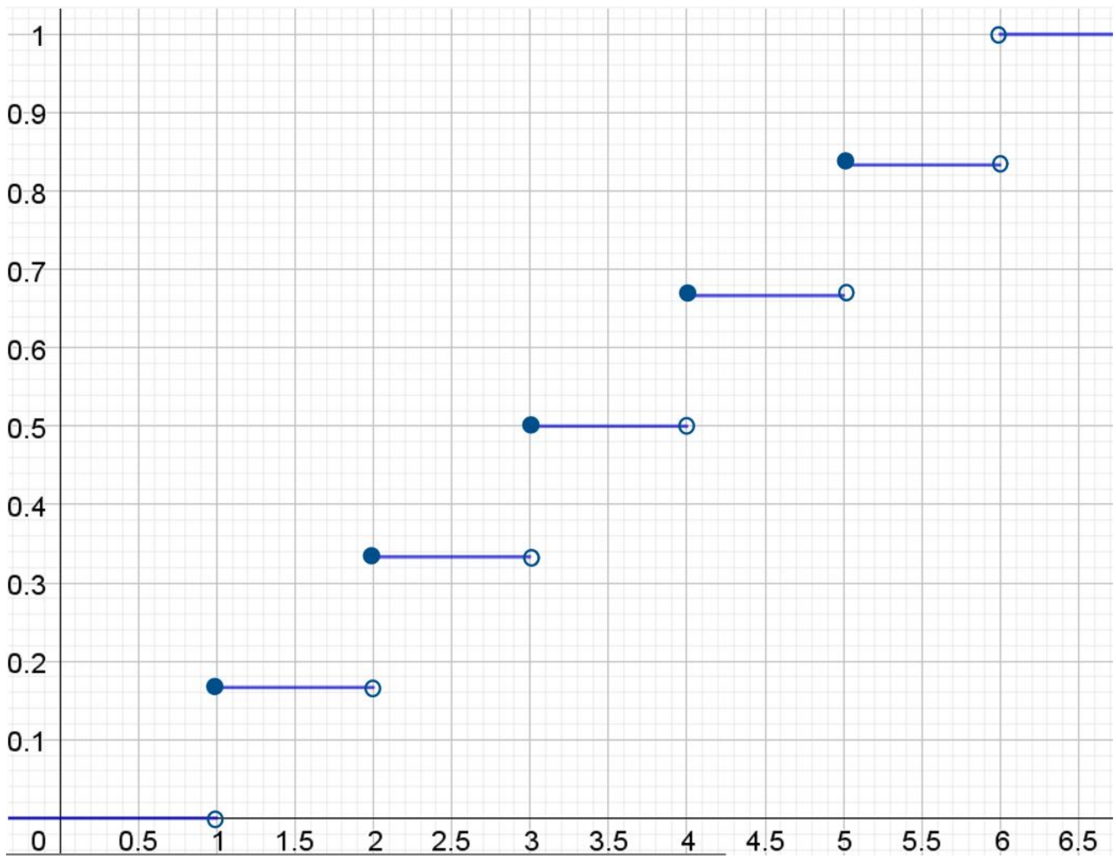
x_i	1	2	3	4	5	6
$F(x_i) = P(X \leq x_i)$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{4}{6}$	$\frac{5}{6}$	$\frac{6}{6} = 1$

Aufgrund Ihrer Definition ist die Verteilungsfunktion für den kleinsten Wert aller x_i immer identisch mit der Wahrscheinlichkeitsfunktion, während sie für den größten Wert der x_i immer gleich 1 sein muss. Der allgemeine Zusammenhang zwischen Wahrscheinlichkeitsfunktion und Verteilungsfunktion lautet:

Formel 2.2.2 *Wahrscheinlichkeitsfunktion und Verteilungsfunktion*

$$F(x_k) = \sum_{i=1}^{k} f(x_i) \tag{2.1}$$

Damit sieht der Graph der Verteilungsfunktion beim einfachen Würfeln wie folgt aus:



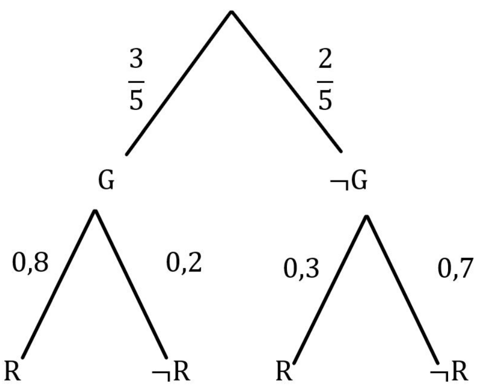
Beispiel 2.2.3 *Prüfung*

In einer mündlichen Prüfung muss ein Prüfling aus fünf Themengebieten zwei ziehen (der Einfachheit halber mit Zurücklegen). Bei drei Themengebieten ist er besser vorbereitet und hat eine Chance von 80% die Frage beantworten zu können. Bei den anderen beiden Themengebieten ist der Prüfling schlechter vorbereitet und hat nur eine Chance 30%. Es sei

X: Anzahl der richtig beantworteten Fragen

Stellen sie die Wahrscheinlichkeitsverteilung sowie die Verteilungsfunktion in Tabellenform sowie grafisch dar.

Lösung:
Für jede der beiden gezogenen und beantworteten Fragen gilt folgendes Baumdiagramm, wobei wir folgende Ereignisse definieren:
G: Das Themengebiet wurde gelernt
R: Die Frage wurde richtig beantwortet



Die Wahrscheinlichkeit eine Frage richtig/nicht richtig zu beantworten lautet also jeweils:

$$P(R) = \frac{3}{5} \cdot 0,8 + \frac{2}{5} \cdot 0,3 = 0,6 = 60\%$$

$$P(\neg R) = 1 - 0,6 = 0,4 = 40\%$$

Für die gesamte Prüfung ergibt sich also:

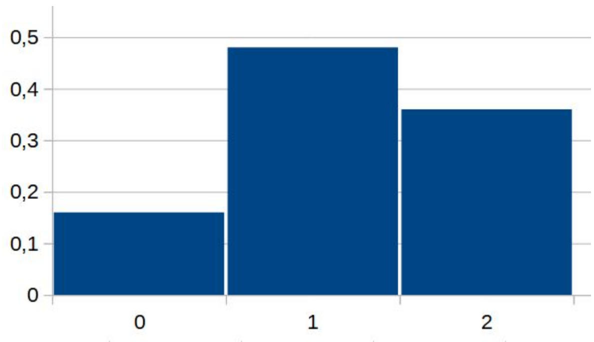
$$P(\text{beide Fragen richtig}) = 0,6^2 = 0,36 = 36\%$$

$$P(\text{beide Fragen falsch}) = 0,4^2 = 0,16 = 16\%$$

$$P(\text{genau eine Frage richtig}) = 1 - 0,36 - 0,16 = 0,48 = 48\%$$

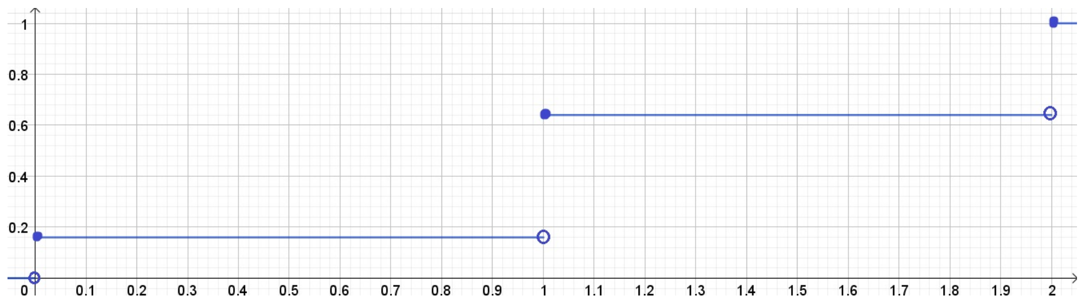
Damit ergibt sich folgende Verteilung der Zufallsvariable:

x_i	0	1	2
$P(X = x_i)$	$0,16$	$0,48$	$0,36$



Es ergibt sich folgende Verteilungsfunktion:

x_i	0	1	2
$F(x_i)$	$0,16$	$0,64$	1



Übung 2.2.4 *Prüfung*
Lösen Sie Beispiel **??** wenn nicht zurückgelegt wird.

2.2.2 Erwartungswert und Varianz

Ähnlich wie in der Statistik gibt es auch in der Wahrscheinlichkeitsrechnung Lage- und Streuungsmaße. Das Lagemaß einer Verteilung nennen wir den Erwartungswert, die Streuungsmaße heißen Varianz und Standardabweichung.

2.2.2.1 Der Erwartungswert

Der Erwartungswert μ bzw. $E(X)$ einer Verteilung ist folgendermaßen definiert:

Formel 2.2.5 *Erwartungswert*

$$\mu = E(X) = \sum_{i=1}^{n} x_i \cdot P(X = x_i) = x_1 \cdot P(X = x_1) + ... + x_n \cdot P(X = x_n) \tag{2.2}$$

Zur Interpretation des Erwartungswertes betrachten wir folgendes

Beispiel 2.2.6 *Werkstücke*

In einer Lieferung von befinden sich Werkstücke verschiedener Massen gemäß folgender Tabelle:

Masse in kg	2	4	6
Anzahl	5	8	7

Die durchschnittliche Masse $\overline{m}$ eines Werkstückes dieser Lieferung beträgt also

$$\begin{aligned}\overline{m} &= \frac{2 \cdot 5 + 4 \cdot 8 + 6 \cdot 7}{5 + 8 + 7} = \\ &= \frac{2 \cdot 5 + 4 \cdot 8 + 6 \cdot 7}{20} = \\ &= 2 \cdot \frac{5}{20} + 4 \cdot \frac{8}{20} + 6 \cdot \frac{7}{20}\end{aligned}$$

Dies entspricht genau unserer Formel des Erwartungswertes da die relativen Häufigkeiten $\frac{5}{20}, \frac{8}{20}, \frac{7}{20}$ auch den Wahrscheinlichkeiten für das Ziehen eines Werkstücks einer gewissen Masse entspricht. Wir interpretieren den Erwartungswert einer Zufallsvariable also als den *erwarteten Mittelwert* bei häufigem Wiederholen des Experiments.

Beispiel 2.2.7 *Prüfung*

Der Erwartungswert für die Anzahl der richtigen Antworten aus Beispiel **??** lautet

$$E(X) = 0 \cdot 0,16 + 1 \cdot 0,48 + 2 \cdot 0,36 = 1,2$$

Ein Prüfling mit einer solchen Vorbereitung wird also im Schnitt 1,2 Fragen pro Prüfung richtig beantworten.

Für den Erwartungswert gelten folgende Rechenregeln:

i) $E(X + a) = E(X) + a$

ii) $E(b \cdot X) = b \cdot E(X)$

iii) $E(X + Y) = E(X) + E(Y)$

Übung 2.2.8 *Roulette* 05_2022_04_10 Roulette_Lösungen.pdf

Beim Roulette gilt ein Mindesteinsatz von 10 €. Beim Setzen auf "rot" oder "schwarz" erhält man den einfachen Einsatz als Gewinn ausbezahlt, beim Setzen auf eines der Drittel den doppelten Einsatz und beim Setzen auf eine Zahl erhält man den 35-fachen Einsatz (bei den anderen Setz-Möglichkeiten dementsprechend).
i) Mit welchem Gewinn kann das Casino pro Spiel und Spieler im Schnitt mindestens rechnen? (Hinweis: Berechnen Sie jeweils den erwarteten Gewinn für das Setzen auf eine Farbe, auf eines der Drittel und für das Setzen auf eine Zahl.)
ii) Mit welchem Gewinn pro Tisch und Spiel kann das Casino rechnen wenn an einem Tisch immer 10 Spieler gleichzeitig spielen?
iii) Ein Spieler setzt immer 10 € auf "rot". Wenn er gewinnt hört er auf, wenn er verliert verdoppelt er seinen Einsatz. Wenn er nach 5 Spielen nicht gewonnen hat muss er aufhören weil er kein Geld mehr hat. Mit welcher Wahrscheinlichkeit erzielt der Spieler mit dieser Taktik einen (positiven) Gewinn und welcher Erwartungswert ergibt sich für seinen Gewinn?

2.2.2.2 Varianz, Standardabweichung und Kovarianz

Analog zur Statistik gibt es in der Wahrscheinlichkeitsrechnung die eng verwandten Streuungsmaße *Varianz V* und *Standardabweichung* $\sigma(X)$. Die Varianz ist definiert als die erwartete quadratische Abweichung vom Mittelwert, die Standardabweichung ist die Wurzel aus der Varianz.

Formel 2.2.9 *Varianz*

$$V(X) = E(X-\mu)^2 = \sum_{i=1}^{n} (x_i - \mu)^2 \cdot P(X = x_i) = \\ = (x_1 - \mu)^2 \cdot P(X = x_1) + ... + (x_n - \mu)^2 \cdot P(X = x_n) \tag{2.3}$$

Formel 2.2.10 *Standardabweichung*

$$\sigma(X) = \sqrt{V(X)} \tag{2.4}$$

Für die Varianz gilt außerdem folgender sog. *Verschiebungssatz* mit dem die Varianz oft etwas leichter berechnet werden kann.

Formel 2.2.11 *Verschiebungssatz der Varianz*

$$V(X) = E(X^2) - E(X)^2 = \sum_{i=1}^{n} (x_i^2 \cdot P(X = x_i)) - \mu^2 \tag{2.5}$$

Die Varianz als solche lässt sich im Allgemeinen nicht interpretieren, sie dient eher zum Vergleich der Streuung zweier Zufallsvariablen.

Beispiel 2.2.12 *Prüfung*

Für den Prüfling aus Beispiel **??** ergibt sich also folgende Varianz bzw. Standardabweichung:

$$V(X) = (0 - 1,2)^2 \cdot 0,16 + (1 - 1,2)^2 \cdot 0,48 + (2 - 1,2)^2 \cdot 0,36 = 0,48$$

bzw.

$$\sigma(X) = \sqrt{0,48} \approx 0,69$$

Alternative Berechnung der Varianz mit dem Verschiebungssatz:

$$V(X) = 0^2 \cdot 0,16 + 1^2 \cdot 0,48 + 2^2 \cdot 0,36 - 1,2^2 = 0,48$$

Für die Varianz gelten folgende Rechenregeln:

i) $V(X + a) = V(X)$

ii) $V(b \cdot X) = b^2 \cdot V(X)$

iii) $V(X + Y) = V(X) + V(Y)$ wenn X und Y unabhängig sind

Beispiel 2.2.13 *Aktienfonds*

Ein Aktienfonds lässt eine Rendite von 1500 € erwarten bei einer Varianz von 60. Ein zweiter, vom ersten unabhängiger Fonds bringt eine erwartete Rendite von 2500 € bei einer Varianz von 110 €. Sie besitzen 3 Fonds der ersten Sorte und zwei der zweiten. Mit welcher Gesamtrendite bei welcher Varianz ist zu rechnen? Wie lautet die Standardabweichung?

Lösung: Sei X: Gesamtrendite

$$E(X) = 3 \cdot 1500 + 2 \cdot 2500 = 9500 \text{ €}$$
$$V(X) = 3^2 \cdot 60 + 2^2 \cdot 110 = 980$$
$$\sigma(X) = \sqrt{980} \approx 31,30 \text{ €}$$

Hinweis: Beachten Sie, dass die Einheit der Varianz in diesem Fall streng genommen €^{2} wäre, was im Sachzusammenhang und bei der Interpretation aber keinen Sinn ergibt. Das ist einer der Gründe warum es daneben den Begriff der Standardabweichung gibt.

Übung 2.2.14 *Roulette*
Berechnen Sie die Varianz und die Standardabweichung für den Gewinn aus Sicht eines Spielers der beim Roulette auf "rot" oder "schwarz" setzt mit einem Einsatz von
i) 10 €
ii) 50 €

Eine stochastische Größe die den Zusammenhang zweier Verteilungen beschreibt ist die sog. *Kovarianz* den wir auch aus der Statistik kennen.

Formel 2.2.15 *Die Kovarianz*

Seien X und Y zwei Wahrscheinlichkeitsverteilungen deren Werte jeweils paarweise auftreten. Dann ist die Kovarianz folgendermaßen definiert:

$$\sigma_{XY} = Cov(X, Y) = E[(X - \mu_X) \cdot (Y - \mu_Y)] \tag{2.6}$$

Auch für die Kovarianz gibt es einen Verschiebungssatz:

Formel 2.2.16 *Verschiebungssatz der Kovarianz*

$$\sigma_{XY} = E(X \cdot Y) - \mu_X \cdot \mu_Y \tag{2.7}$$

Beispiel 2.2.17 *Gegeben ist folgende gemeinsame Verteilung. Berechnen sie die Kovarianz i) ohne und ii) mit Verschiebungssatz.*

x_i	1	2	4
y_i	2	4	7
$P(X = x_i \wedge Y = y_i)$	0,25	0,25	0,5

Lösung:

Wir berechnen zunächst:

$$\mu_X = 1 \cdot 0,25 + 2 \cdot 0,25 + 4 \cdot 0,5 = 2,75$$
$$\mu_Y = 2 \cdot 0,25 + 4 \cdot 0,25 + 7 \cdot 0,5 = 5$$

Und damit

i)

$$\sigma_{XY} = (1 - 2,75) \cdot (2 - 5) \cdot 0,25 + (2 - 2,75) \cdot (4 - 5) \cdot 0,25 + (4 - 2,75) \cdot (7 - 5) \cdot 0,5 = 2,75$$

ii)

$$\sigma_{XY} = 1 \cdot 2 \cdot 0,25 + 2 \cdot 4 \cdot 0,25 + 4 \cdot 7 \cdot 0,5 - 2,75 \cdot 5 = 2,75$$

Hinweis: Wie in der Statistik ergibt sich für die Kovarianz eine umso größere positive Zahl wenn zwischen den beiden Zufallsvariablen ein starker positiver Zusammenhang (je größer X desto größer Y) besteht, und eine umso kleinere negative Zahl bei einem starken negativen Zusammenhang (je größer X desto kleiner Y).
Besteht kein solcher Zusammenhang so ist das Ergebnis der Kovarianz nahe 0.

Übung 2.2.18 *Kovarianz*
Für zwei Zufallsvariablen X und Y sei folgendes bekannt:
$P(X = 1 \wedge Y = 2) = 0,2$, $P(X = 2 \wedge Y = 2) = 0,3$ und $P(X = 2 \wedge Y = 3) = 0,5$
Berechnen Sie die Kovarianz mit und ohne Verschiebungssatz.

Lösiung:
0,1 ?

Für die Kovarianz gilt folgende Rechenregel:

Formel 2.2.19

$$Cov(a \cdot X + b, c \cdot Y + d) = ac \cdot Cov(X, Y) \tag{2.8}$$

Gewisse Verteilungen die einem bestimmten Muster folgen und häufig anwendbar sind haben eigene Bezeichnungen und für sie gibt es spezielle Formeln. Davon werden wir nun einige kennenlernen.

2.2.3 Die Binomialverteilung

Um uns mit dieser Verteilung beschäftigen zu können benötigen wir vorher noch den sog.

Formel 2.2.20 *Binomialkoeffizient*

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!} \tag{2.9}$$

wobei

$$n! = 1 \cdot 2 \cdot ... \cdot n \quad \text{("n Fakultät“ oder "n Faktorielle“)} \tag{2.10}$$

Hinweis: An Taschenrechner gibt es für den Binomialkoeffizienten den Befehl nCr.
$\binom{n}{k}$ ("n über k“ oder "k aus n“) gibt die Anzahl der Möglichkeiten an aus einer n-elementigen Menge unterschiedlicher Elemente eine ungeordnete, k-elementige Stichprobe zu ziehen. Z.B. gibt es $\binom{45}{6}$ mögliche Ergebnisse einer Lottoziehung "6 aus 45“.

Nun können wir uns die Binomialverteilung ansehen.

Formel 2.2.21 *Formel der Binomialverteilung*

Ein Experiment mit 2 möglichen (bzw. unterschiedenen) Ergebnissen (genannt Treffer bzw. Nichttreffer) wird n mal wiederholt. Dabei sind die Wahrscheinlichkeiten für einen Treffer bzw. Nichttreffer konstant und werden mit p bzw. $q = 1 - p$ bezeichnet. Die Zufallsvariable sei

X:Anzahl der erzielten Treffer

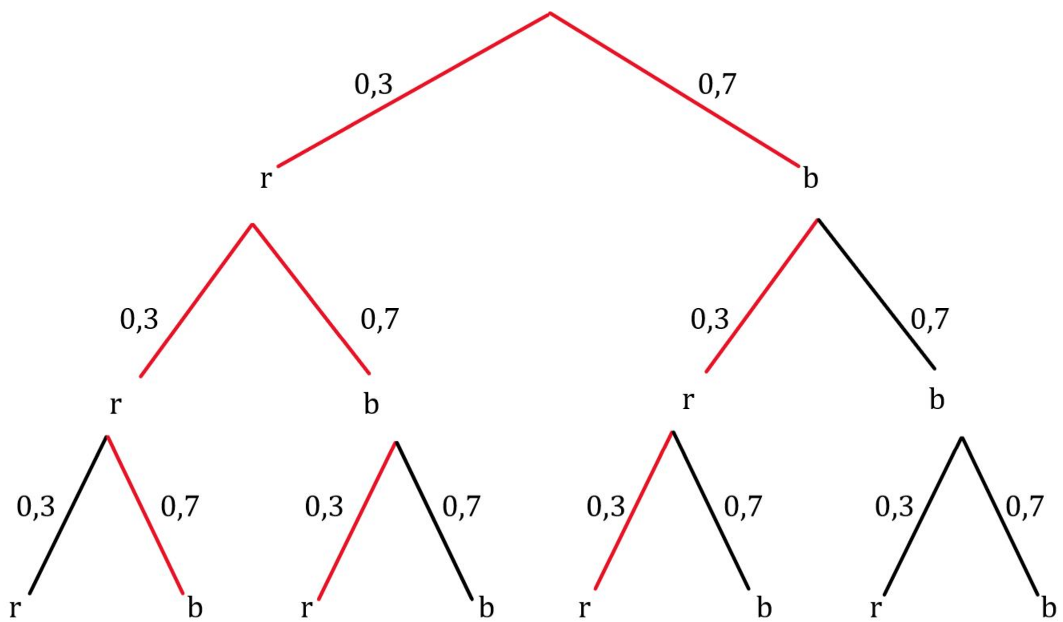
Dann heißt X binomialverteilt und für die Wahrscheinlichkeit für genau k Treffer gilt folgende

$$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k} = \binom{n}{k} \cdot p^k \cdot (q)^{n-k} \tag{2.11}$$

Eine binomialverteilte Zufallsvariable erhalten wir mit dem Urnenmodell bei mehrfachem Ziehen ohne Zurücklegen und ohne Beachtung der Reihenfolge. Durch das Zurücklegen ist die wichtige Voraussetzung der konstanten Werte für p und $1 - p$ erfüllt. Damit wollen wir uns anhand eines Beispiels die Formel herleiten.

Beispiel 2.2.22 *3-maliges Ziehen mit Zurücklegen*

Aus einer Urne mit 3 roten und 7 blauen Kugeln werden 3 Kugeln ohne Zurücklegen gezogen. Mit welcher Wahrscheinlichkeit zieht man genau 2 rote Kugeln?



Lösung:

Es sei *X: Anzahl der roten Kugeln*

Dann gilt:

Taschenrechner Taste: nCr

$$P(X=2) = 0,3 \cdot 0,3 \cdot 0,7 + 0,3 \cdot 0,7 \cdot 0,3 + 0,7 \cdot 0,3 \cdot 0,3 = 3 \cdot 0,3^2 \cdot 0,7^1 = \binom{3}{2} \cdot 0,3^2 \cdot 0,7^1$$ = 0.189??

Erläuterung: Es kommen die 3 markierten Pfade in Frage, die alle einzeln betrachtet das Ergebnis $0,3^2 \cdot 0,7^1$ liefern. Diese Anzahl 3 ergibt sich weil die beiden roten Kugeln an den Positionen 1 und 2, 1 und 3 oder 2 und 3 gezogen werden können. Das entspricht genau der Anzahl der möglichen Ergebnisse bei einer Lottoziehung "2 aus 3". Der Binomialkoeffizient $\binom{n}{k}$ berechnet in diesem Fall also die Anzahl der in Frage kommenden Pfade.

Bezeichnen wir das Ziehen einer roten Kugel als "Treffer" und damit die Trefferwahrscheinlichkeit mit $p = 0,3$ so entspricht das Ergebnis genau der Formel der Binomialverteilung.

Beispiel 2.2.23 *Flugtickets*

Bei einem Linienflug stehen 100 Sitzplätze zur Verfügung. Erfahrungsgemäß wird ein gebuchtes Ticket mit einer Wahrscheinlichkeit von 2% nicht in Anspruch genommen. Daher ist es gängige Praxis etwas mehr Tickets zu verkaufen um möglichst volle Auslastung zu erreichen.
Die Fluglinie verkauft 103 Tickets. Mit welcher Wahrscheinlichkeit geht die Rechnung der Fluglinie auf, sodass kein Fluggast mit Ticket den Flug trotzdem nicht antreten kann und entschädigt werden muss?

Es wurden 103 Tickets verkauft, 103 mal entscheidet sich also ob ein gebuchter Flug auch wahrgenommen wird oder nicht, und zwar jeweils mit einer konstanten Wahrscheinlichkeit von 2% den Flug nicht anzutreten. Damit handelt es ich um eine Binomialverteilung.

Es soll an diesem Beispiel verdeutlicht werden, dass es wichtig ist darauf zu achten, dass Zufallsvariable und Trefferwahrscheinlichkeit p zusammenpassen. Daher wird die Aufgabe auf 2 Arten gelöst.

Lösung 1
Es sei *X: Anzahl der Fluggäste die den Flug antreten möchten.*
Dann gilt: $p = 0,98$ und $n = 103$.
Die Rechnung der Fluggesellschaft geht auf wenn höchstens 100 Passagiere den Flug antreten. Das entspricht dem Ereignis $X \leq 100$.

$$
\begin{aligned}
P(X \leq 100) &= P(X=0) + ... + P(X=100) = \\
&= 1 - P(X \geq 101) = \\
&= 1 - (P(X=101) + ... + P(X=103)) = \\
&= 1 - \left(\binom{103}{101} \cdot 0,98^{101} \cdot 0,02^{2} + \binom{103}{102} \cdot 0,98^{102} \cdot 0,02^{1} + \binom{103}{103} \cdot 0,98^{103} \cdot 0,02^{0} \right) = \\
&= \cancel{0,3397 = 33,97\%}
\end{aligned}
$$

¿ =1,11411 ??

Lösung 2
Es sei *Y: Anzahl der Fluggäste die den Flug <u>nicht</u> antreten möchten.*
Dann gilt: $p = 0,02$ und $n = 103$.
Die Rechnung der Fluggesellschaft geht auf wenn höchstens 100 Passagiere den Flug antreten. Das entspricht dem Ereignis $Y \geq 3$.

$$
\begin{aligned}
P(Y \geq 3) &= P(Y=3) + ... + P(Y=103) = \\
&= 1 - P(Y \leq 2) = \\
&= 1 - (P(Y=0) + ... + P(Y=2)) = \\
&= 1 - \left(\binom{103}{0} \cdot 0,02^{0} \cdot 0,02^{103} + \binom{103}{1} \cdot 0,02^{1} \cdot 0,02^{102} + \binom{103}{2} \cdot 0,02^{2} \cdot 0,98^{101} \right) = \\
&= 0,3397 = 33,97\%
\end{aligned}
$$

Für Erwartungswert, Varianz und Standardabweichung einer Binomialverteilung gibt es sehr einfache Formeln:

$$E(X) = n \cdot p \tag{2.12}$$

$$V(X) = n \cdot p \cdot (1-p) = n \cdot p \cdot q \tag{2.13}$$

$$\sigma(X) = \sqrt{n \cdot p \cdot (1-p)} = \sqrt{n \cdot p \cdot q} \tag{2.14}$$

Beispiel 2.2.24 *Flugtickets*

Bei $n = 103$ verkauften Tickets und einer Absagequote von 2% werden beträgt die Anzahl der im Schnitt zu erwartenden Absagen

$$E(X) = 103 \cdot 0,02 = 2,06$$

Abschließend sehen wir uns noch 2 Beispiele von Binomialverteilungen veranschaulicht als Säulendiagramm an:

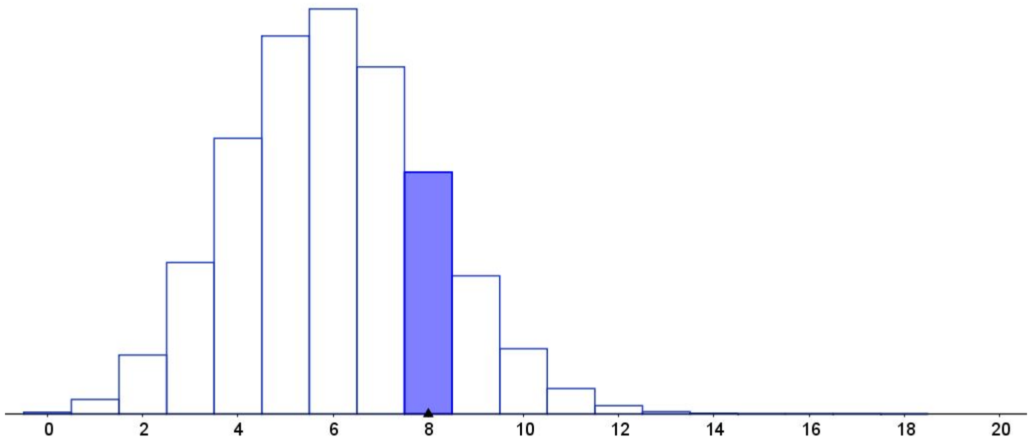


Abbildung 2.1: n=20; p=0,3; Dargestellt ist $P(X = 8)$

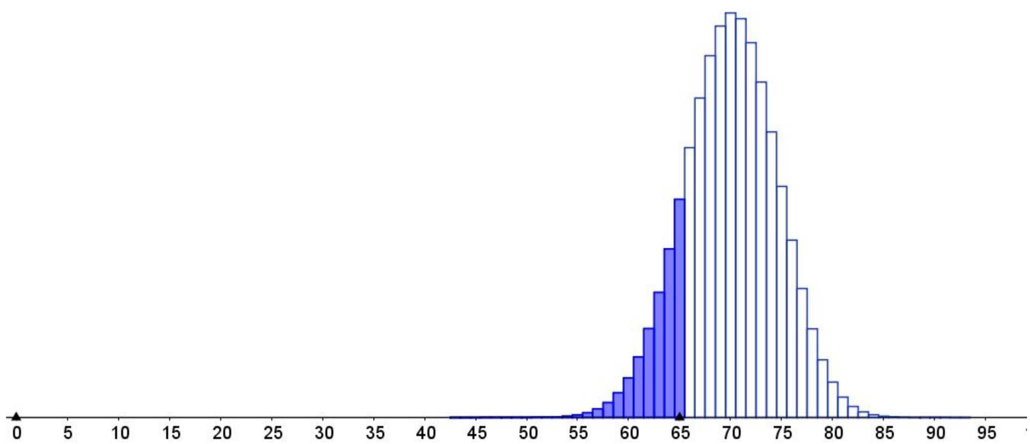


Abbildung 2.2: n=100; p=0,7; Dargestellt ist $F(65) = P(X \leq 65)$

An dieser Stelle soll nochmals darauf hingewiesen werden, dass ein Balken zum x-Wert k die Fläche $P(X = k)$ besitzt. Damit entspricht die markierte Fläche in Abbildung **??** dem Wert der Verteilungsfunktion $F(65) = P(X \leq 65)$.

Lösung Skript_05_2022_04_10 Roulette_Lösung.pdf

Übung 2.2.25 *Schadensumme*
In einem Betrieb kommen drei identische Maschinen zum Einsatz. Bei jeder Maschine kann es nach einem Stromausfall zu einem Ausfall des Bauteils A bzw. B kommen das sofort ersetzt werden muss. Berechnen Sie die Verteilung der Gesamtschadensumme S wenn
i) alle Bauteile mit einer Wahrscheinlichkeit von $p = 0,1$ ausfallen und jeweils 5000 € kosten
ii) alle Bauteile mit einer Wahrscheinlichkeit von $p = 0,1$ ausfallen, Bauteil A 5000 € und Bauteil B 10000 € kostet
iii) Bauteil A mit einer Wahrscheinlichkeit von $p = 0,1$ ausfällt, Bauteil B mit $p = 0,2$ und beide Bauteile jeweils 5000 € kosten
iv) Bauteil A mit einer Wahrscheinlichkeit von $p = 0,1$ ausfällt, Bauteil B mit $p = 0,2$, Bauteil A 5000 € und Bauteil B 10000 € kostet.
Berechnen Sie auch jeweils den Erwartungswert.

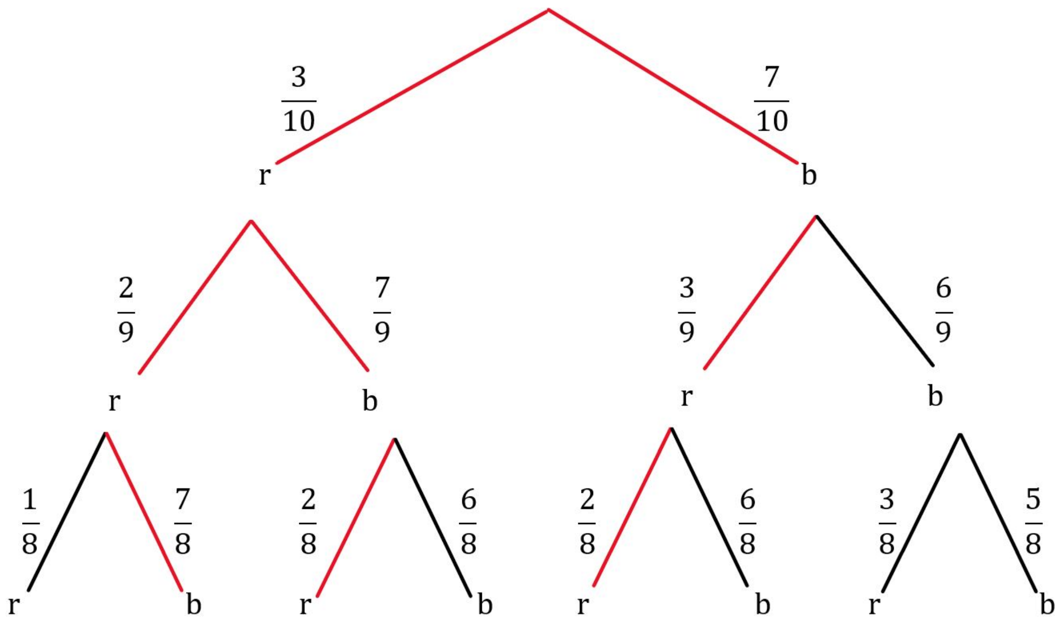
Als nächstes betrachten wir das Urnenmodell ohne Zurücklegen.

2.2.4 Die Hypergeometrische Verteilung

Wieder wollen wir den Formalismus anhand eines Beispiels betrachten:

Beispiel 2.2.26 *3-maliges Ziehen ohne Zurücklegen*

Aus einer Urne mit 3 roten und 7 blauen Kugeln werden 3 Kugeln ohne Zurücklegen gezogen. Mit welcher Wahrscheinlichkeit zieht man genau 2 rote Kugeln?



Es sei wieder *X: Anzahl der roten Kugeln*

Dann gilt:

$$\begin{aligned} P(X=2) &= \frac{3}{10}\cdot\frac{2}{9}\cdot\frac{7}{8}+\frac{3}{10}\cdot\frac{7}{9}\cdot\frac{2}{8}+\frac{7}{10}\cdot\frac{3}{9}\cdot\frac{2}{8}= \\ &= 3\cdot\frac{3\cdot 2\cdot 7}{10\cdot 9\cdot 8} \end{aligned}$$

Durch Umformen werden wir die allgemeine Formel erhalten. Der Faktor 3 beschreibt wieder die Anzahl der Pfade, dafür können wir also wieder $\binom{3}{2}=\frac{3!}{2!\cdot 1!}$ schreiben. Außerdem gilt $10\cdot 9\cdot 8=\frac{10!}{7!}$, $3\cdot 2=\frac{3!}{1!}$ und $7=\frac{7!}{6!}$ und damit:

$$\begin{aligned} P(X=2) &= \frac{3!}{2!\cdot 1!}\cdot\frac{\frac{3!}{1!}\cdot\frac{7!}{6!}}{\frac{10!}{7!}}= \\ &= \frac{\frac{3!}{2!\cdot 1!}\cdot\frac{7!}{6!\cdot 1!}}{\frac{10!}{7!\cdot 3!}}= \\ &= \frac{\binom{3}{2}\cdot\binom{7}{1}}{\binom{10}{3}} \end{aligned}$$

Diese Darstellung entspricht folgender Überlegung:
Es gibt $\binom{3}{2}$ Möglichkeiten welche 2 der 3 roten Kugeln gezogen werden und $\binom{7}{1}$ Möglichkeiten welche der 7 blauen Kugeln gezogen wird, also ist die Anzahl der günstigen Fälle gleich $\binom{3}{2}\cdot\binom{7}{1}$. Außerdem gibt es $\binom{10}{3}$ welche 3 Kugeln überhaupt gezogen werden, das ist also die Anzahl der möglichen Fälle. Der Term entspricht also genau der Formel (**??**) der Laplace Wahrscheinlichkeit. Damit können wir eine allgemeine Formel der Hypergeometrischen Verteilung formulieren:

Formel 2.2.27 *Formel der Hypergeometrischen Verteilung Gegeben sei eine Menge mit N Elementen die sich in 2 Kategorien mit je M Elementen der 1. Kategorie und $N-M$ Elementen der 2. Kategorie aufteilen lässt. Es werden n Elemente ohne Zurücklegen gezogen. Es sei:*

X: Anzahl der gezogenen Elemente der 1. Kategorie

Dann gilt:

$$P(X=m)=\frac{\binom{M}{m}\cdot\binom{N-M}{n-m}}{\binom{N}{n}} \tag{2.15}$$

Beispiel 2.2.28 *Fahrkartenkontrolle*

In einem U-Bahn-Waggon befinden sich 40 Fahrgäste, 5 davon haben kein gültiges Ticket. Bei einer Kontrolle werden 5 Personen kontrolliert. Mit welcher Wahrscheinlichkeit werden *i) genau 2* bzw. *ii) höchtens 1* SchwarzfahrerInnen erwischt?

Lösung:

Es sei: *X: Anzahl der erwischten SchwarzfahrerInnen*

i. Die $N = 40$ Fahrgäste teilen sich auf zu $M = 5$ Personen ohne bzw. $N - M = 35$ mit Fahrschein.Von den 5 SchwarzfahrerInnen sollen $m = 2$ erwischt werden.

$$P(X = 2) = \frac{\binom{5}{2} \cdot \binom{35}{3}}{\binom{40}{5}} = 0,0995 = 9,95\%$$

5 über 2 = 5 nCr 2 im Taschenrechner

ii. Das Ereignis "höchstens 1 erwischte SchwarzfahrerIn“ formulieren wir als $X \leq 1$ und addieren die Wahrscheinlichkeiten der (unvereinbaren) Einzelereignisse $X = 0$ und $X = 1$:

$$P(X \leq 1) = \frac{\binom{5}{0} \cdot \binom{35}{5}}{\binom{40}{5}} + \frac{\binom{5}{1} \cdot \binom{35}{4}}{\binom{40}{5}} = 0,8912 = 89,12\%$$

Auch für Erwartungswert, Varianz und Standardabweichung einer Hypergeometrischen Verteilung gibt es einfache Formeln:

$$E(X) = n \cdot \frac{M}{N} \tag{2.16}$$

$$V(X) = n \cdot \frac{M}{N} \cdot (1 - \frac{M}{N}) \cdot \frac{N-n}{N-1} \tag{2.17}$$

$$\sigma(X) = \sqrt{n \cdot \frac{M}{N} \cdot (1 - \frac{M}{N}) \cdot \frac{N-n}{N-1}} \tag{2.18}$$

Beispiel 2.2.29 *Fahrkartenkontrolle*

Die im Schnitt zu erwartende Anzahl der erwischten SchwarzfahrerInnen bei der Kontrolle wie aus Beispiel **??** beträgt:

$$E(X) = 5 \cdot \frac{5}{40} = 0,625$$

Die Formel der Hypergeometrischen Verteilung lässt sich auch noch etwas verallgemeinern:

Formel 2.2.30 *Formel der verallgemeinerten Hypergeometrischen Verteilung*

Gegeben sei eine Menge mit N Elementen die sich in k Kategorien mit M_i Elementen der i-ten Kategorie aufteilen lassen mit $i = 1, ..., k$. Es werden n Elemente ohne Zurücklegen gezogen. Es sei:

X_i: Anzahl der gezogenen Elemente der i-ten Kategorie

Dann gilt:

$$P(X_1 = m_1, ..., X_k = m_k) = \frac{\binom{M_1}{m_1} \cdots \binom{M_k}{m_k}}{\binom{N}{n}} \tag{2.19}$$

Beispiel 2.2.31 *Poker*

Ein Kartenspiel beim Poker besteht aus 52 Karten, jeder Spieler erhält 5 Karten. Mit welcher Wahrscheinlichkeit bekommt man einen *i) Full House der höchsten Wertigkeit (3 Asse, 2 Könige)* bzw. *ii) einen allgemeinen Full House (Drilling+Zwilling)*?

Lösung:

i. Es gibt $M_1 = 4$ Asse, $M_2 = 4$ Könige und $M_3 = 44$ andere Karten.

$$P(\text{3 Asse, 2 Könige}) = \frac{\binom{4}{3} \cdot \binom{4}{2} \cdot \binom{44}{0}}{\binom{52}{5}} = 0,00000923 = 0,000923\%$$

ii. Um die Anzahl der günstigen Fälle zu berechnen überlegen wir uns, dass es nochmal so viele Möglichkeiten gibt einen Full House aus Assen und Königen zu bilden (nämlich 3 Könige, 2 Asse) und $\binom{13}{2}$ Möglichkeiten welche beiden Kartenwerte beteiligt sein könnten.

$$P(\text{Full House}) = \frac{\binom{4}{3} \cdot \binom{4}{2} \cdot \binom{44}{0}}{\binom{52}{5}} \cdot 2 \cdot \begin{pmatrix} 13 \\ 2 \end{pmatrix} = 0,00144 = 0,144\%$$

Auch hier sehen wir uns noch eine Abbildung eines Säulendiagramms fest und erkennen eine große Ähnlichkeit zur Binomialverteilung.

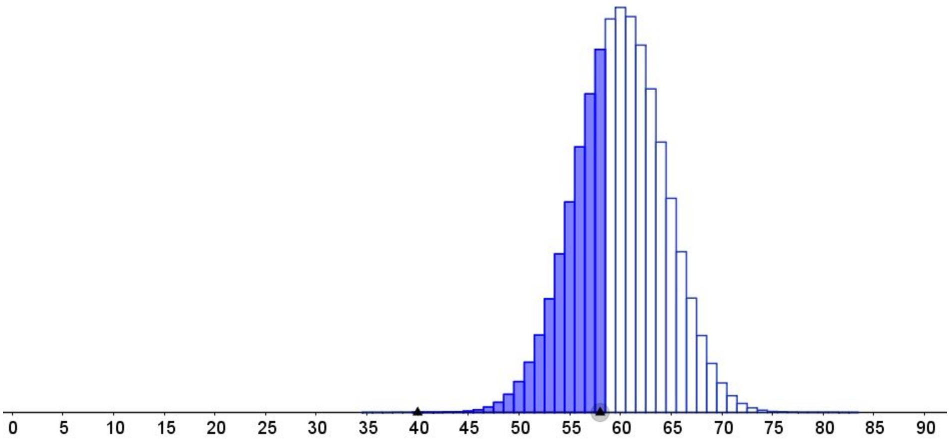


Abbildung 2.3: N=500; M=300; n=100; Dargestellt ist $F(65) = P(X \leq 65)$

Übung 2.2.32 *Versandhandel*
Ein Versandhandel hat in seinem Lager 500 Stück eines Produktes vorrätig von denen 20 jedoch defekt sind. Ein Kunde bestellt 5 Stück dieses Produkts. Mit welcher Wahrscheinlichkeit erhält der Kunde
i) genau 2
ii) mindestens 1
iii) höchstens 1
defekte Produkte zugestellt?

2.2.5 Die Poisson-Verteilung

Als letzten Vertreter der diskreten Verteilungen lernen wir nun noch die Poisson-Verteilung kennen. Mit der Poisson-Verteilung kann man Situationen beschreiben bei der die Wahrscheinlichkeit der Häufigkeit eines Ereignisses in einer gewissen vorgegeben Einheit berechnet werden soll (sie wird auch "Verteilung der seltenen Ereignisse“ genannt). Wir werden gleich sämtliche relevanten Formeln kennenlernen.

Formel 2.2.33 *Die Poisson-Verteilung*

Es sei λ der Mittelwert der Häufigkeit eines gewissen Ereignisses in einer vorgegeben Einheit.

Außerdem sei X: Häufigkeit des Ereignisses pro Einheit

Dann gilt:

$$P(X = x) = \frac{\lambda^x}{x!} \cdot e^{-\lambda} \tag{2.20}$$

Taschenrechner x! (2nd3)

wobei $e = 2,718281828459...$ eine irrationale Konstante, genannt *Euler'sche Zahl* ist.

Außerdem gilt:

$$E(X) = \lambda \tag{2.21}$$

$$V(X) = \lambda \tag{2.22}$$

$$\sigma(X) = \sqrt{\lambda} \tag{2.23}$$

Auch die Poisson-Verteilung hat also eine Verwandtschaft zur Binomialverteilung, allerdings ist zu beachten, dass der Zufallsvariable X bei der Poisson-Verteilung nach oben keine Grenze gesetzt ist während die Treffer-Anzahl bei der Binomialverteilung nach oben durch die Anzahl der Wiederholungen n beschränkt ist.

Beispiel 2.2.34 *Gewitter*

In Österreich kommt es im Schnitt alle 146 Tage zu einem Todesfall durch einen Blitzeinschlag. Mit welcher Wahrscheinlichkeit kommt es in einem Jahr zu *i) gar keinem* bzw. *ii) höchsten 3* Blitztoten?

Lösung:

Zunächst berechnen wir den Parameter λ als Durchschnittswert der Blitztoten pro Jahr:

$$\lambda = \frac{365}{146} = 2,5$$

Man erhält also:

i. $P(X = 0) = \frac{2,5^0}{0!} \cdot e^{-2,5} = e^{-2,5} = 0,082 = 8,2\%$

   Hinweis: $a^0 = 1$ und $0! = 1$

ii. $P(X \leq 3) = \frac{2,5^0}{0!} \cdot e^{-2,5} + \frac{2,5^1}{1!} \cdot e^{-2,5} + \frac{2,5^2}{2!} \cdot e^{-2,5} + \frac{2,5^3}{3!} \cdot e^{-2,5} = 0,7576 = 75,76\%$

Die Einheit auf die sich die Häufigkeit eines Ereignisses bezieht ist sehr oft eine Zeiteinheit, es kommen aber auch andere Einheiten in Frage.

Beispiel 2.2.35 *Buchdruck*

Beim Druck eines Buches kommt es auf 100 Seiten im Schnitt zu 7 Druckfehlern. Mit welcher Wahrscheinlichkeit hat ein 250-seitiges Buch höchstens 1 Druckfehler?

Lösung:

Es sei: *X: Anzahl der Druckfehler pro 250 Seiten*

Die Einheit auf die sich die Fragestellung bezieht ist also in dem Fall ein Buch mit 250 Seiten. Für die Anzahl der Druckfehler pro 250 Seiten beträgt der Mittelwert also:

$$\lambda = 7 \cdot 2,5 = 17,5$$

Damit erhalten wir:

$$\begin{aligned} P(X \leq 1) = & \frac{17,5^0}{0!} \cdot e^{-17,5} + \frac{17,5^1}{1!} \cdot e^{-17,5} = \\ & = e^{-17,5} + 17,5 \cdot e^{-17,5} = \\ & = 18,5 \cdot e^{-17,5} = \\ & = 0,0000000251 = 0,00000251\% \end{aligned}$$

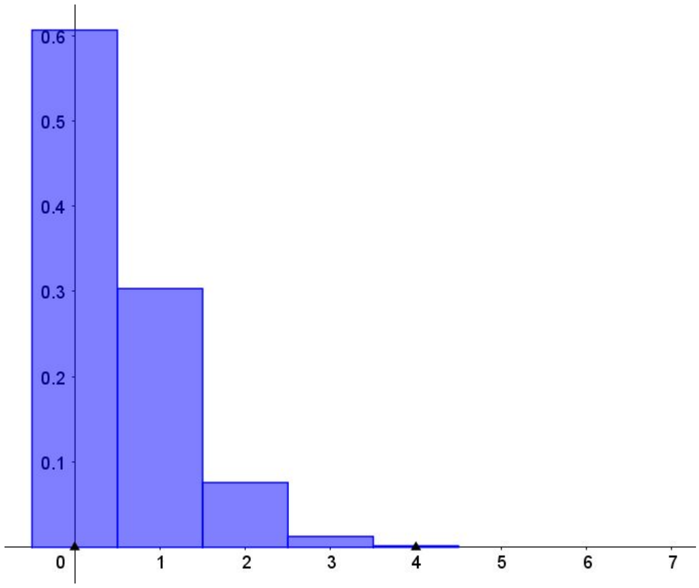


Abbildung 2.4: Poisson-Verteilung mit $\lambda = 0,5$

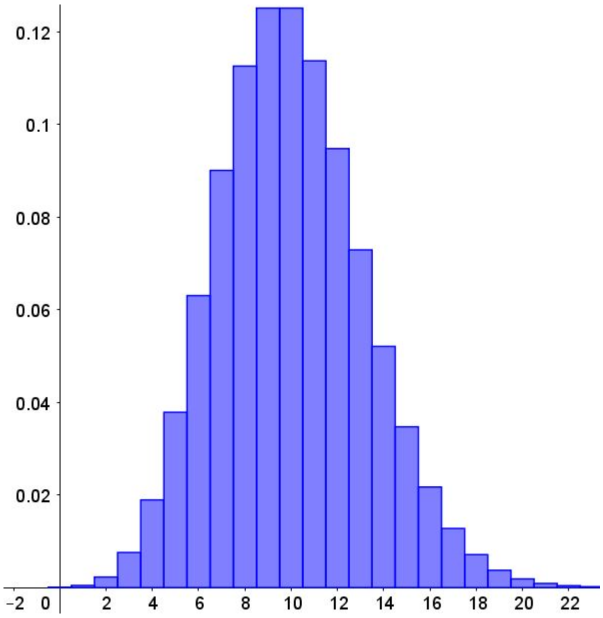


Abbildung 2.5: Poisson-Verteilung mit $\lambda = 10$

Übung 2.2.36 *Erdbeben*
Ein Versicherungsunternehmen interessiert sich für die Anzahl der Erdbeben in einer gewissen Region. Erfahrungsgemäß kommt es alle 2,7 Jahre zu einem Erbeben. Mit welcher Wahrscheinlichkeit gibt es

i) ein Erdbeben im kommenden Jahr
ii) zwei Erdbeben im kommenden Jahr
iii) höchstens ein Erdbeben in den kommenden zwei Jahren.

2.3 Stetige Verteilungen

Diskrete Verteilungen ordnen einem oder mehreren einzelnen (diskreten) Werten eine Wahrscheinlichkeit zu. Es gibt aber auch Zufallsvariablen die in einem gewissen Bereich der reellen Zahlen jede beliebige reelle Zahl annehmen können. Da in jedem noch so kleinen Intervall der reellen Zahlen unendlich viele Zahlen enthalten sind und somit einem einzelnen (günstigen) Wert x unendlich viele andere (mögliche) Werte gegenüber stehen gilt für eine stetige Zufallsvariable X grundsätzlich:

Formel 2.3.1

Sei X eine stetige Zufallsvariable. Dann gilt:

$$P(X = a) = 0 \tag{2.24}$$

Es werden also nur Zahlenbereichen, d.h. Ereignissen der Form $X > a$, $X \leq b$, $a \leq X \leq b$ usw. Wahrscheinlichkeiten zugeordnet.

2.3.1 Dichtefunktion und Verteilungsfunktion

Wir haben bei den diskreten Verteilungen schon gesehen, dass eine Wahrscheinlichkeit der Form $P(X \leq x)$ der Gesamtfläche aller Balken links des Balkens zum Wert x entspricht (vergleiche z.B. Abbildung **??**).
Bei stetigen Verteilungen ist es in gewisser Weise ähnlich. Jeder stetigen Verteilung liegt eine sog. *Dichtefunktion* $f(x)$ zugrunde, die Wahrscheinlichkeit eines Ereignisses der Form $X \leq x$ entspricht der Fläche unter dem Graphen dieser Funktion links des Wertes x. Die Funktion die $F(x) = P(X \leq x)$ nennen wir wieder die *Verteilungsfunktion*, die Berechnung der Fläche und damit der Wahrscheinlichkeit geschieht mittels Integralrechnung (ist aber bei gewissen Verteilungen auch mittels geometrischer Formeln möglich).

Formel 2.3.2 *Verteilungsfunktion einer stetigen Verteilung*

Sei X eine stetige Zufallsvariable mit Dichtefunktion $f(x)$. Dann gilt für ihre Verteilungsfunktion $F(x)$:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt \tag{2.25}$$

Da die Gesamtwahrscheinlichkeit einer Verteilung immer gleich $100\% = 1$ sein muss gilt dies also für die Gesamtfläche unter dem Graphen einer Dichtefunktion.

Formel 2.3.3 *Dichtefunktion einer stetigen Verteilung*

Die Funktion $f(x)$ heißt Dichtefunktion wenn gilt:

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{2.26}$$

Beispiel 2.3.4 *Dichtefunktion*

Die Funktion

$$f(x) = \begin{cases} -\frac{1}{2}x + 1 & \textit{für } 0 < x < 2 \\ 0 & \textit{sonst} \end{cases}$$

ist eine Dichtefunktion. Der Graph dieser Funktion sieht folgendermaßen aus:

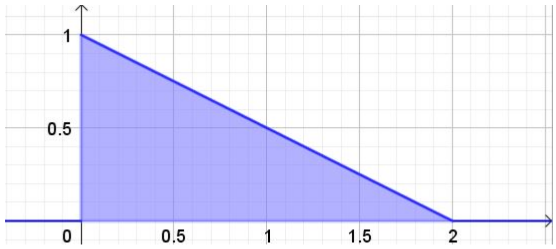


Abbildung 2.6: Graph der Dichtefunktion f aus Beispiel **??**

Begründung:

- rechnerisch: Das Integral muss nur in dem Bereich berechnet werden in dem die Funktion nicht 0 ist:

$$\int_0^2 (-\frac{1}{2}x + 1)dx = \left[-\frac{1}{4}x^2 + x\right]_0^2 = \left(-\frac{1}{4} \cdot 2^2 + 2\right) - \left(-\frac{1}{4} \cdot 0^2 + 0\right) = 1$$

- geometrisch: Die markierte Fläche ist ein rechtwinkliges Dreieck mit den Kathetenlängen 1 und 2 (bzw. ein halbes Rechteck mit der Länge 1 und der Breite 2, also folgt für die Fläche:

$$A = \frac{1}{2} \cdot 1 \cdot 2 = 1$$

Drei häufig gesuchte Wahrscheinlichkeiten können mit Dichtefunktion bzw. Verteilungsfunktion folgendermaßen dargestellt werden:

Formel 2.3.5 *Wahrscheinlichkeiten stetiger Verteilungen*

$$i) \quad P(X \leq a) = \int_{-\infty}^{a} f(x)dx = F(a) \tag{2.27}$$

$$ii) \quad P(X \geq a) = \int_{a}^{\infty} f(x)dx = 1 - F(a) \tag{2.28}$$

$$iii) \quad P(a \leq X \leq b) = \int_{a}^{b} f(x)dx = F(b) - F(a) \tag{2.29}$$

Hinweis: Die Verteilungsfunktion F ist also Stammfunktion der Dichtefunktion f, bzw.

$$F'(x) = f(x)$$

Beispiel 2.3.6 *Verteilungsfunktion*

Für die Verteilungsfunktion $F(x)$ zur Dichtefunktion $f(x)$ aus Beispiel **??** gilt also:

$$F(x) = \begin{cases} 0 & \textit{für } x < 0 \\ -\frac{1}{4}x^2 + x & \textit{für } 0 \leq x \leq 2 \\ 1 & \textit{für } x > 2 \end{cases}$$

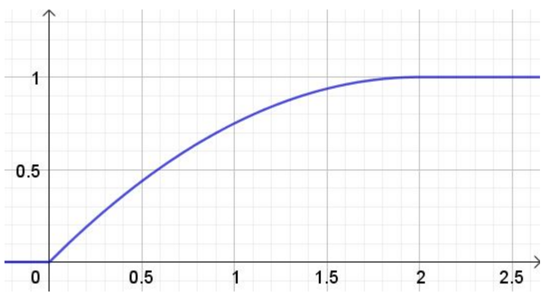


Abbildung 2.7: Graph der Verteilungsfunktion F aus Beispiel **??**

Hinweis: *Eigenschaften der Verteilungsfunktion*

i) Für kleine x-Werte tendiert $F(x)$ gegen 0
ii) Für große x-Werte tendiert $F(x)$ gegen 1
iii) $F(x)$ ist monoton steigend

Beispiel 2.3.7 *Berechnung von Wahrscheinlichkeiten*

Berechnen sie für die Verteilung aus Beispiel **??** folgende Wahrscheinlichkeiten und stellen sie diese mittels Dichtefunktion und Verteilungsfunktion grafisch dar.

i) $P(X \leq 0,5)$
ii) $P(X \geq 1)$
iii) $P(0,5 \leq X \leq 1)$

Lösung:

i)

$$P(X \leq 0,5) = F(0,5) = \left(-\frac{1}{4} \cdot 0,5^2 + 0,5\right) = 0,4375 = 43,75\%$$

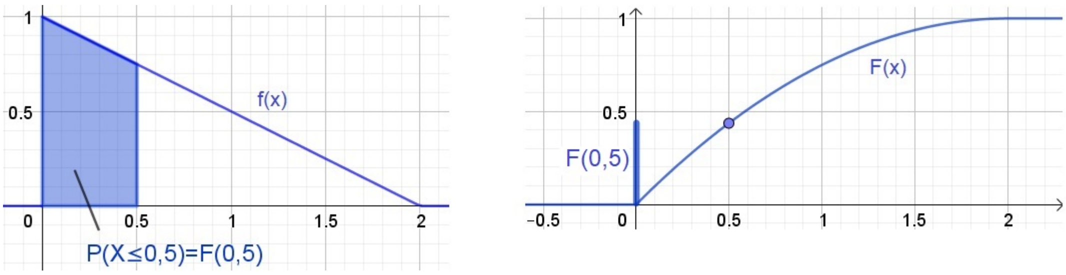


Abbildung 2.8: $P(X \leq 0,5) = F(0,5)$ dargestellt anhand $f(x)$ bzw. $F(x)$

ii)

$$P(X \geq 1) = 1 - F(1) = 1 - \left(-\frac{1}{4} \cdot 1^2 + 1\right) = 0,25 = 25\%$$

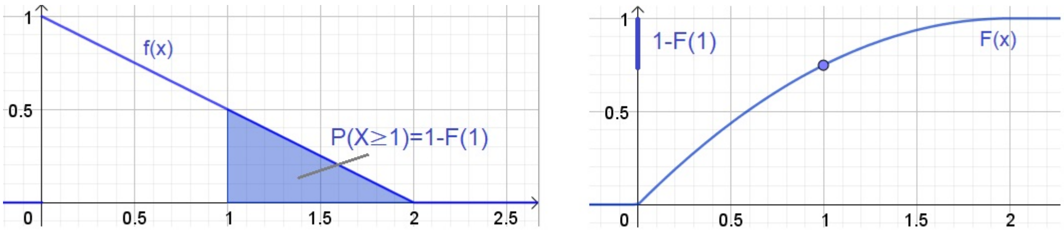


Abbildung 2.9: $P(X \geq 1) = 1 - F(1)$ dargestellt anhand $f(x)$ bzw. $F(x)$

iii)

$$P(0,5 \leq X \leq 1) = F(1) - F(0,5) = \left(-\frac{1}{4} \cdot 1^2 + 1\right) - \left(-\frac{1}{4} \cdot 0,5^2 + 0,5\right) =$$
$$= 0,75 - 0,4375 = 0,3125 = 31,25\%$$

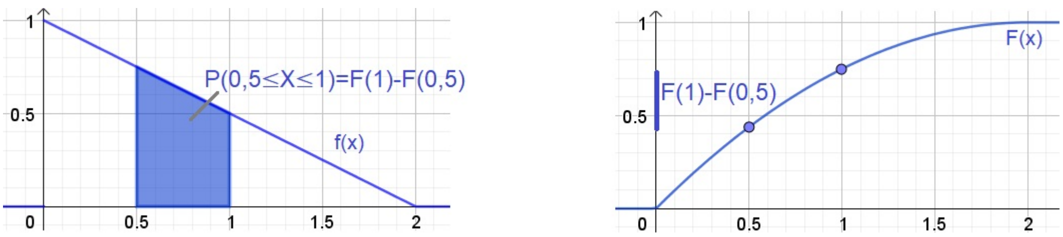


Abbildung 2.10: $P(0,5 \leq X \leq 1) = F(1) - F(0,5)$ dargestellt anhand $f(x)$ bzw. $F(x)$

Übung 2.3.8 *Dichtefunktion und Verteilungsfunktion*
i) Zeigen Sie, dass die die folgende Funktion eine Dichtefunktion ist und bestimmen Sie die zugehörige Verteilungsfunktion

$$f(x) = \begin{cases} 0 & \textit{für } x < -2 \\ \frac{1}{8}x + \frac{1}{4} & \textit{für } -2 \leq x \leq 2 \\ 0 & \textit{für } x > 2 \end{cases}$$

ii) Berechnen Sie für diese Verteilung die Wahrscheinlichkeiten $P(X < 0)$ und $P(X \geq 1)$.

2.3.2 Erwartungswert, Varianz und Standardabweichung

Der Graph der Dichtefunktion ist die stetige Entsprechung zum Säulendiagramm im diskreten Fall da in beiden Fällen die Gesamtfläche aller Säulen bzw. die Gesamtfläche unter dem Graphen der Gesamtwahrscheinlichkeit von 100% entspricht. Daher sind die Formeln für Erwartungswert und Varianz in ähnlicher Weise definiert: An die Stelle der Wahrscheinlichkeiten treten die Funktionswerte der Dichtefunktion, statt der Summenbildung wird integriert.
Hinweis: Vergleiche dazu die Herleitung des Integrals mit Ober- und Untersummen!

2.3.2.1 Der Erwartungswert

Formel 2.3.9 *Erwartungswert einer stetigen Verteilung*

Sei $f(x)$ die Dichtefunktion einer stetigen Zufallsvariable. Dann lässt sich der Erwartungswert mit folgender Formel berechnen:

$$\mu = E(X) = \int_{-\infty}^{\infty} (x \cdot f(x))dx \tag{2.30}$$

Beispiel 2.3.10 *Erwartungswert*

Für die Dichtefunktion $f(x) = \begin{cases} -\frac{1}{2}x + 1 & \textit{für } 0 < x < 2 \\ 0 & \textit{sonst} \end{cases}$ aus Beispiel **??** ergibt sich folgender Erwartungswert:

$$\begin{aligned} E(X) &= \int_0^2 (x \cdot (-\frac{1}{2}x + 1))dx = \\ &= \int_0^2 (-\frac{1}{2}x^2 + x))dx = \\ &= \left[-\frac{1}{2} \cdot \frac{x^3}{3} + \frac{x^2}{2} \right]_0^2 = \\ &= \left(-\frac{2^3}{6} + \frac{2^2}{2} \right) - 0 = \\ &= -\frac{4}{3} + \frac{6}{3} = \\ &= \frac{2}{3} \end{aligned}$$

2.3.2.2 Die Varianz

Formel 2.3.11 *Varianz einer stetigen Verteilung*

Sei $f(x)$ die Dichtefunktion einer stetigen Zufallsvariable. Dann lässt sich die Varianz mit folgender Formel berechnen:

$$\sigma^2 = V(X) = \int_{-\infty}^{\infty} \Big((x - \mu)^2 \cdot f(x) \Big) \, dx \tag{2.31}$$

Und auch im stetigen Fall gibt es einen entsprechend formulierten Verschiebungssatz:

Formel 2.3.12 *Verschiebungssatz für stetige Verteilungen*

Die Varianz einer stetigen Verteilung mit Dichtefunktion $f(x)$ kann berechnet werden mit der Formel:

$$\sigma^2 = V(X) = \int_{-\infty}^{\infty} \Big(x^2 \cdot f(x) \Big) \, dx - \mu^2 \tag{2.32}$$

Beispiel 2.3.13 *Varianz*

Berechne für die Dichtefunktion $f(x) = \begin{cases} -\frac{1}{2}x + 1 & \textit{für } 0 < x < 2 \\ 0 & \textit{sonst} \end{cases}$ aus Beispiel **??** die Varianz.

Lösung:

Oft (aber nicht immer) ist die Berechnung des Integrals mit dem Verschiebungssatz etwas weniger aufwendig, so auch hier.

$$\begin{aligned}
\sigma^2 &= \int_0^2 \left(x^2 \cdot (-\frac{1}{2}x+1)\right) dx - \left(\frac{2}{3}\right)^2 = \\
&= \int_0^2 \left(-\frac{1}{2}x^3 + x^2\right) dx - \frac{4}{9} = \\
&= \left[-\frac{1}{2} \cdot \frac{x^4}{4} + \frac{x^3}{3}\right]_0^2 - \frac{4}{9} = \\
&= \left(-2 + \frac{8}{3}\right) - \frac{4}{9} = \\
&= \frac{2}{9}
\end{aligned}$$

Hinweis: Die Rechenregeln für Erwartungswert und Varianz die wir im Kapitel **??** für diskrete Verteilungen kennengelernt haben gelten für stetige Verteilungen genau so.

Übung 2.3.14 Berechnen Sie Erwartungswert und Varianz für die Verteilung aus Übung **??**.

Als nächstes werden wir uns einige spezielle stetige Verteilungen genauer ansehen.

2.3.3 Die Gleichverteilung

Sind alle Werte in einem Intervall $[a; b]$ gleich wahrscheinlich so spricht man von einer *Gleichverteilung*. Aufgrund des Aussehens der zugehörigen Dichtefunktion spricht man auch von der *Rechteckverteilung*.

Formel 2.3.15 *Dichtefunktion der Gleichverteilung*

$$f(x) = \begin{cases} \frac{1}{b-a} & \textit{für } a \leq x \leq b \\ 0 & \textit{sonst} \end{cases} \tag{2.33}$$

Formel 2.3.16 *Verteilungsfunktion der Gleichverteilung*

$$F(x) = \begin{cases} 0 & \textit{für } x < a \\ \frac{x-a}{b-a} & \textit{für } a \leq x \leq b \\ 1 & \textit{für } x > b \end{cases} \tag{2.34}$$

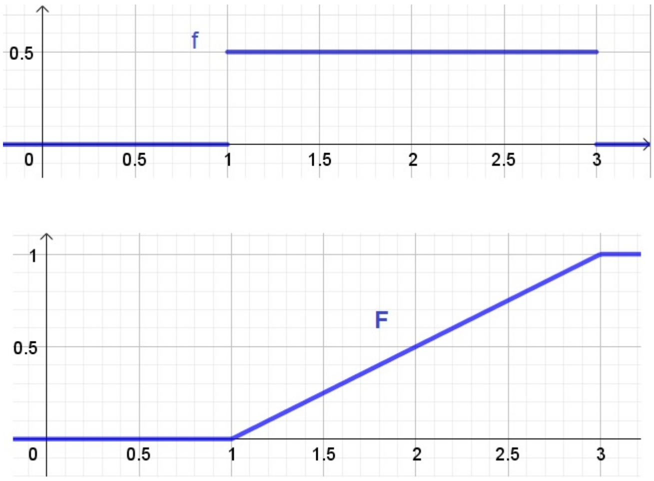


Abbildung 2.11: Graphen der Dichtefunktion und Verteilungsfunktion einer Gleichverteilung für $a = 1$ und $b = 3$

Formel 2.3.17 *Erwartungswert der Gleichverteilung*

Für eine Gleichverteilung im Intervall $[a; b]$ *lautet der Erwartungswert*

$$E(X) = \frac{a+b}{2} \tag{2.35}$$

Formel 2.3.18 *Varianz der Gleichverteilung*

Für eine Gleichverteilung im Intervall $[a; b]$ *lautet die Varianz*

$$\sigma^2 = V(X) = \frac{1}{12} \cdot (b-a)^2 \tag{2.36}$$

Übung 2.3.19
Leiten Sie die Verteilungsfunktion der Gleichverteilung allgemein her indem Sie das Integral $\int_a^x (\frac{1}{b-a})dt$ berechnen.
Leiten Sie außerdem Erwartungswert und Varianz her, verwenden Sei als Integrationsgrenzen die Variablen a und b.

Übung 2.3.20 *Rendite*
Die Rendite einer Investition in % sei gleichverteilt mit der Dichte

$$f(x) = \begin{cases} 0,125 & \textit{für } -2 \leq x \leq b \\ 0 & \textit{sonst} \end{cases}$$

Bestimmen Sie den Parameter b sowie Erwartungswert und Varianz.

2.3.4 Die Dreiecksverteilung

Namensgebend ist bei dieser Verteilung das Aussehen der Dichtefunktion das die Form eines Dreiecks hat. Sie wird bestimmt durch drei Werte a, b und c, wobei a und b die Ränder markieren und c die Position der Spitze.

Formel 2.3.21 *Dichtefunktion der Dreiecksverteilung*

$$f(x) = \begin{cases} \frac{2(x-a)}{(b-a)(c-a)} & \textit{für } a \leq x \leq c \\ \frac{2(b-x)}{(b-a)(b-c)} & \textit{für } c \leq x \leq b \\ 0 & \textit{sonst} \end{cases} \tag{2.37}$$

Formel 2.3.22 *Verteilungsfunktion der Dreiecksverteilung*

$$F(x) = \begin{cases} 0 & \textit{für } x < a \\ \frac{(x-a)^2}{(b-a)(c-a)} & \textit{für } a \leq x \leq c \\ 1 - \frac{(b-x)^2}{(b-a)(b-c)} & \textit{für } c \leq x \leq b \\ 1 & \textit{für } x > b \end{cases} \tag{2.38}$$

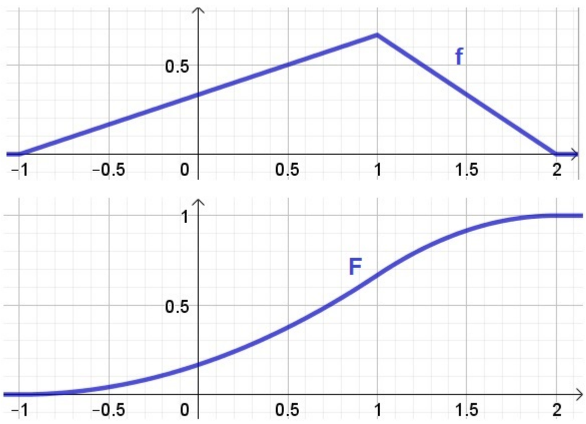


Abbildung 2.12: Graphen der Dichtefunktion und Verteilungsfunktion einer Dreiecksverteilung für $a = -1$ $c = 1$ und $b = 2$

Formel 2.3.23 *Erwartungswert der Dreiecksverteilung*

Für eine Dreiecksverteilung mit den Rändern a und b und der Spitze bei c lautet der Erwartungswert

$$E(X) = \frac{a+b+c}{3} \tag{2.39}$$

Formel 2.3.24 *Varianz der Dreiecksverteilung*

Für eine Dreiecksverteilung mit den Rändern a und b und der Spitze bei c lautet die Varianz

$$\sigma^2 = V(X) = \frac{(c-a)^2 + (b-a)^2 + (b-c)^2}{36} \tag{2.40}$$

Übung 2.3.25
Zeigen Sie, dass die Gesamtfläche unter dem Graphen einer Dreiecksverteilung mit den Werten $a = 10$, $c = 15$ und $b = 20$ den Wert 1=100% hat (geometrisch oder durch Integration).

Übung 2.3.26
i) Bestimmen Sie für die folgende Dreiecks-Verteilung die Parameter a, b und k.

$$f(x) = \begin{cases} k \cdot x + \frac{2}{3} & \textit{für } a \leq x \leq 0 \\ -\frac{1}{3} \cdot x + \frac{2}{3} & \textit{für } 0 \leq x \leq b \\ 0 & \textit{sonst} \end{cases}$$

ii) Bestimmen Sie Erwartungswert und Varianz.

2.3.5 Die Exponentialverteilung

Die Exponentialverteilung wird vor allem verwendet um die Länge von Zeitintervallen zu beschreiben, z.B. die Dauer bis zum Ausfall eines Bauteils (ohne Berücksichtigung von Alterserscheinungen) oder die Dauer einer Reparatur.

Formel 2.3.27 *Dichtefunktion der Exponentialverteilung*

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda x} & \textit{für } x \geq 0 \\ 0 & \textit{sonst} \end{cases} \tag{2.41}$$

Formel 2.3.28 *Verteilungsfunktion der Exponentialverteilung*

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \textit{für } x \geq 0 \\ 0 & \textit{sonst} \end{cases} \tag{2.42}$$

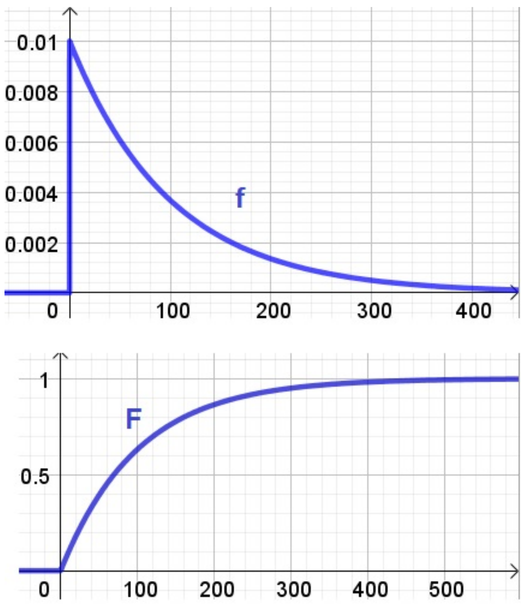


Abbildung 2.13: Graphen der Dichtefunktion und Verteilungsfunktion einer Exponentialverteilung für $\lambda = 0,01$

Formel 2.3.29 *Erwartungswert der Exponentialverteilung*

Für eine Exponentialverteilung mit Parameter λ lautet der Erwartungswert

$$E(X) = \frac{1}{\lambda} \tag{2.43}$$

Formel 2.3.30 *Varianz der Exponentialverteilung*

Für eine Exponentialverteilung mit Parameter λ lautet die Varianz

$$\sigma^2 = V(X) = \frac{1}{\lambda^2} \tag{2.44}$$

Hinweis zur Interpretation des Parameters λ:
Mit einer Exponentialverteilung wird typischerweise die Lebensdauer einer Maschine oder eines Bauteils beschrieben.
$\frac{1}{\lambda}$ ist dann also die *mittlere Lebensdauer*, und damit wäre λ die *mittlere Ausfallrate* bezogen auf die Zeiteinheit in der die Lebensdauer angegeben ist.
Wäre die mittlere Lebensdauer beispielsweise 3 Jahre dann heißt das, dass im Schnitt 1 Ausfall alle 3 Jahre und damit eine Ausfallrate von $\frac{1}{3} = 0,33 = 33\%$ pro Jahr zu verzeichnen ist.

Beispiel 2.3.31
In einem produzierenden Unternehmen kommen 3 identische Maschinen zum Einsatz die immer wieder ausgetauscht werden müssen. Aus Erfahrung weiß man, dass die mittlere Lebensdauer einer solchen Maschine 8 Jahre beträgt, die Restlebensdauer einer solchen Maschine kann mit einer Exponentialverteilung modelliert werden. Fällt eine Maschine aus muss sie sofort ersetzt werden was jeweils Kosten in Höhe von 5000 € verursacht.
i) Mit welchen mittleren jährlichen Kosten für zu ersetzende Maschinen muss das Unternehmen rechnen?
ii) Das Unternehmen gerät in eine kurzfristige finanzielle Schieflage und kann sich den Ausfall einer solchen Maschine im nächsten halben Jahr nicht leisten. Wie groß ist die Wahrscheinlichkeit, dass das Unternehmen überlebt?

Lösung:
i) Zunächst berechnen wir den Parameter λ. Für die mittlere Lebensdauer gilt:

$$\frac{1}{\lambda} = 8 \quad \Rightarrow \quad \lambda = \frac{1}{8} = 0,125 = 12,5\%$$

Da das Unternehmen 3 solche Maschinen besitzt muss folgender Betrag jährlich einkalkuliert werden:

$$0,125 \cdot 3 \cdot 5000 = 1875 \text{ €}$$

ii) Es sei *X: Lebensdauer einer Maschine*
Die Verteilungsfunktion lautet:

$$F(x) = \begin{cases} 1 - e^{-0,125 \cdot x} & \textit{für } x \geq 0 \\ 0 & \textit{sonst} \end{cases}$$

Die Wahrscheinlichkeit, dass eine (!) Maschine noch mindestens ein halbes Jahr funktioniert ergibt sich als

$$P(X \geq 0,5) = 1 - F(0,5) = 1 - (1 - e^{-0,125 \cdot 0,5}) = e^{-0,0625} = 0,9394 = 93,94\%$$

Es kann davon ausgegegangen werden, dass die Maschinen unabhängig voneinander ausfallen oder nicht, somit ergibt sich als Wahrscheinlichkeit, dass alle 3 Maschinen frühestens in einem halben Jahr ausfallen und ersetzt werden müssen

$$P(\text{"Das Unternehmen überlebt"}) = 0,9394^3 = 0,8290 = 82,9\%$$

Übung 2.3.32 *Zeit bis zum nächsten Auftrag*
Ein Unternehmen erhält durchschnittlich 8 Aufträge pro Quartal (ein Quartal wird zu 90 Tagen angenommen), die Zeit bis zum nächsten Auftrag kann mit einer Exponentialverteilung modelliert werden. Berechnen Sie die Wahrscheinlichkeit, dass die Zeit bis zum nächsten Auftrag
i) mindestens 10 Tage
ii) höchstens 15 Tage
beträgt.

2.3.6 Die Normalverteilung

Zum Abschluss der stetigen Verteilungen kommen wir zur Normalverteilung die als die wichtigste gelten darf. Dies liegt zum Einem daran, dass viele Verteilungen sich unter gewissen Umständen einer Normalverteilung immer mehr angleichen. Das gilt beispielsweise bei der Binomialverteilung für große Werte der Anzahl der Wiederholungen n, bei der Poisson-Verteilung für große Wert der mittleren Anzahl an Ereignissen λ. Das werden wir später noch genauer sehen wenn wir die Dichtefunktion der Normalverteilung betrachten.

Ein weiter Grund für die besondere Bedeutung der Normalverteilung kommt besteht in der Aussage des folgenden Satzes.

Satz 2.3.33 *Zentraler Grenzwertsatz*

Die Summe und der Mittelwert n identischer und unabhängig verteilter Zufallsvariablen nähern sich für große Werte von n immer mehr der Normalverteilung an.

oder etwas anders formuliert:

Entsteht eine Zufallsvariable aus einer Überlagerung vieler einzelner unabhängiger Zufallseffekte von denen keiner dominierend ist, so nähert sich diese Zufallsvariable immer mehr einer Normalverteilung an.

Da in der Praxis viele Vorgänge die dem Zufall unterliegen von vielen verschiedenen Einflussfaktoren abhängig sind führt das dazu, dass die Normalverteilung besonders häufig für stochastische Modelle geeignet ist.

Wie die Normalverteilung genau aussieht sehen wir uns jetzt an, dabei beginnen wir mit der sog. *Standardnormalverteilung.*

Formel 2.3.34 *Dichtefunktion der Standardnormalverteilung*

Eine Verteilung heißt standardnormalverteilt wenn ihre Dichtefunktion folgende Gestalt hat

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}} \tag{2.45}$$

Formel 2.3.35 *Verteilungsfunktion der Standardnormalverteilung*

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt \tag{2.46}$$

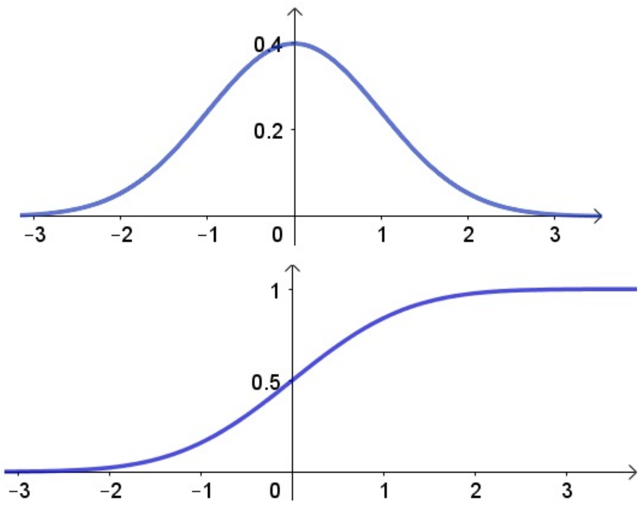


Abbildung 2.14: Graphen der Dichtefunktion und der Verteilungsfunktion der Standardnormalverteilung

Hinweise: Zur Abgrenzung zur anschließend folgenden allgemeinen Normalverteilung verwenden wir bei der Standardnormalverteilung die Bezeichnungen φ bzw. Φ ("Phi“) statt f und F sowie z statt x.
Für die Verteilungsfunktion der Standarnormaverteilung gibt es keine einfache Darstellung die ohne Integral auskommt, wir werden aber auch nicht mit diesen Formeln rechnen, sie sind eher der Vollständigkeit halber erwähnt.
Auch wenn konkrete Berechnungen heutzutage natürlich am Computer geschehen (was allerdings auf praktisch alle Inhalte dieser Vorlesung zutrifft) werden wir für ein echtes Verständnis nach der "altmodischen“ Variante vorgehen. D.h. wir werden $\Phi(z)$ für alle relevanten Werte $-3 \leq z \leq 3$ einer Tabelle im Anhang entnehmen.
In dieser Tabelle finden wir in der 1. Spalte und in der 1. Zeile die Werte durch deren Addition der betreffende z-Wert gebildet wird. Im Inneren der Tabelle finden wir dann den betreffen Wert der Verteilungsfunktion der Standardnormalverteilung.

A.2. Standardnormalverteilungstabelle

$\Phi_{o;1}(z) \rightarrow$

z	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,5279	0,53188	0,53586
0,1	0,53983	0,5438	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,6293	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,6591	0,66276	0,6664	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,7054	0,70884	0,71226	0,71566	0,71904	0,7224
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,7549

Abbildung 2.15: Ausschnitt der Tabelle zur Verteilungsfunktion der Standardnormalverteilung

Im Beispiel ist dargestellt: $\Phi(0,47) = P(Z \leq 0,47) = 0,68082 = 68,082\%$

Aufgrund der Symmetrie der Dichtefunktion zur y-Achse gilt für negative z-Werte mit $z < 0$:

Formel 2.3.36 *Symmetrieregel der Standardnormalverteilung*

$$\Phi(-z) = 1 - \Phi(z) \tag{2.47}$$

Das wird bei der Verwendung besagter Tabelle noch eine wichtige Rolle spielen.

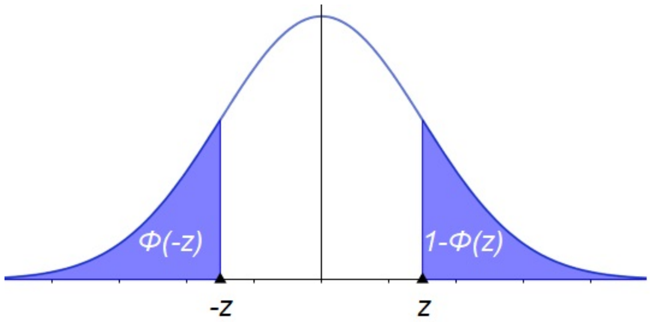


Abbildung 2.16: Graphische Darstellung der Symmetrieregel der Standardnormalverteilung

Formel 2.3.37 *Erwartungswert und Varianz der Standardnormalverteilung*

Für eine Standardnormalverteilung gilt

$$\mu = E(X) = 0 \quad und \quad \sigma^2 = V(X) = 1 \tag{2.48}$$

Die Besonderheit der standardisierten Form einer Normalverteilung, also der Standardnormalverteilung ist also, dass der Mittelwert genau 0 ist und die Standardabweichung genau 1. Bei einer allgemeinen Normalverteilung kommt jeder beliebige reelle Zahl als Erwartungswert und jede beliebige positive reelle Zahl als Standardabweichung in Frage. Die Grundsätzliche Form der Graphen der Dichtefunktion ändert sich dabei allerdings nicht.

Formel 2.3.38 *Dichtefunktion der allgemeinen Normalverteilung*

Die Dichtefunktion einer Normalverteilung mit Erwartungswert μ und Standardabweichung σ lautet

$$\varphi(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma^2} \cdot e^{-\frac{(x-\mu)^2}{2\sigma}} \tag{2.49}$$

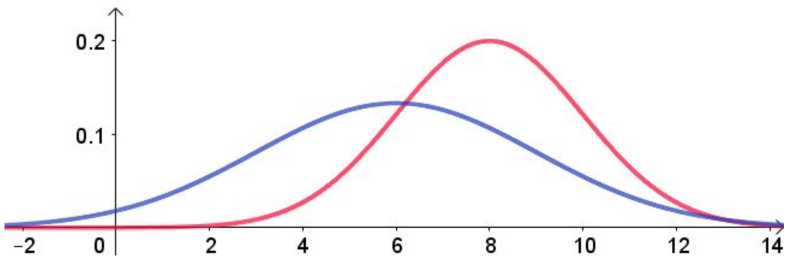


Abbildung 2.17: Graphen der Dichtefunktionen zweier Normalverteilungen mit $\mu = 6$ und $\sigma = 3$ (blau) bzw. $\mu = 8$ und $\sigma = 2$ (rot)

Hinweis: Der Erwartungswert hat einen Einfluss auf die Position des Graphen (je größer desto weiter rechts), die Standardabweichung auf die Form (je größer desto breiter).

Der Standardabweichung σ kommt bei der Normalverteilung eine etwas konkretere Bedeutung zu, es gilt grundsätzlich:

Formel 2.3.39 *68%-Regel*

$$P(\mu - \sigma \leq X \leq \mu + \sigma) = F(\mu + \sigma) - F(\mu - \sigma) \approx 0,68 = 68\% \tag{2.50}$$

Das heißt die Wahrscheinlichkeit, dass eine normalverteilte Zufallsvariable um höchstens $\pm\sigma$ vom Erwartungswert μ abweicht ist immer ca. 68%.

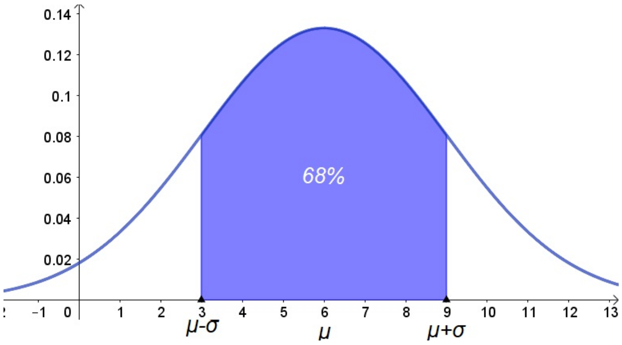


Abbildung 2.18: Graphische Darstellung der 68%-Regel für $\mu = 6$ und $\sigma = 3$

Um konkrete Berechnungen anstellen zu können benötigen wir noch folgende Eigenschaft normalverteilter Zufallsvariablen:
Eine standardnormalverteilte Zufallsgröße Z erhält man indem man eine normalverteilte Zufallsvariable X mit Erwartungswert μ und Standardabweichung σ mittels folgender Formel transformiert:

Formel 2.3.40 *Transformationsformel der Normalverteilung*

$$Z = \frac{X - \mu}{\sigma} \tag{2.51}$$

Und weiters:

Formel 2.3.41

Sei X eine normalverteilte Zufallsvariable mit Erwartungswert μ und Standardabweichung σ, dann gilt:

$$P(X \leq x) = F(x) = \Phi(\frac{x - \mu}{\sigma}) \tag{2.52}$$

Beispiel 2.3.42 *Normalverteilter Gewinn*
Sei X (in 1000 €) die normalverteilte Verteilungsfunktion des Gewinns einer Investition mir $\mu = 5$ und $\sigma = 7$. Mit welcher Wahrscheinlichkeit wird der Verlust nicht höher als 1000 € sein?[2]

Lösung:

Gesucht ist also folgende Wahrscheinlichkeit:

$$P(X \geq -1) = 1 - \Phi\left(\frac{-1-5}{7}\right) = 1 - \Phi\left(-0,8571\right) = 1 - (1 - \Phi\left(0,8571\right) \approx \Phi\left(0,86\right) = 0,8051 = 80,51\%$$

Man kann sich also zu ca. 80% sicher sein, dass der Verlust höchstens 1000 € beträgt.

[2] Negativer Gewinn wird als Verlust bezeichnet und umgekehrt.

Übung 2.3.43 *Rendite*
Eine Rendite in % sei normalverteilt mit $\mu = 1,5$ und $\sigma = 2$. Die Investitionssumme beträgt 100 000 €. Mit welcher Wahrscheinlichkeit beträgt der Verlust bei dieser Investition höchstens 2000 €?
Hinweis: Bestimmen Sie zunächst die Verteilung für den Verlust als Produkt aus Investitionssumme und Rendite und beachten Sie dabei die Rechenregeln für Erwartungswert und Varianz sowie folgenden Zusammenhang zwischen Gewinn G und Verlust V: $V = -G$

3 Risikomaße

Im folgenden Kapitel werden wir anhand verschiedener Beispiele einige praktische Anwendungen im Bereich Risikomanagement kennenlernen. Dabei werden die möglichen Verluste einer Investition als Zufallsvariablen betrachtet, so dass berechnet werden kann mit welcher Wahrscheinlichkeit ein gewisser Mindestverlust eintritt oder wie hoch dieser Verlust langfristig im Schnitt sein wird. Weiters werden wir exemplarisch Methoden kennenlernen wie verschiedene sich überlagernde Einflussgrößen gehandhabt werden, dabei geht es also um Methoden der sog. Risikoaggregation.

3.1 Quantile und der Value at Risk

Ein Risikomaß ist der sog. *Value at Risk* . Dieses bezeichnet jenen Wert des möglichen Verlustes der mit einer bestimmten, vorgegebenen Wahrscheinlichkeit nicht überschritten wird. In der Wahrscheinlichkeitsrechnung und Statistik wird dafür der Begriff der sog. *Quantile* benutzt. Auch hier unterscheiden wir zwischen diskreten und stetigen Verteilungen.

3.1.1 Quantile

Formel 3.1.1 *p-Quantile*

Sei X eine Zufallsvariable mit Verteilungsfunktion F. Dann nennt man eine Zahl Q_p das p-Quantil wenn gilt:

$$P(X \leq Q_p) = F(Q_p) \geq p \qquad \text{und} \qquad P(X \geq Q_p) \geq 1 - p \tag{3.1}$$

Die Wahrscheinlichkeit, dass die Zufallsvariable höchstens den Wert Q_p annimmt muss also mindestens p sein, und gleichzeitig muss die Wahrscheinlichkeit, dass die Zufallsvariable mindestens den Wert Q_p annimmt mindestens $1 - p$ sein.
Diese Definition wirkt auf den ersten Blick vielleicht etwas übertrieben genau, sie entspricht aber der Definition wie wir es aus der Statistik kennen. Betrachten wir z.B. folgenden kleinen Datensatz einer Stichprobe: 2, 2, 3, **4**, 4, 5, 7
Das 0,5-Quantil (Man sagt auch allgemein 50%-Quantil, das ist in diesem speziellen Fall der Median) wäre "der Wert in der Mitte", also $Q_{0,5} = 4$. $\frac{5}{7} \approx 71\%$ aller Werte sind kleiner oder gleich 4, $\frac{4}{7} \approx 57\%$ aller Werte sind größer oder gleich 4, das sind also jeweils mehr als die geforderten $p = 0,5 = 50\%$
Würden wir also einen Wert $Q_{0,5}$ suchen für den gilt: Genau die Hälfte der Werte ist kleiner oder gleich $Q_{0,5}$, so würden wir überhaupt keinen finden, was schon an der ungeraden Anzahl

$n = 7$ der Werte liegt.
Würden wir nur fordern, dass mindestens die Hälfte der Wert kleiner oder gleich $Q_{0,5}$ sein soll, so kämen alle Werte 4, 5 und 7 für $Q_{0,5}$ in Frage.

Trotzdem sehen wir anhand des Würfel-Beispiels, dass gewisse Quantile immer noch nicht eindeutig sind.

Beispiel 3.1.2 *Würfeln*
Ein fairer Würfel wird einmal geworfen mit *X: Augenzahl.* Die Verteilungsfunktion lautet:

$$F(x) = \begin{cases} 0 & \textit{für } x < 1 \\ \frac{1}{6} & \textit{für } 1 \leq x < 2 \\ \frac{2}{6} & \textit{für } 2 \leq x < 3 \\ \frac{3}{6} & \textit{für } 3 \leq x < 4 \\ \frac{4}{6} & \textit{für } 4 \leq x < 5 \\ \frac{5}{6} & \textit{für } 5 \leq x < 6 \\ 1 & \textit{für } x \geq 6 \end{cases}$$

i) Wir betrachten den Fall $p = 0,7$, wir suchen also eine Zahl $Q_{0,7}$ die mit mindestens 70%iger Wahrscheinlichkeit höchstens geworfen wird, aber auch mit mindestens 30%iger Wahrscheinlichkeit mindestens geworfen wird. Es gilt $\frac{4}{6} < 0,7 < \frac{5}{6}$. Die gesuchte Zahl ist also $Q_{0,7}$=5 weil:

$P(X \leq 5) = \frac{5}{6} \geq 0,7$ und $P(X \geq 5) = \frac{2}{6} \geq 1 - 0,7 = 0,3$.

Das selbe Ergebnis würden wir allerdings für jeden Wert $p \in (\frac{4}{6}; \frac{5}{6})$ erhalten!

ii) Nun betrachten wir $p = \frac{4}{6}$. In diesem Fall gilt, dass jede Zahl aus dem Bereich $[4; 5]$ als $\frac{4}{6}$-Quantil in Frage kommt. Es gilt nämlich:

$P(X \leq 4) = \frac{4}{6} \geq \frac{4}{6}$ und $P(X \geq 4) = \frac{3}{6} \geq 1 - \frac{4}{6} = \frac{2}{6}$

aber auch

$P(X \leq 5) = \frac{5}{6} \geq \frac{4}{6}$ und $P(X \geq 5) = \frac{2}{6} \geq 1 - \frac{4}{6} = \frac{2}{6}$

Hinweis: Im diskreten Fall kann eine Zahl Q_p das p-Quantil zu verschiedenen Werten p sein, es kann aber auch zu einem Wert p mehre p-Quantile Q_p geben!

Die Uneindeutigkeit liegt in der Tatsache begründet, dass die Graphen diskreter Verteilungen einerseits waagrechte Abschnitte und andererseits Sprungstellen besitzen. Im stetigen Fall vereinfacht sich das daher etwas.

Formel 3.1.3 *p-Quantile für stetige Verteilungen*

Sei X eine stetige Zufallsvariable mit Verteilungsfunktion F. Dann nennt man eine Zahl Q_p das
p-Quantil wenn gilt:

$$F(Q_p) = p \quad \Leftrightarrow \quad Q_p = F^{-1}(p) \tag{3.2}$$

wobei F^{-1} die Umkehrfunktion von F ist.

Beispiel 3.1.4 *Dauer zwischen zwei Aufträgen*

Ein Unternehmen erhält durchschnittlich 5 Aufträge pro Monat, die Zeit zwischen zwei Aufträgen kann mit einer Exponentialverteilung modelliert werden. Wie lange muss der Unternehmer nach Erhalt eines Auftrages mit einer Wahrscheinlichkeit von 95% höchstens warten bis der nächste Auftrag erteilt wird?

Lösung:

Gesucht ist also das 95%-Quantil. Für die durchschnittliche Rate der Aufträge pro Monat gilt $\lambda = 5$. Die Verteilungsfunktion lautet:

$$F(x) = \begin{cases} 1 - e^{-5x} & \text{\textit{für} } x \geq 0 \\ 0 & \textit{sonst} \end{cases}$$

Wir setzen

$$F(x) = 1 - e^{5x} = 0,95$$

und erhalten

$$\begin{aligned} 1 - e^{-5x} &= 0,95 \quad | -1 \\ -e^{-5x} &= -0,05 \quad | \cdot (-1) \\ e^{-5x} &= 0,05 \quad | \ln(...) \\ -5x &= \ln(0,05) \quad | \cdot (-\frac{1}{5}) \\ x &= 0,599 \approx 0,6 \end{aligned}$$

Der Unternehmer kann also zu 95% sicher sein, dass der nächste Auftrag innerhalb der nächsten 0,6 Monate erteilt wird.

3.1.2 Value at Risk

Im Risikomanagement können Quantile benutzt werden um den sog. *Value at Risk* zu bestimmen. Der Value at Risk VaR_p zum sog. *Konfidenzniveau* p bezeichnet jenen Verlust der höchstens mit einer Wahrscheinlichkeit $\alpha = 1 - p$ überschritten wird.[1]
Der Value at Risk entspricht damit dem p-Quantil der Verlustverteilung, wobei die Definition in leichter Abwandlung dazu so formuliert ist, dass Eindeutigkeit besteht.

Formel 3.1.5 *Der Value at Risk für diskrete Verteilungen*

Sei X: Verlust

Dann gilt:

$$VaR_p = \min\{x \in \mathbb{R} \mid P(X \leq x) \geq p\} = \min\{x \in \mathbb{R} \mid F(x) \geq p\} \tag{3.3}$$

Dabei bezeichnet "min" das Minimum der genannten Menge. Der Var_p ist also das kleinste aller p-Quantile.

Beispiel 3.1.6 *Diskreter Verlust*

Für den Verlust X einer Investition geltende folgende Wahrscheinlichkeiten:

Verlust x in €	-100	50	75
$P(X = x)$	70%	20%	10%

Bestimmen sie $VaR_{0{,}8}$, $VaR_{0{,}9}$,$VaR_{0{,}95}$ und $VaR_{0{,}99}$

Lösung:

Die Verteilungsfunktion lautet:

$$F(x) = \begin{cases} 0 & \textit{für } x < -100 \\ 0,7 & \textit{für } -100 \leq x < 50 \\ 0,9 & \textit{für } 50 \leq x < 75 \\ 1 & \textit{für } x \geq 75 \end{cases}$$

Es ergibt sich also: $VaR_{0{,}8} = 50$, $VaR_{0{,}9} = 50$, $VaR_{0{,}95} = 75 = VaR_{0{,}99}$

[1] Die Bezeichnungsweise für α und p ist in der Literatur nicht eindeutig und wird von manchen Autoren genau umgekehrt verwendet.

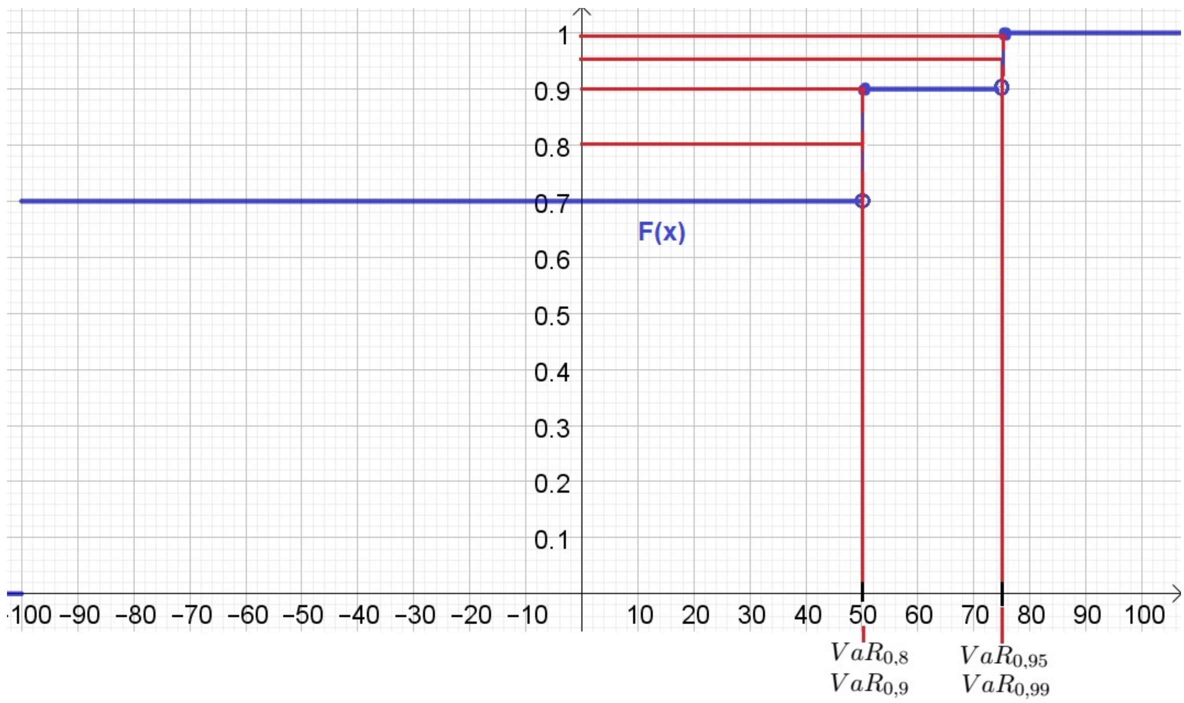


Abbildung 3.1: Graphische Darstellung der VaR für diskrete Verlustverteilung

Für stetige Verlustverteilungen ist der Value at Risk zum Konfidenzniveau p identisch mit dem p-Quantil, das dieses ja bereits eindeutig definiert war.

Formel 3.1.7 *Der Value at Risk für stetige Verteilungen*

Sei X eine stetige Verlustverteilungen mit Verteilungsfunktion F. Dann gilt für den Value at Risk zum Konfidenzniveau p:

$$VaR_p = F^{-1}(p) \tag{3.4}$$

wobei F^{-1} die Umkehrfunktion von F ist.

Beispiel 3.1.8 *Portfolio*
Der potentielle Verlust eines Portfolios sei dreiecksverteilt mit der den Parametern $a = -100$, $b = 100$ und $c = 0$ (alle Werte in 1000 €). Bestimmen Sie die Dichtefunktion, die Verteilungsfunktion und den Verlust der mit einer Wahrscheinlichkeit von 95% nicht überschritten wird.

Lösung:

Die Dichtefunktion lautet:

$$f(x) = \begin{cases} \frac{2(x+100)}{200 \cdot 100} = \frac{x}{10000} + \frac{1}{100} & \textit{für } -100 \leq x \leq 0 \\ \frac{2(100-x)}{200 \cdot 100} = -\frac{x}{10000} + \frac{1}{100} & \textit{für } 0 \leq x \leq 100 \\ 0 & \textit{sonst} \end{cases}$$

Zur Übung bestimmen wir die Verteilungsfunktion per Integral:

$i)$ $\quad -100 \leq x \leq 0: \quad \int_{-100}^{x} (\frac{t}{10000} + \frac{1}{100}) dt = \left[\frac{t^2}{20000} + \frac{t}{100}\right]_{-100}^{x} = \frac{x^2}{20000} + \frac{x}{100} + \frac{1}{2}$

$i)$ $\quad 0 \leq x \leq 100: \quad \frac{1}{2} + \int_0^x (-\frac{t}{10000} + \frac{1}{100}) dt = \frac{1}{2} + \left[-\frac{t^2}{20000} + \frac{t}{100}\right]_0^x = \frac{1}{2} - \frac{x^2}{20000} + \frac{x}{100}$

$$\Rightarrow \quad F(x) = \begin{cases} \frac{x^2}{20000} + \frac{x}{100} + \frac{1}{2} & \textit{für } -100 \leq x \leq 0 \\ \frac{1}{2} - \frac{x^2}{20000} + \frac{x}{100} & \textit{für } 0 \leq x \leq 100 \\ 0 & \textit{sonst} \end{cases}$$

Der Wert der rechts 5% der Gesamtfläche abgrenzt ist sicher im Bereich $0 \leq x \leq 100$ zu finden. Wir setzen also:

$\frac{1}{2} - \frac{x^2}{20000} + \frac{x}{100} = 0,95 \quad | - 0,95$

$-\frac{x^2}{20000} + \frac{x}{100} - \frac{45}{100} = 0 \quad | \cdot 20000$

$-x^2 + 200x - 9000 = 0$

Mit der quadratischen Lösungsformel erhalten wir:

$$x_{1,2} = \frac{-200 \pm \sqrt{200^2 - 4 \cdot 9000}}{-2} = \begin{cases} \mathbf{68{,}4} \\ (131,6) \end{cases}$$

Es gilt: $131,6 \notin [-100; 100]$, daher folgt: $VaR_{0,95} = 68,4$.

Der Verlust ist also mit 95%iger Sicherheit höchstens 68400 € bzw. ein höher Verlust als 68400 € tritt nur mit einer Wahrscheinlichkeit von 5% ein.

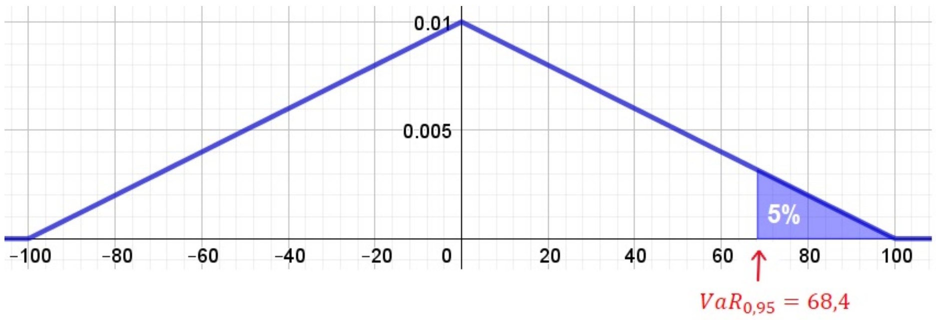


Abbildung 3.2: Graphische Darstellung des VaR für stetige Verlustverteilung anhand der Dichtefunktion

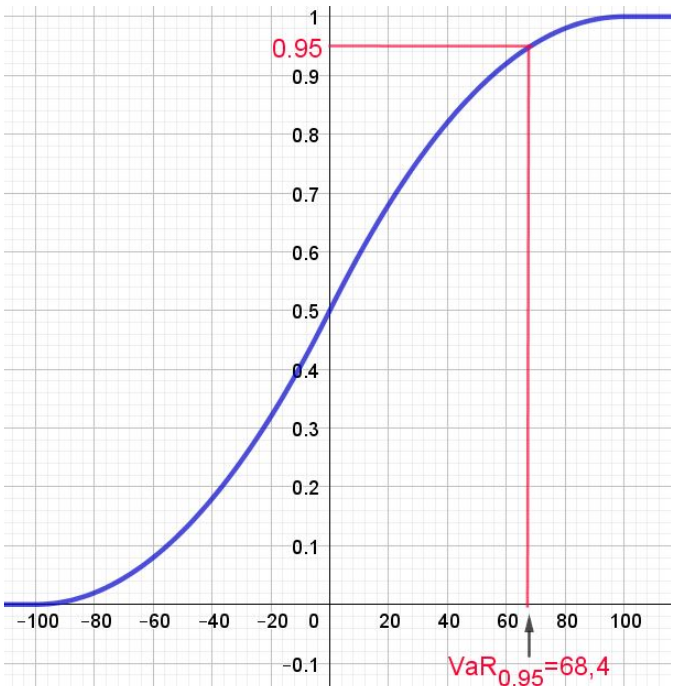


Abbildung 3.3: Graphische Darstellung des VaR für stetige Verlustverteilung anhand der Verteilungsfunktion

Übung 3.1.9 *Value At Risk einer Investition*
Eine Rendite in % sei normalverteilt mit $\mu = 2,5$ und $\sigma = 4$. Die Investitionssumme beträgt 500 000 €. Bestimmen Sie die Werte $VaR_{0,9}$ und $VaR_{0,95}$.
Hinweis: Bestimmen Sie zunächst die Verteilung für den Verlust als Produkt aus Investitionssumme und Rendite und beachten Sie dabei die Rechenregeln für Erwartungswert und Varianz sowie folgenden Zusammenhang zwischen Gewinn G und Verlust V: $V = -G$

3.2 Bedingte Erwartungswerte und Expected Shortfall

Der Value at Risk gibt an mit welcher Wahrscheinlichkeit ein gewisser Verlust nicht überschritten wird. Es sagt aber nicht darüber mit welcher Verlusthöhe für diesen Fall zu rechnen wäre. Das fällt insbesondere ins Gewicht wenn sehr hohe Verluste mit eher geringen Wahrscheinlichkeiten möglich sind und ist für die Einschätzung des Risikos natürlich von Bedeutung. Wir interessieren uns im Folgenden also für den mittleren Verlust für den Fall, dass der VaR doch überschritten wird. Das ist aus stochastischer Sicht ein sog. *bedingter Erwartungswert.*

3.2.1 Bedingter Erwartungswert

Als bedingten Erwartungswert bezeichnen wir den erwartetem Mittelwert einer Zufallsvariable der unter Annahme einer gewissen Voraussetzung eintritt. Auch hier müssen wir wieder zwischen diskreten und stetigen Verteilungen unterscheiden.

Formel 3.2.1 *Bedingter Erwartungswert für diskrete Verteilungen*

Sei X eine diskrete Zufallsvariable. Dann ist der bedingte Erwartungswert unter der Bedingung B

$$E(X|B) = \frac{\sum_{x_i \in B} (x_i \cdot P(X = x_i))}{P(B)} = \frac{\sum_{x_i \in B} (x_i \cdot P(X = x_i))}{\sum_{x_i \in B} (P(X = x_i))} \tag{3.5}$$

Wir bilden also den Erwartungswert für den Teilbereich der Werte die zur Bedingung B gehören und dividieren durch die Wahrscheinlichkeit der Bedingung.

Beispiel 3.2.2 *Würfeln*
Berechne den bedingten Erwartungswert $E(X|X \geq 5)$ für einen einmaligen Wurf mit einem fairen Würfel.

Lösung:

$$E(X|X \geq 5) = \frac{5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6}}{\frac{1}{6} + \frac{1}{6}} = 5,5$$

Bei allen Würfen die mindestens die Augenzahl 5 aufweisen erwartet man bei einem fairen Würfel gleich viele 5er und 6er und damit den Mittelwert 5,5.

In analoger Weise lautet der bedingte Erwartungswert für stetige Verteilungen

Formel 3.2.3 *Bedingter Erwartungswert für stetige Verteilungen*

Sei X eine diskrete Zufallsvariable mit Dichtefunktion. Dann ist der bedingte Erwartungswert unter der Bedingung B

$$E(X|B) = \frac{\int_{x \in B} (x \cdot f(x))\, dx}{P(B)} = \frac{\int_{x \in B} (x \cdot f(x))\, dx}{\int_{x \in B} (f(x))\, dx} \tag{3.6}$$

Beispiele hierzu werden wir uns im nächsten Unterkapitel ansehen.

3.2.2 Expected Shortfall und Tail Value at Risk

Mit dem Value at Risk haben wir bereits ein Risikomaß kennengelernt das den maximalen Verlust der mit einer gewissen Wahrscheinlichkeit p eintritt angibt. Etwas genauer ist der sog. *Expected Shortfall*. Er gibt den erwarten mittleren Verlust unter der Bedingung an, dass der Verlust doch größer als es der VaR angibt ausfällt, was mit einer Wahrscheinlichkeit von $1 - p$ der Fall ist. Es gibt mehrere Shortfall-Maße die für stetige Verteilungen des Risikos allerdings oft identisch sind. Die Unterscheidungen existieren aber weil es vor allem im diskreten Fall Unterschiede geben kann. Es sei an dieser Stelle erwähnt, dass die verwendeten Bezeichnungen in der Literatur voneinander abweichen und nicht einheitlich gebraucht werden.

Formel 3.2.4 *Expected Shortfall*

Sei X eine Verlustverteilung. Dann berechnet sich der Expected Shortfall zum Konfidenzniveau p nach der Formel

$$ES_{1-p} = E(X|X > VaR_p) \tag{3.7}$$

Beispiel 3.2.5 *Stromausfall*
In einer Fertigungsanlage werden drei identische Maschinen benutzt. Nach einem Stromausfall besteht für jede der drei Maschinen eine Wahrscheinlichkeit von 7% ausgetauscht werden zu müssen, was jeweils einen Verlust in Höhe von 3000 € verursacht. Berechnen Sie den Expected Shortfall zu einem Konfidenzniveau von $p = 90\%$.

Lösung:

Die Anzahl der ausgefallenen Maschinen ist binomialverteilt mit $n = 3$ und $p = 0,07$. Somit erhalten wir folgende Verteilung für den Verlust X:

$$\begin{aligned} P(X = 0) &= 0,93^3 = 0,80436 = 80,436\% \\ P(X = 3000) &= \binom{3}{1} \cdot 0,07 \cdot 0,93^2 = 0,18163 = 18,163\% \\ P(X = 6000) &= \binom{3}{2} \cdot 0,07^2 \cdot 0,93 = 0,01367 = 1,367\% \\ P(X = 9000) &= 0,07^3 = 0,00034 = 0,034\% \end{aligned}$$

Damit ergibt sich für den Value at Risk für $p = 0,9$:

$$VaR_{0,9} = 3000$$

und damit für den Expected Shortfall

$$ES_{0,1} = \frac{6000 \cdot 0,01367 + 9000 \cdot 0,00034}{0,01367 + 0,00034} = 6072,81 \tag{3.8}$$

Im Falle eines Stromausfalls ist also in $1,367\% + 0,034\% = 1,401\%$ aller Fälle mit einem mittleren Verlust von 6072,81 € zu rechnen.

Es fällt allerdings auf, dass sich zu einem Konfidenzniveau von z.B. $p = 0,95$ der selbe Wert für den Expected Shortfall ergeben würde, genau genommen sogar für alle $p \in (0,80436; 0,98599]$. Daher ist die tatsächliche Wahrscheinlichkeit der Bedingung $1,404\%$ auch nicht identisch mit dem vorher festgelegten Wert $1 - p = 10\%$. Das hängt damit zusammen, dass sich für verschiedene Konfidenzniveaus bei diskreten Verteilungen der selbe VaR ergeben kann wie wir bereits bei Beispiel **??** gesehen haben.

Daher definieren wir auch noch folgendes Shortfall-Maß bei dem sich auch bei diskreten Verteilungen unterschiedliche Werte zu unterschiedlichen Konfidenzniveaus ergeben.

Formel 3.2.6 *Tail Value at Risk für diskrete Verteilungen*

Sei X eine diskrete Verlustverteilung mit den Wahrscheinlichkeiten p_i für die verschiedenen Verlustwerte x_i und p ein vorgegebenes Konfidenzniveau. Weiters sei k der kleinste Wert für den gilt $\sum_{i=1}^{k} p_k > p$. Dann berechnet sich der Tail Value at Risk nach der Formel

$$TVaR_{1-p} =$$

$$\frac{1}{1-p} \cdot \Big[\Big((1-p) - (p_{k+1} + \cdots + p_n) \Big) \cdot P\left(X = x_k\right) + p_{k+1} \cdot P\left(X = x_{k+1}\right) + \cdots + p_n \cdot P\left(X = x_n\right) \Big] \tag{3.9}$$

Die Idee ist, dass der Expected Shortfall so ergänzt wird, dass die Wahrscheinlichkeit der Bedingung tatsächlich mit dem Wert $1 - p$ identisch wird indem man sich den fehlenden Anteil von dem Bereich mit dem größten nicht berücksichtigen Wert x_i "ausborgt".

Beispiel 3.2.7 *Stromausfall*
Wir betrachten die selbe Verteilung wie bei Beispiel **??**. Der Tail Value at Risk zu einem Konfidenzniveau von $p = 90\%$ ergibt sich als:

$$TVaR_{0,1} = \frac{1}{0,1} \cdot (0,08599 \cdot 3000 + 0,01367 \cdot 6000 + 0,00034 \cdot 9000) = 3430,50$$

Hinweis: Der TVaR berücksichtigt also im Gegensatz zum ES auch die fehlenden $8,599\%$ bei denen noch ein Verlust von 3000 € eintritt und kommt somit der Idee den mittleren Verlust in den 10% aller ungünstigsten Ausgänge auszudrücken näher.

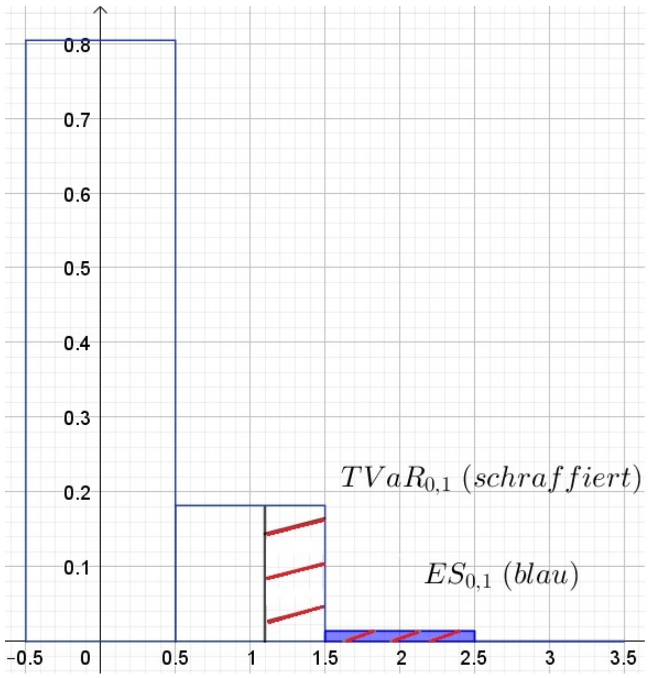


Abbildung 3.4: Graphische Darstellung der von TVaR und ES berücksichtigten Bereiche

Übung 3.2.8 *Expected Shortfall und Tail Value at Risk*
Der Verlust einer Investition sei folgendermaßen verteilt:
$P(X = -1000) = 0,5$; $P(X = 500) = 0,2$; $P(X = 650) = 0,18$; $P(X = 700) = 0,1$ und $P(X = 2000) = 0,02$. Bestimmen Sie die Werte $Var_{0,8}$, $ES_{0,2}$ und $TVaR_{0,2}$ sowie $Var_{0,9}$, $ES_{0,1}$ und $TVaR_{0,1}$.

Für stetige Verteilungen ist keine solche Unterscheidung nötig. Die genau Formel für den Expected Shortfall formulieren wir zunächst mit Hilfe der Dichtefunktion.

Formel 3.2.9 *Expected Shortfall für stetige Verlustverteilungen*

Sei X eine stetige Verlustverteilung mit Dichtefunktion f. Dann berechnet sich der Expected Shortfall zum Konfidenzniveau p nach der Formel

$$ES_{1-p} = \frac{1}{1-p} \cdot \int_{VaR_p}^{\infty} x \cdot f(x)dx \tag{3.10}$$

Mir Hilfe der Substitutionsmethode und dem Umstand, dass $F'(x) = f(x)$ lässt sich auch eine Formel entwickeln die die Verteilungsfunktion benutzt. Wir setzten

$$\begin{aligned} u(x) = F(x) \quad &\Rightarrow \quad i) \quad F'(x) = f(x) = \frac{du}{dx} \\ &\Rightarrow \quad ii) \quad x = F^{-1}(u) \quad \text{(wobei } F^{-1} \text{ die Umkehrfunktion von F ist)} \end{aligned}$$

Für die Integrationsgrenzen folgt

$$u(Var_p) = F(VaR_p) = p \quad und \quad u(\infty) = F(\infty) = 1$$

Einsetzen liefert

$$ES_{1-p} = \frac{1}{1-p} \cdot \int_p^1 F^{-1}(u) \cdot \frac{du}{dx} \cdot dx = \frac{1}{1-p} \cdot \int_p^1 F^{-1}(u) du$$

Also:

Formel 3.2.10 *Expected Shortfall für stetige Verlustverteilungen*

Sei X eine stetige Verlustverteilung mit Verteilungsfunktion F. Dann berechnet sich der Expected Shortfall zum Konfidenzniveau p nach der Formel

$$ES_{1-p} = \frac{1}{1-p} \cdot \int_p^1 F^{-1}(u) du \tag{3.11}$$

Beispiel 3.2.11 *Portfolio*
Wir betrachten nochmals das Portfolio von Beispiel **??** mit folgender Dichtefunktion

$$f(x) = \begin{cases} \frac{x}{10000} + \frac{1}{100} & f\ddot{u}r\ -100 \leq x \leq 0 \\ -\frac{x}{10000} + \frac{1}{100} & f\ddot{u}r\ 0 \leq x \leq 100 \\ 0 & sonst \end{cases}$$

Bestimmen Sie den Expected Shortfall zu einem Konfidenzniveau von $p = 0,95$

Lösung:
Wir wissen bereits, dass $VaR_{0,95} = 68,4$. Der relevante Bereich ist wieder $0 \leq x \leq 100$. Es gilt:

$$\begin{aligned}
ES_{0,05} &= \frac{1}{1-0,95} \cdot \int_{68,4}^{100} x \cdot (-\frac{x}{10000} + \frac{1}{100}) dx = \\
&= \frac{1}{0,05} \cdot \int_{68,4}^{100} (-\frac{x^2}{10000} + \frac{x}{100}) dx = \\
&= \left[-\frac{x^3}{30000} + \frac{x^2}{200} \right]_{68,4}^{100} = \\
&= \left(-\frac{100^3}{30000} + \frac{100^2}{200} \right) - \left(-\frac{68,4^3}{30000} + \frac{68,4^2}{200} \right) = \\
&= 78,8
\end{aligned}$$

In den 5% der ungünstigsten Ausgänge ist also mit einem mittleren Verlust von 78800 € zu rechnen.

Formel 3.2.12 *Expected Shortfall für normalverteilte Verluste*

Sei V eine normalverteilte Zufallsvariable des Verlustes mit Erwartungswert μ und Standardabweichung σ. Dann ist

$$ES_{1-p} = \mu + \frac{\sigma}{1-p} \cdot \varphi(z_p) \tag{3.12}$$

wobei $\varphi(x)$ die Dichtefunktion der Standardnormalverteilung und z_p das p-Quantil der Standardnormalverteilung sind.

Hinweis: Die Herleitung kann für Interessierte im Tutorium vorgeführt werden.

Beispiel 3.2.13 *Aktienpaket*
Der Gewinn eines Aktienpaketes sei normalverteilt mit $\mu_X = 2$ und $\sigma_X^2 = V(X) = 9$ (in 1000 €). Berechnen Sie den Expected Shortfall zu einem Konfidenzniveau von 99%.

Lösung:

Sei X: Verteilung des Gewinns

Verlust ist negativer Gewinn, daher gilt für die Verlustverteilung $Y = -X$.
Mit den Rechenregeln für Erwartungswert und Varianz folgt für die Verlustverteilung Y:

$$E(Y) = E(-X) = -E(X) = -2 = \mu_Y$$
$$V(Y) = V(-X) = (-1)^2 \cdot V(X) = 9 \quad \Rightarrow \quad \sigma_Y = \sqrt{9} = 3$$

Aus der Tabelle der Verteilungsfunktion der Standardnormalverteilung entnehmen wir

$$z_{0,99} = 2,33$$

Außerdem berechnen wir

$$\varphi(2,33) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{2,33^2}{2}} = 0,02643$$

und erhalten

$$ES_{0,01} = -2 + \frac{3}{0,01} \cdot 0,02643 \approx 5,92$$

In den ungünstigsten 1% der Fälle ist also mit einem mittleren Verlust von ca. 5920 € zu rechnen.

Zum Vergleich: Der Value at Risk zu $p = 0,99$ wäre

$$VaR_{0,99} = \mu + \sigma z_{0,99} = -2 + 3 \cdot 2,33 = 4,99$$

Hinweis: Es gilt grundsätzlich

$$ES_{1-p} > VaR_p$$

Übung 3.2.14
Die Rendite einer Investition sei gleichverteilt innerhalb der Grenzen -2 und 2 (in %) bei einem Investitionsvolumen von 50000 €.
Bestimmen Sie denjenigen Verlust der mit einer Wahrscheinlichkeit von 90% nicht überschritten wird sowie den durchschnittlichen Verlust für den unwahrscheinlichen Fall, dass dies doch eintritt.
Hinweis: Bestimmen Sie zunächst die Verteilung für den Verlust als Produkt aus Investitionssumme und Rendite und beachten Sie dabei die Rechenregeln für Erwartungswert und Varianz sowie folgenden Zusammenhang zwischen Gewinn G und Verlust V: $V = -G$

Übung 3.2.15
Die Rendite einer Investition sei normalverteilt mit einem Erwartungswert von 0 und einer Standardabweichung von 2,5 (in %) bei einem Investitionsvolumen von 750000 €.
Bestimmen Sie denjenigen Verlust der mit einer Wahrscheinlichkeit von 95% nicht überschritten wird sowie den durchschnittlichen Verlust für den unwahrscheinlichen Fall, dass dies doch eintritt.
Hinweis: Bestimmen Sie zunächst die Verteilung für den Verlust als Produkt aus Investitionssumme und Rendite und beachten Sie dabei die Rechenregeln für Erwartungswert und Varianz sowie folgenden Zusammenhang zwischen Gewinn G und Verlust V: $V = -G$

4 Methoden der Risikoaggregation

Risikoaggregation bedeutet mehrere Einzelrisiken zu einem Gesamtrisiko zusammenzufassen. Dabei ist zu berücksichtigen, dass die verschiedenen Einzelrisiken sich gegenseitig beeinflussen oder gemeinsame Abhängigkeiten aufweisen können, wie z.B. die Kurse verschiedener Aktien aus ähnlichen Branchen.
Grundsätzlich gibt es die Möglichkeit das Gesamtrisiko analytisch, sprich rechnerisch abzuschätzen (was aber nicht immer möglich ist), oder durch Simulation. Die jeweiligen Grundprinzipien werden wir im folgenden Kapitel kennenlernen.

4.1 Der Varianz-Kovarianz-Ansatz

Mit Hilfe des *Varianz-Kovarianz-Ansatzes* (auch Delta-Normal-Ansatz genannt) können die Risiokomaße VaR und ES von Gesamtrisiken berechnet werden unter Berücksichtigung linearer Zusammenhänge der Einzelrisiken, wobei man bei diesen jeweils von einer Normalverteilung ausgeht.
Vorher benötigen wir aber noch einige wenige Grundlagen der Vektor- und Matrizenrechnung.

Formel 4.1.1 *Eine $m \times n$ Matrix A mit m Zeilen und n Spalten ist eine rechteckige Anordnung von Elementen in der folgenden Form*

$$A_{m\times n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{4.1}$$

Eine $n \times 1$ Matrix nennen wir auch einen Spaltenvektor $\vec{v}$

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \tag{4.2}$$

und den zugehörigen Zeilenvektor bezeichnen wir als $\vec{v}^\top$ ("v transponiert)"

$$\vec{v}^\top = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} \tag{4.3}$$

Formel 4.1.2 *Multiplikation zweier Matrizen*
Aus zwei Matrizen A und B lässt sich das Produkt $C = A \cdot B$ bilden wenn die Anzahl der Zeilen von A mit der Anzahl der Spalten von B übereinstimmt. Ist diese Zeilen- bzw. Spaltenanzahl gleich n, dann ergibt sich das Element der i-ten Zeile und der j-ten Spalte von C nach der Formel

$$c_{ij} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj} \tag{4.4}$$

Beispiel 4.1.3 *Multiplikation zweier Matrizen*

$$\begin{pmatrix} 2 & 4 \\ -1 & 3 \\ -2 & 6 \end{pmatrix} \cdot \begin{pmatrix} 5 & -3 \\ -4 & 7 \end{pmatrix} = \begin{pmatrix} -6 & 22 \\ -17 & 24 \\ -34 & 48 \end{pmatrix}$$

wobei $c_{11} = 2 \cdot 5 + 4 \cdot (-4) = -6$, $c_{12} = 2 \cdot (-3) + 4 \cdot 7 = 22$, ...

An dieser Stelle führen wir noch die sog. *Kovarianz-Matrix* ein.

Formel 4.1.4 *Die Kovarianz-Matrix*
Seien $X_1, ..., X_n$ n verschiedene Wahrscheinlichkeitsverteilungen mit den Varianzen σ^2 und den Kovarianzen σ_{ij}. Dann lautet die sog. Kovarianz-Matrix Σ

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{pmatrix} \tag{4.5}$$

Nun können wir den Varianz-Kovarianz-Ansatz definieren.

Formel 4.1.5 *Der Varianz-Kovarianz-Ansatz*

Seien $X_1, ..., X_n$ n verschiedene Verlustverteilungen eines Portfolios die alle normalverteilt sind mit den Erwartungswerten $\mu_1, ..., \mu_n$ und den Varianzen $\sigma_1, ...\sigma_n$.
Weiters sind die sog. Gewichte $w_1, ..., w_n$ die jeweiligen Anteile der Zufallsvariablen am Gesamtwert des Portfolios. Wir definieren folgende Vektoren:

$$\vec{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \quad \textit{und} \quad \vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \tag{4.6}$$

Dann berechnen sich der Value at Risk und der Expected Shortfall wie folgt:

$$Var_p = \vec{w}^\top \cdot \vec{\mu} + \sqrt{\vec{w}^\top \cdot \Sigma \cdot \vec{w}} \cdot z_p \tag{4.7}$$

$$ES_{1-p} = \vec{w}^\top \cdot \vec{\mu} + \sqrt{\vec{w}^\top \cdot \Sigma \cdot \vec{w}} \cdot \frac{\varphi(z_p)}{1-p} \tag{4.8}$$

wobei z_p das p-Quantil der Standardnormalverteilung ist.

Hinweis: In der Praxis müssen die enthaltenen Größen in der Regel statistisch ermittelt werden. Das geschieht mit dem arithmetischen Mittel für die Erwartungswerte und den entsprechenden statistischen Formeln für die empirische Varianz und die empirische Kovarianz.

Beispiel 4.1.6 *Portfolio*
Ein Aktienportfolio besteht aus 10 Aktien der Sorte A im Wert von jeweils 1500 € und 15 Aktien der Sorte B im Wert von jeweils 2000 €. In den vergangenen 30 Tagen wurden die täglichen Gewinne und Verluste beobachtet und damit folgende Werte für die Verlustverteilungen X_A und Y_B der 1. und 2. Aktie statistisch ermittelt (pro Aktie):

$\overline{x}_A = -10$, $s^2_{x_A} = 80$, $\overline{y}_B = 10$, $s^2_{y_B} = 40$ und $s_{x_A y_B} = s_{y_B x_A} = 38$
(Dabei bezeichnen $s^2_{x_A}$ und $s^2_{y_B}$ die empirischen Varianzen und $s_{x_A y_B}$ bzw. $s_{y_B x_A}$ die empirischen Kovarianzen.)

X_A und Y_B können als normalverteilt angenommen werden. Berechnen sie den VaR und den ES zu $p = 0,95$

Lösung:

Wir berechnen zunächst die Gewichte erhalten

$$w_x = \frac{10 \cdot 1500}{10 \cdot 1500 + 15 \cdot 2000} = 0,33 \quad \Rightarrow \quad w_y = 0,67$$

Gemäß den Rechenregeln für Erwartungswert, Varianz und Kovarianz erhalten wir für die Gesamtverlustverteilungen X aus Aktie A bzw. Y aus Aktie B:

$$\begin{aligned}
\mu_X &= -10 \cdot 10 = -100 \\
\mu_Y &= 10 \cdot 15 = 150 \\
\sigma^2_X &= 80 \cdot 10^2 = 8000 \\
\sigma^2_Y &= 40 \cdot 15^2 = 9000 \\
\sigma_{XY} &= 38 \cdot 10 \cdot 15 = 5700
\end{aligned}$$

Aus der Tabelle der Verteilungsfunktion der Standardnormalverteilung entnehmen wir $z_{0,95} = 1,65$ und erhalten damit:

$$VaR_{0,95} = \begin{pmatrix} 0,33 \\ 0,67 \end{pmatrix}^\top \cdot \begin{pmatrix} -100 \\ 150 \end{pmatrix} + \sqrt{\begin{pmatrix} 0,33 \\ 0,67 \end{pmatrix}^\top \cdot \begin{pmatrix} 8000 & 5700 \\ 5700 & 9000 \end{pmatrix} \cdot \begin{pmatrix} 0,33 \\ 0,67 \end{pmatrix}} \cdot 1,645 =$$
$$= 208,39$$

Für den ES berechnen wir zunächst:

$$\varphi(1,645) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1,645^2}{2}} = 0,1031$$

und weiters:

$$ES_{0,05} = \begin{pmatrix} 0,33 \\ 0,67 \end{pmatrix}^\top \cdot \begin{pmatrix} -100 \\ 150 \end{pmatrix} + \sqrt{\begin{pmatrix} 0,33 \\ 0,67 \end{pmatrix}^\top \cdot \begin{pmatrix} 8000 & 5700 \\ 5700 & 9000 \end{pmatrix} \cdot \begin{pmatrix} 0,33 \\ 0,67 \end{pmatrix}} \cdot \frac{0,1031}{0,05} =$$
$$= 244,31$$

Übung 4.1.7
Ein Aktienportfolio besteht aus 30 Aktien der Sorte A im Wert von jeweils 1200 € und 20 Aktien der Sorte B im Wert von jeweils 1500 €. In den vergangenen 30 Tagen wurden die täglichen Renditen beobachtet und damit folgende Werte für deren Verteilungen R_A und R_R statistisch ermittelt (Werte jeweils in Prozent):
$\overline{r}_A = 1,2$, $s^2_{R_A} = 1,4$, $\overline{r}_B = -1$, $s^2_{R_B} = 1,3$ und $s_{R_A R_B} = s_{R_B R_A} = 1,2$.
Die Renditen können als normalverteilt angenommen werden. Berechnen sie den VaR und den ES zu $p = 0,9$.
Hinweis: Berechnen Sie zunächst die Parameter der Verteilungen der Gesamtverluste X und Y durch Aktie A bzw. Aktie B und beachten Sie, dass für den Verlust V und den Gewinn G gilt: $V = -G$

4.2 Monte-Carlo-Simulation

Mit dem Monte-Carlo Verfahren werden grundsätzlich Probleme gelöst die auf analytischem Weg nicht oder nur sehr schwer zu lösen sind. Dabei nähert man sich der Lösung mit Hilfe von Zufallszahlen an und nützt das Gesetz der großen Zahl aus das im Wesentlich besagt, dass sich

der Mittelwert einer Stichprobe umso mehr dem Erwartungswert der Zufallsgröße annähert je größer der Stichprobenumfang ist.
Mit Monte-Carlo-Simulationen können z.B. nicht oder schwer lösbare Integrale bestimmt werden oder Zufallszahlen unbekannter Verteilung generiert werden. Dabei ist zu beachten, dass je nach benutzter Software evtl. nur Zufallszahlen im Bereich $[0; 1]$ zur Verfügung stehen.

4.2.1 Integration mit Monte-Carlo-Simulation

Um die Methode kennenzulernen und zu verstehen und wenden wir uns zunächst der Abschätzung von Integralen an.

Formel 4.2.1 *Integration mit Monte-Carlo-Simulation*

Sei f(x) eine Funktion und das bestimmte Integral $\int_a^b f(x)dx$ *gesucht. Dann lässt sich das Ergebnis mit Hilfe von n Zufallszahlen* $u_1, ..., u_n \in [0; 1]$ *folgendermaßen abschätzen.*

i) Erzeuge Zufallszahlen im Bereich $r_i \in [a; b]$ *durch die Transformation* $r_i = u_i \cdot (b-a)+a$
ii) Die Abschätzung des Integrals lautet dann:

$$\int_a^b f(x)dx = \sum_{i=1}^{n} \frac{b-a}{n} \cdot f(r_i) \tag{4.9}$$

Beispiel 4.2.2 *Value at Risk*
Wir berechnen nochmals den Expected Shortfall aus Beispiel **??**. Zu berechnen war das folgende Integral:

$$ES_{0,05} = \frac{1}{0,05} \cdot \int_{68,4}^{100} \left(-\frac{x^2}{10000} + \frac{x}{100}\right)dx$$

Die Realisierung mit EXCEL mit n=1000 Zufallszahlen sieht folgendermaßen aus:

	A	B	C	D	E
1	**Zufallszahlen**	**transformierte Zufallszahlen**	**Summanden**	**Summe=Integral**	**Endergebnis**
2	**ZUFALLSZAHL()**	**Ai*31,6+68,4**	**31,6/1000*(-Bi^2/10000+Bi/100)**	**SUMME(C3:C1002)**	**D3*0,05**
3	0,9407756848504	98,1285116412736	0,000580322520338703	3,93444592670863	78,688918534
4	0,746923252731	92,0027747862981	0,0023250234363948		
5	0,6992898594166	90,4975595575634	0,00271743463682588		
6	0,6605136334799	89,2722308179633	0,00302630636159473		
7	0,2273510236157	75,5842923462561	0,00583160299304651		
8	0,0700881235907	70,6147847054654	0,0065570968588895		
9	0,8512071443044	95,2981457600191	0,00141592645062483		
10	0,5997403859932	87,3517961973835	0,00349130489382004		

Abbildung 4.1: Abschätzung eines bestimmten Integrals mit Monte-Carlo-Simulation

Hinweis: Genauere Werte lassen sich entweder durch noch mehr Zufallszahlen erreichen, oder indem man mehrere Werte simuliert und davon den Mittelwert bildet.

4.2.2 Risikoaggregation mit Monte-Carlo-Simulation

In der Risikoaggregation hat man es oft mit mehreren, sich überlagernden Einzelrisiken zu tun. Wenn keine Normalverteilungsannahme getroffen werden kann ist in der Regel nicht klar welcher Verteilung das Gesamtrisiko folgt. Dann kann mit Hilfe der Monte-Carlo-Methode eine simulierte Stichprobe erzeugt werden anhand derer VaR und ES berechnet wird.
Dabei wird den einzelnen Teilrisiken jeweils ein gewisse Verteilung unterstellt, sodass es zunächst nötig ist Zufallszahlen zu erzeugen die einer gewissen Verteilung folgen.

Für das Erzeugen diskret verteilter Zufallszahlen gehen wir folgendermaßen vor:

Formel 4.2.3 *Zufallszahlen einer diskreten Verteilung*

Gegeben sei eine diskrete Verteilung

x_i	x_1	x_2	...	x_{n-1}	x_n
$P(X = x_i)$	p_1	p_2	...	p_{n-1}	p_n

Dann erzeugt folgende Funktion auf der Definitionsmenge $D = [0; 1]$ Zufallszahlen gemäß dieser Verteilung:

$$f(x) = \begin{cases} x_1 & \text{für } x \leq p_1 \\ x_2 & \text{für } p_1 < x \leq p_1 + p_2 \\ \cdots & \cdots \\ x_{n-1} & \text{für } p_1 + \cdots + p_{n-2} < x \leq p_1 + \cdots + p_{n-1} \\ x_n & \text{sonst} \end{cases}$$

Beispiel 4.2.4 *Zufallszahlen einer diskreten Verteilung*

Gegeben sei folgende diskrete Verlustverteilung:

x_i	-2	-1	1
$P(X = x_i)$	0,5	0,3	0,2

Wenn wir nun Zufallszahlen aus dem Intervall $[0; 1]$ in folgende Funktion einsetzen, dann erhalten wir Zufallszahlen die der obigen Verteilung folgen.

$$f(x) = \begin{cases} -2 & \text{für } x \leq 0,5 \\ -1 & \text{für } 0,5 < x \leq 0,8 \\ 1 & \text{sonst} \end{cases}$$

Die Realisierung mit EXCEL sieht folgendermaßen aus:

	A	B
1	Zufallszahl	Zufallszahlen in Verteilung
2	ZUFALLSZAHL()	WENN(A3<0,5;-2;WENN(A3<0,8;-1;1))
3	0,8737198959474	1
4	0,618177560326	-1
5	0,8113413072995	1
6	0,0017585109886	-2
7	0,9884834602351	1
8	0,6932429021565	-1
9	0,2776123495841	-2
10	0,2479243758643	-2

Abbildung 4.2: Zufallszahlen einer diskreten Verteilung

Für stetige Verteilungen gilt:

Formel 4.2.5 *Zufallszahlen einer stetigen Verteilung*

Sei U eine auf dem Intervall [0;1] gleichverteilte Zufallsvariable mit der Verteilungsfunktion $F_U(x) = x$ für $x \in [0;1]$ und $F(x)$ eine beliebige Verteilungsfunktion. Dann ist die Zufallsvariable

$$X = F^{-1}(U) \tag{4.10}$$

verteilt nach der zu F gehörenden Verteilung.

Begründung:

$$P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F_U(F(x)) = F(x)$$

Beispiel 4.2.6 *Exponentialverteilte Zufallszahlen*

Gegeben sie eine Exponentialverteilung mit Verteilungsfunktion

$$F(x) = \begin{cases} 1 - e^{-2x} & \textit{für } x \geq 0 \\ 0 & \textit{sonst} \end{cases}$$

Wir berechnen zunächst die Umkehrfunktion $F^{-1}(x)$

$$
\begin{aligned}
x &= 1 - e^{-2y} \quad &&| + e^{-2y} - x \\
e^{-2y} &= 1 - x &&| \ln(\cdots) \\
-2y &= \ln(1-x) &&| : (-2) \\
y &= -\frac{\ln(1-x)}{2} \\
F^{-1}(x) &= -\frac{\ln(1-x)}{2}
\end{aligned}
$$

und erhalten folgende Realisierung mit Excel:

	A	B
1	Zufallszahlen	Zufallszahlen in Verteilung
2	ZUFALLSZAHL()	-LN((1-A3)/2)
3	0,490040984357	1,36657209853714
4	0,7037502404935	1,90969957888534
5	0,471154580941	1,33020628388544
6	0,318614353649	1,0767740194755
7	0,761156707061	2,12509480074096
8	0,2792249545684	1,02057537450259
9	0,4373068481013	1,26816800325202
10	0,3372692657013	1,1045336841298

Abbildung 4.3: Zufallszahlen einer stetigen Verteilung

Übung 4.2.7 *Zufallszahlen* Die Verteilungsfunktion aus Beispiel **??** lässt sich im Bereich $x \in [0; 2]$ auch folgendermaßen darstellen:

$$F(x) = -\frac{1}{4} \cdot (x-2)^2 + 1$$

Berechnen Sie die Umkehrfunktion F^{-1} die also gleichverteilte Zufallszahlen aus dem Intervall $(0; 1)$ in Zufallszahlen der Verteilung von F umwandelt.

Die Umkehrung der Verteilungsfunktion einer Normalverteilung ist nicht so einfach zu bewerkstelligen, wegen ihrer besonderen Bedeutung wird an dieser Stelle noch ein Verfahren zu Erzeugung normalverteilter Zufallszahlen vorgestellt.

Formel 4.2.8 *Algorithmus zur Erzeugung normalverteilter Zufallszahlen mit μ und σ*

1) *Erzeuge Realisierungen zweier auf [0;1] gleichverteilter Zufallsvariablen U_1 und U_2*

2) *Setze*

$$Z = \sqrt{-2 \cdot \ln(U_1)} \cdot \sin(2\pi \cdot U_2)$$

3) *Setze*

$$X = \mu + \sigma \cdot Z$$

Beispiel 4.2.9 *Erzeugung normalverteilter Zufallszahlen*

Es sollen Zufallszahlen nach einer Normalverteilung mit $\mu = -2$ und $\sigma = 3$ erzeugt werden.

Die Realisierung mit EXCEL sieht folgendermaßen aus:

	A	B	C	D
1	Zufallszahl 1	Zufallszahl 2	Zufallszahlen in Z-Verteilung	Zufallszahlen in Verteilung
2	ZUFALLSZAHL()	ZUFALLSZAHL()	WURZEL(-2*LN(A3))*SIN(2*PI()*B3)	-2+3*C3
3	0,3573775041472	0,6503141553474	-1,16223508214723	-5,48670524644169
4	0,7297491362642	0,7364236422665	-0,790907699399826	-4,37272309819948
5	0,0027466019809	0,2189906281494	3,36937215988146	8,10811647964438
6	0,9537034440607	0,0984469920962	0,178542320373172	-1,46437303888049
7	0,7340436588641	0,7384475640364	-0,78429663423863	-4,35288990271589
8	0,6320401434547	0,9987267185584	-0,00766350540950472	-2,02299051622851
9	0,0676036341043	0,2842368802869	2,26774561588923	4,80323684766769
10	0,7327929898207	0,5369951080734	-0,181646047339058	-2,54493814201717

Abbildung 4.4: Zufallszahlen einer Normalverteilung mit $\mu = -2$ und $\sigma = 3$

Beispiel 4.2.10 *Risikoaggregation mit Monte-Carlo-Simulation*

Als Beispiel betrachten wir nun die Summe der beiden Verlustverteilungen aus den Beispielen **??** und **??**. Es wurden $n = 1000$ Zufallszahlen generiert. Wir erhalten sozusagen eine fiktive Stichprobe mit der wir dann mit den aus der Statistik bekannten Formeln das 95%-Quantil und damit den $VaR : 0,95$ ermitteln können. Oder wir lassen auch dies von EXCEL erledigen:

	A	B	C	D
1	**diskrete Zufallszahlen**	**normalverteilte Zufallszahlen**	**Summe**	**VaR_0,95**
2			**A3+B3**	QUANTIL(C$3:C$1002;0,95)
3	-1	-3,40718761481232	-4,40718761	**2,18759816371548**
4	-2	-6,07114907318502	-8,07114907	
5	-2	5,58189620204825	3,581896202	
6	-2	-0,0724326718788055	-2,07243267	
7	-2	-4,43685978584332	-6,43685979	
8	-1	-1,79461989195952	-2,79461989	
9	1	-3,36611849459025	-2,36611849	
10	-2	-9,67816201307619	-11,678162	

Abbildung 4.5: VaR mit Monte-Carlo Simulation

Um den $ES_{0,05}$ zu berechnen sortieren alle Werte aus die größer als der $Var_{0,95}$ sind. Bei $n = 1000$ Werten tritt jeder Wert mit einer relativen Häufigkeit von $h_i = \frac{1}{1000}$ auf was als Wahrscheinlichkeit bzw. Gewichtung genommen wird. Die Realisierung auf EXCEL sieht dann folgendermaßen aus:

	A	B	C	D	E	F
1	**diskrete Zufallszahlen**	**normalverteilte Zufallszahlen**	**Summe**	**VaR_0,95**		
2			**A3+B3**	**QUANTIL(C$3:C$1002;0,95)**	**WENN(C3>D$3;1;0)**	**E3*(1/0,05)*(1/1000)*C3**
3	-1	-3,40718761481232	-4,40718761	**2,18759816371548**	0	0
4	-2	-6,07114907318502	-8,07114907		0	0
5	-2	5,58189620204825	3,5818962		1	0,071637924040965
6	-2	-0,0724326718788055	-2,07243267		0	0
7	-2	-4,43685978584332	-6,43685979		0	0
8	-1	-1,79461989195952	-2,79461989		0	0
9	1	-3,36611849459025	-2,36611849		0	0
10	-2	-9,67816201307619	-11,678162		0	0
. . .						
993	-2	-1,95664030375712	-3,9566403		0	0
994	1	-4,15333767697505	-3,15333768		0	0
995	-2	0,345705122545643	-1,65429488		0	0
996	-2	-0,10637070453918	-2,1063707		0	0
997	1	1,70898149211374	2,70898149		1	0,0541796298422749
998	-2	-1,84504488677268	-3,84504489		0	0
999	-1	0,454230965338987	-0,54576903		0	0
1000	1	4,00348630756639	5,00348631		1	0,100069726151328
1001	-2	-3,21564873455032	-5,21564873		0	0
1002	1	-6,65298190812507	-5,65298191		0	0
1003					**Summe=ES_0,05:**	**3,69137322945509**

Abbildung 4.6: ES mit Monte-Carlo Simulation

Angenommen die Werte der beiden Verteilungen sind jeweils in 1000 €, dann wird der Gesamtverlust mit 95%iger Sicherheit nicht höher als ca. 2188 € liegen und in den 5% der ungünstigsten Ausgänge im Schnitt ca. 3691 € betragen.

	A	B	C	D	E	F
1	diskrete Zufallszahlen	normalverteilte Zufallszahlen	Summe	VaR_0,95		
2			A3+B3	QUANTIL(C$3:C$1002;0,95)	WENN(C3>D$3;1;0)	E3*(1/0,05)*(1/1000)*C3
3	-1	-3,40718761481232	-4,40718761	2,18759816371548	0	0
4	-2	-6,07114907318502	-8,07114907		0	0
5	-2	5,58189620204825	3,5818962		1	0,071637924040965
6	-2	-0,0724326718788055	-2,07243267		0	0
7	-2	-4,43685978584332	-6,43685979		0	0
8	-1	-1,79461989195952	-2,79461989		0	0
9	1	-3,36611849459025	-2,36611849		0	0
10	-2	-9,67816201307619	-11,678162		0	0
...						
993	-2	-1,95664030375712	-3,9566403		0	0
994	1	-4,15333767697505	-3,15333768		0	0
995	-2	0,345705122545643	-1,65429488		0	0
996	-2	-0,10637070453918	-2,1063707		0	0
997	1	1,70898149211374	2,70898149		1	0,0541796298422749
998	-2	-1,84504488677268	-3,84504489		0	0
999	-1	0,454230965338987	-0,54576903		0	0
1000	1	4,00348630756639	5,00348631		1	0,100069726151328
1001	-2	-3,21564873455032	-5,21564873		0	0
1002	1	-6,65298190812507	-5,65298191		0	0
1003					Summe=ES_0,05:	3,69137322945509

Abbildung 4.7: ES mit Monte-Carlo Simulation

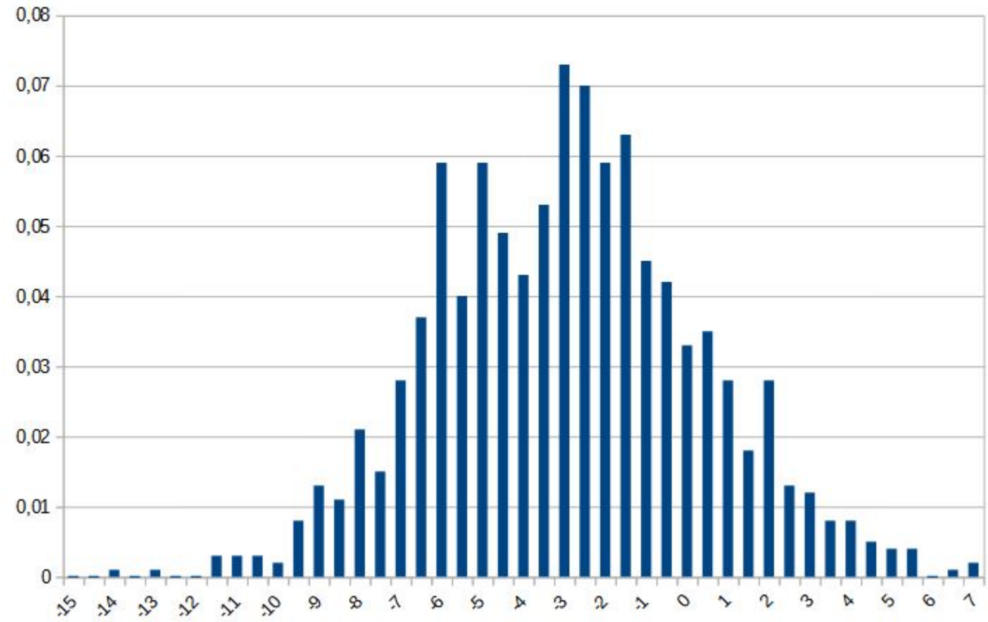


Abbildung 4.8: Visualisierung der Daten aus Beispiel ??

A Anhang

A.1 Standardnormalverteilungstabelle

$\Phi_{o;1}(z) \longrightarrow$

z	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,5279	0,53188	0,53586
0,1	0,53983	0,5438	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,6293	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,6591	0,66276	0,6664	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,7054	0,70884	0,71226	0,71566	0,71904	0,7224
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,7549
0,7	0,75804	0,76115	0,76424	0,7673	0,77035	0,77337	0,77637	0,77935	0,7823	0,78524
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1,0	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,8665	0,86864	0,87076	0,87286	0,87493	0,87698	0,879	0,881	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,9032	0,9049	0,90658	0,90824	0,90988	0,91149	0,91309	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,9222	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,9452	0,9463	0,94738	0,94845	0,9495	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,9608	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,9732	0,97381	0,97441	0,975	0,97558	0,97615	0,9767
2,0	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,9803	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,983	0,98341	0,98382	0,98422	0,98461	0,985	0,98537	0,98574
2,2	0,9861	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,9884	0,9887	0,98899
2,3	0,98928	0,98956	0,98983	0,9901	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,9918	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,9943	0,99446	0,99461	0,99477	0,99492	0,99506	0,9952
2,6	0,99534	0,99547	0,9956	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,9972	0,99728	0,99736
2,8	0,99744	0,99752	0,9976	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861
3,0	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99896	0,999
3,1	0,99903	0,99906	0,9991	0,99913	0,99916	0,99918	0,99921	0,99924	0,99926	0,99929
3,2	0,99931	0,99934	0,99936	0,99938	0,9994	0,99942	0,99944	0,99946	0,99948	0,9995
3,3	0,99952	0,99953	0,99955	0,99957	0,99958	0,9996	0,99961	0,99962	0,99964	0,99965
3,4	0,99966	0,99968	0,99969	0,9997	0,99971	0,99972	0,99973	0,99974	0,99975	0,99976
3,5	0,99977	0,99978	0,99978	0,99979	0,9998	0,99981	0,99981	0,99982	0,99983	0,99983
3,6	0,99984	0,99985	0,99985	0,99986	0,99986	0,99987	0,99987	0,99988	0,99988	0,99989
3,7	0,99989	0,9999	0,9999	0,9999	0,99991	0,99991	0,99992	0,99992	0,99992	0,99992
3,8	0,99993	0,99993	0,99993	0,99994	0,99994	0,99994	0,99994	0,99995	0,99995	0,99995
3,9	0,99995	0,99995	0,99996	0,99996	0,99996	0,99996	0,99996	0,99996	0,99997	0,99997