1 Grundlagen der Wahrscheinlichkeitsrechnung

1.1 Einführung in den Wahrscheinlichkeitsbegriff

Die Verfahren zur Aufbereitung und Auswertung statistischer Daten beziehen sich in der deskriptiven Statistik stets auf die jeweils gegebene Menge von Beobachtungswerten. Die Verallgemeinerung auf übergeordnete statistische Massen ist nicht möglich.

Beispiel 1.1.1:

a) 550 *Studenten einer Universität werden nach ihrem monatlich verfügbaren Einkommen befragt. Es ergibt sich ein durchschnittliches Einkommen von* 642 *€/Monat. Auf das durchschnittlich verfügbare Monatseinkommen* ***aller Studenten*** *dieser Universität kann nicht geschlossen werden.*

b) Eine Befragung von 1200 *Bundesbürgern über* 18 *Jahre hat ergeben, dass* 4% *der Befragten Parteimitglieder sind. Für die Befragten steht damit der Anteil der Parteimitglieder fest. Über den* ***Anteil*** *der Parteimitglieder* ***in der Gesamtbevölkerung*** *wird damit noch nichts ausgesagt.*

Für viele Fragestellungen sind die Verfahren der deskriptiven Statistik nicht geeignet. Insbesondere, wenn nicht die gesamte Masse untersucht werden kann, sondern anstelle einer Vollerhebung nur eine Teilerhebung durchgeführt wird, z.B. aus Kostengründen, aus Zeitgründen oder aus technischen Gründen (zerstörende Prüfung). Die bei einer Teilerhebung erhobene Datenmenge heißt Stichprobe, so dass allgemein von einer **Stichprobenerhebung** gesprochen wird. **Stichprobenerhebung**

Aufgabe der **induktiven Statistik** ist, von den Ergebnissen einer Stichprobenuntersuchung auf die statistische Masse (Grundgesamtheit) zurückzuschließen. Eine wichtige Grundlage solcher Verfahren, die den Rückschluss auf die Grundgesamtheit zulassen, ist die **Wahrscheinlichkeitsrechnung**. Zur Erinnerung sei erwähnt, dass die Anzahl der Elemente der Grundgesamtheit mit N angegeben wird, während die Stichprobe n Elemente umfasst.

Beispiel 1.1.2:

a) *Mit den Regeln der Wahrscheinlichkeitsrechnung kann berechnet werden, wie groß die Chance bzw. Wahrscheinlichkeit ist,* 6 *Richtige beim Zahlenlotto* 6 *aus* 49 *zu haben.*

b) *Für viele Bauteile kann zum Zeitpunkt der Installation nicht gesagt werden, wie groß die Lebensdauer sein wird. Die Lebensdauer ist eine Zufallsgröße. Mittels der Wahrscheinlichkeitsrechnung kann unter bestimmten Voraussetzungen gesagt werden, wie groß die Wahrscheinlichkeit ist, dass die Lebensdauer des Bauteils wenigstens* 100 *Stunden beträgt.*

Im alltäglichen Sprachgebrauch werden die Begriffe „Wahrscheinlichkeit“ oder „wahrscheinlich“ sehr häufig angewendet, und zwar immer dann, wenn ein Sachverhalt oder ein Zusammenhang nicht sicher ist. In der Umgangssprache finden sich Sätze wie diese:

- Wahrscheinlich wird es morgen regnen.
- Wahrscheinlich verliert Eintracht Braunschweig das Fußballspiel gegen Bayern München.

In den folgenden Sätzen hat der Wahrscheinlichkeitsbegriff schon einen konkreteren Inhalt:

- Die Wahrscheinlichkeit im Lotto 6 Richtige zu haben ist sehr klein.
- Es ist unwahrscheinlich, dass ein Säugling bei der Geburt mehr als 5 Kilo wiegt.

Die Aussagen der angeführten Sätze können noch konkretisiert werden, in dem der in diesen Beispielen verwendete Begriff „wahrscheinlich“ als Häufigkeitsaussage formuliert wird.

- Es gibt 13983816 Möglichkeiten 6 aus 49 Zahlen anzukreuzen. Die Gewinnaussichten beim Zahlenlotto 6 Richtige zu haben betragen also, wenn einmal 6 Zahlen ankreuzt werden 1 : 13983816.
- Nur 3 von 1000 Säuglingen wiegen bei der Geburt mehr als 5 Kilo.

Die Wahrscheinlichkeit kann als Maß für die „Chance“ angesehen werden, dass ein bestimmtes Ergebnis von mehreren möglichen Ausgängen eintritt. Dabei kommt es in der Wahrscheinlichkeitsrechnung darauf an, Wahrscheinlichkeiten durch Zahlen auszudrücken.

1.2 Mengen und Mengenoperationen

Die in der Wahrscheinlichkeitsrechnung betrachteten Ereignisse lassen sich als Mengen schreiben, so dass die in der Mengenlehre gültigen Mengenoperationen auch für die dort betrachteten Ereignisse gelten.

Ein Zusammenschluss verschiedener Elemente (Objekte) wird als **Menge** bezeichnet. Ist ein Objekt x Element einer Menge A, so lautet die Schreibweise $x \in A$. Ist x kein Element von A, lautet die Schreibweise $x \notin A$.

Menge

Beispiel 1.2.1:

a) Die Menge A der Primzahlen bis zur Zahl 10, $A = \{2, 3, 5, 7\}$.

b) Die Menge B der Grundfarben, $B = \{Gelb, Rot, Blau\}$.

Die Menge, die alle möglichen Elemente enthält, wird in der Mengenlehre als **Grundraum Ω** bezeichnet[1]. Dagegen heißt die Menge, die kein Element enthält, **leere Menge**. Sie wird mit $\emptyset$ bezeichnet. Im Folgenden werden wichtige Notationen und Mengenoperationen eingeführt:

Grundraum Ω
leere Menge $\emptyset$

- Die zu der Menge A gehörende **Komplementärmenge** $\overline{A}$, wobei $A \subset \Omega$ gilt, enthält alle Elemente von Ω, die nicht in A enthalten sind, $\overline{A} = \{x : x \in \Omega \backslash A\}$.
- Die **Schnittmenge** $A \cap B$ enthält alle Elemente, die in A *und* B enthalten sind, $A \cap B = \{x : x \in A \text{ und } x \in B\}$.
- Die **Vereinigungsmenge** $A \cup B$ enthält alle Elemente, die in A *oder* B enthalten sind, $A \cup B = \{x : x \in A \text{ oder } x \in B\}$.
- Die **Dífferenzmenge** $A \backslash B$ enthält alle Elemente, die in A aber nicht in B enthalten sind, $A \backslash B = \{x : x \in A \text{ und } x \notin B\}$.
- Die Menge B ist **Teilmenge** von A, $B \subset A$, wenn jedes Element aus B auch zu A gehört.
- Die **Potenzmenge** der Menge A, $\mathcal{P}(A)$, enthält alle Teilmengen von A, $\mathcal{P}(A) = \{B : B \subset A\}$ ($\emptyset$ ist auch eine Teimenge von A).
- Die **Mächtigkeit** der Menge A, $|A|$, entspricht der Anzahl (in Zeichen: #) der Elemente, die in der Menge A enthalten sind, $|A| = \#\{x : x \in A\}$.

[1] Im Zusammenhang mit Zufallsexperimenten wird Ω als Ergebnisraum bezeichnet. Ω ist das große O (Omega) des griechischen Alphabets.

Die Mengenoperationen können anschaulich in Form von sogenannten Venn-Diagrammen dargestellt werden.[2]

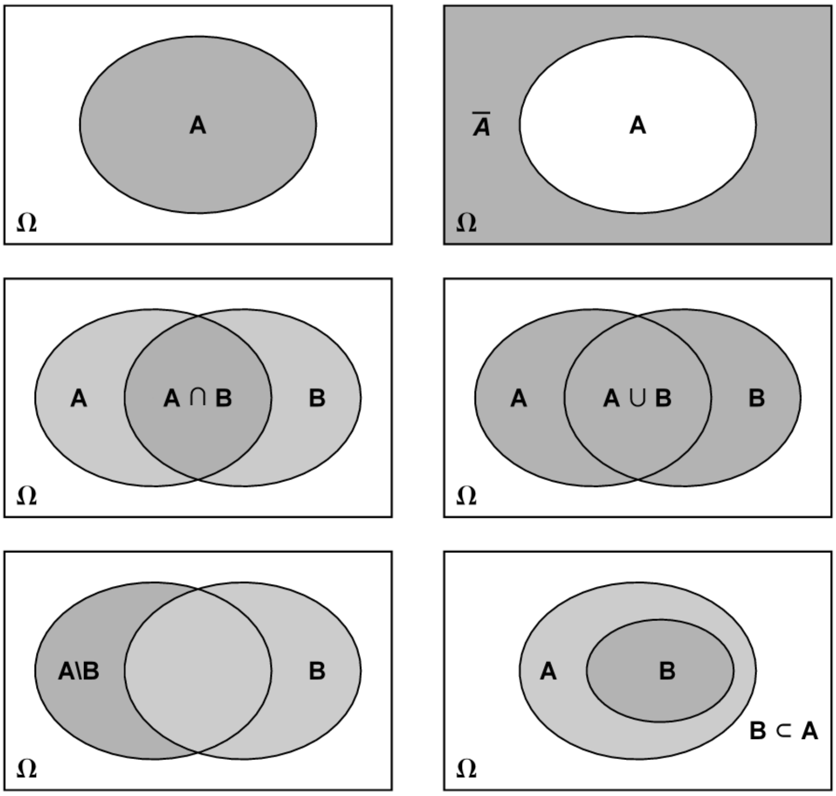


Abbildung 1.2.1: *Venn-Diagramme*

Wichtig für den Umgang mit Mengen sind folgende Rechenregeln, die auch für die später betrachteten Ereignisse gelten.

Rechenregeln für Mengen	
$A \cup B = B \cup A, \quad A \cap B = B \cap A$	(Kommutativgesetz)
$(A \cup B) \cup C = A \cup (B \cup C)$ $(A \cap B) \cap C = A \cap (B \cap C)$	(Assoziativgesetz)
$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$ $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$	(Distributivgesetz)
$\overline{(A \cup B)} = \overline{A} \cap \overline{B}, \quad \overline{(A \cap B)} = \overline{A} \cup \overline{B}$	(De Morgansche Regeln)
$A \backslash B = A \cap \overline{B}, \quad A \subset B \Rightarrow \overline{B} \subset \overline{A}$	

[2]siehe interaktive Mathematica-Applets auf der Homepage des Lehrstuhls http://www.fernuni-hagen.de/ls_statistik/forschung/multimedia/

1.3 Zufallsexperimente und Ereignisse

Gegenstand der Wahrscheinlichkeitsrechnung sind Fragestellungen, deren Ausgang oder dessen Ergebnis vom Zufall abhängt. In diesem Zusammenhang wird von Zufallsexperimenten gesprochen.

Ein **Zufallsexperiment** ist ein beliebig oft wiederholbarer, nach einer ganz bestimmten Vorschrift auszuführender Vorgang mit mindestens zwei sich gegenseitig ausschließenden Ergebnissen, wobei im Voraus nicht eindeutig bestimmbar ist, welches Ergebnis eintreten wird.

Zufallsexperiment

Beispiel 1.3.1:

a) In einer gut durchgemischten Lostrommel befinden sich Gewinnlose und Nieten. Wird ein Los gezogen, kann nicht vorhergesagt werden, ob ein Gewinnlos oder eine Niete gezogen wird.

b) Während der Ausspielung der Lottozahlen 6 *aus* 49 *werden Kugeln, die mit den Zahlen von* $1 - 49$ *beschrieben sind, in einem Gefäß gemischt. Es werden nacheinander* 7 *Kugeln (*6 *Zahlen und die Zusatzzahl) gezogen. Es ist nicht möglich, im Voraus anzugeben, welche Kugeln gezogen werden.*

Nicht nur im Bereich der Wahrscheinlichkeitsrechnung, sondern auch für Stichprobenverfahren ist es von grundlegender Bedeutung, dass die Ergebnisse oder Beobachtungen **wirklich zufällig** eintreten.

Ein Zufallsexperiment ist u.a. dadurch charakterisiert, dass es wenigstens zwei sich gegenseitig ausschließende Ausgänge des Experimentes gibt.

Jeder der möglichen Ausgänge eines Zufallsexperimentes wird als **Ereignis** bezeichnet.

Ereignis

Beispiel 1.3.2:

a) Viele Zusammenhänge aus der Wahrscheinlichkeitsrechnung können besonders gut mit Hilfe des sogenannten Urnenmodells veranschaulicht werden. Betrachtet wird ein Gefäß (Urne), in dem sich verschiedenartige Kugeln befinden. Die Kugeln werden gemischt und es wird zufällig eine Kugel herausgezogen. Beim Ziehen einer

Kugel aus einer Urne mit roten und grünen Kugeln gibt es zwei mögliche voneinander verschiedene Ereignisse: „Die Kugel ist rot" und „Die Kugel ist grün".

b) *Beim Werfen eines Würfels sind Ereignisse die Augenzahlen „1", „2", „3", „4", „5", „6". Ereignisse bei diesem Zufallsexperiment sind aber auch: „Die Augenzahl ist gerade" oder „Die Augenzahl ist ungerade" oder „Die Augenzahl ist höher als 4".*

Der Teil b) des vorhergehenden Beispiels zeigt, dass es Ereignisse gibt, die sich in „kleinere" Ereignisse oder Teilereignisse zerlegen lassen. Das ist z.B. bei dem Ereignis „die Augenzahl ist gerade" der Fall. Dieses Ereignis kann in die Ereignisse „2", „4" und „6" zerlegt werden.

Ergebnis
Elementarereignis

Die einzelnen nicht mehr zerlegbaren und sich gegenseitig ausschließenden Möglichkeiten für den Ausgang eines Zufallsexperimentes heißen **Ergebnisse** oder **Elementarereignisse**. Sie werden mit $\boldsymbol{\omega_1, \omega_2, \omega_3, ..., \omega_n}$ bezeichnet.[3]

Beispiel 1.3.3:

a) *Beim Werfen eines Würfels sind die Elementarereignisse „1", „2", „3", „4", „5", „6".*

b) *Beim Werfen einer Münze sind die Elementarereignisse „Kopf" und „Zahl".*

Ergebnisraum

Sämtliche Ergebnisse bzw. Elementarereignisse als mögliche Ausgänge eines Zufallsexperimentes können zu einer Menge $\Omega = \{\omega_1, \omega_2, \omega_3, ..., \omega_n\}$ zusammengefasst werden. Diese Menge $\boldsymbol{\Omega}$ aller Elementarereignisse heißt **Ergebnisraum**.

Beispiel 1.3.4:

a) *Für das Zufallsexperiment „Werfen eines Würfels" ergibt sich als Ergebnisraum* $\Omega = \{1, 2, 3, 4, 5, 6\}$.

b) *Für das Zufallsexperiment „Werfen einer Münze" lautet der Ergebnisraum* $\Omega = \{\textit{Kopf, Zahl}\}$.

[3] ω_1: Omega 1; es handelt sich hier um das kleine o (Omega) des griechischen Alphabets.

Die Elementarereignisse eines Zufallsexperimentes lassen sich nicht weiter zerlegen. Wie das Beispiel mit dem Würfelwurf in Beispiel 1.3.1 zeigt, gibt es jedoch Ereignisse, die aus mehreren „kleineren“ Ereignissen bestehen.

Aus der Vereinigung von k Elementarereignissen ($k \geq 2$) entsteht ein **zusammengesetztes Ereignis**:

$$A = \{\omega_i\} \cup \{\omega_j\} = \{\omega_i, \omega_j\} \quad \text{bzw.} \quad A = \bigcup_{i=1}^{k} \{\omega_i\}.$$

Ein zusammengesetztes Ereignis A beschreibt eine Teilmenge des Ergebnisraumes Ω.

zusammengesetztes Ereignis

In den folgenden Ausführungen werden zusammengesetzte Ereignisse kurz mit „Ereignis“ bezeichnet.

Beispiel 1.3.5:

Beim Würfelwurf entspricht das Ereignis „Augenzahl ist gerade“ der Teilmenge $A = \{2, 4, 6\}$ des Ergebnisraumes $\Omega = \{1, 2, 3, 4, 5, 6\}$. Die Menge A setzt sich aus den Elementarereignissen $\{2\}, \{4\}$ und $\{6\}$ zusammen.

Die Vereinigung von Ereignissen ist nicht nur für Elementarereignisse, sondern auch für beliebige Ereignisse, definiert.

Unter der **Vereinigung** der Ereignisse A und B wird das Ereignis $\boldsymbol{A \cup B}$ verstanden, das eintritt, wenn wenigstens eines der beiden Ereignisse A und B eintritt.

Vereinigung von Ereignissen

Beispiel 1.3.6:

a) *In einer Urne befinden sich rote, blaue, grüne, gelbe, weiße und schwarze Kugeln. Das Ereignis „Eine gezogene Kugel ist rot oder grün“ tritt ein, wenn die Kugel rot ist, oder wenn die Kugel grün ist. Es ergibt sich als Vereinigung der beiden Ereignisse „Die gezogene Kugel ist rot“ und „Die gezogene Kugel ist grün“.*

b) *Beim Werfen eines Würfels sei G das Ereignis „Die Augenzahl ist gerade“ und A das Ereignis „Die Augenzahl ist kleiner als 4“, d.h. $G = \{2, 4, 6\}$ und $A = \{1, 2, 3\}$. Für die Vereinigung der beiden Ereignisse gilt dann $G \cup A = \{1, 2, 3, 4, 6\}$.*

Manchmal ist auch das gleichzeitige Eintreten zweier Ereignisse von Interesse.

Durchschnitt von Ereignissen

Der **Durchschnitt $\boldsymbol{A \cap B}$** der Ereignisse A und B beschreibt das Ereignis, das eintritt, wenn sowohl das Ereignis A als auch das Ereignis B eintritt.

Beispiel 1.3.7:

a) In einer Urne befinden sich rote und grüne Kugeln. Unabhängig von der Farbe ist ein Teil der Kugeln mit einer 1 *beschriftet. Die übrigen Kugeln sind unbeschriftet. Es gibt also sowohl bei den roten als auch bei den grünen Kugeln solche, die mit einer* 1 *beschriftet sind und andere, die unbeschriftet sind.*
Das Ereignis „Eine gezogene Kugel ist rot und mit einer 1 *beschriftet“ ergibt sich dann als Durchschnitt der beiden Ereignisse „Die gezogene Kugel ist rot“ und „Die gezogene Kugel ist mit einer* 1 *beschriftet“.*

b) Beim Werfen eines Würfels ist das Ereignis „Die Augenzahl ist gerade und kleiner als 5*“ der Durchschnitt der Ereignisse* $G = \{2, 4, 6\}$ *und* $A = \{1, 2, 3, 4\}$*; also* $G \cap A = \{2, 4\}$.

sich gegenseitig ausschließende Ereignisse

Zwei Ereignisse A und B sind **zwei sich gegenseitig ausschließende Ereignisse**, wenn für den Durchschnitt gilt

$$A \cap B = \emptyset.$$

Spezielle Ereignisse sind das komplementäre Ereignis, das sichere Ereignis und das unmögliche Ereignis.

komplementäres Ereignis

Gegeben sei das Ereignis A. Das Ereignis, das eintritt, wenn A nicht eintritt, heißt das zu A **komplementäre Ereignis** und wird mit $\overline{A}$ bezeichnet.

Beispiel 1.3.8:

Beim Werfen eines Würfels sei G *das Ereignis „Die Augenzahl ist gerade“. Das dazu komplementäre Ereignis ist* $\overline{G}$ = *„Die Augenzahl ist ungerade“.*

Ein Ereignis, das immer eintritt, heißt **sicheres Ereignis** und wird mit Ω bezeichnet.

sicheres Ereignis

Ein Ereignis, das nie eintritt, heißt **unmögliches Ereignis** und wird mit $\emptyset$ bezeichnet.

unmögliches Ereignis

Beispiel 1.3.9:

Beim Werfen eines Würfels sei G das Ereignis „Die Augenzahl ist gerade" und U das Ereignis „Die Augenzahl ist ungerade".

*Das zusammengesetzte Ereignis $G \cup U = \Omega = \{1, 2, 3, 4, 5, 6\}$ tritt **immer** ein, da immer eine der Augenzahlen 1, 2, 3, 4, 5 oder 6 oben liegt. Das immer eintretende Ereignis Ω bezeichnet das sichere Ereignis.*

*Der Durchschnitt der beiden Ereignisse $G \cap U$ kann **nie** eintreten. Gerade und ungerade Augenzahlen schließen sich gegenseitig aus. Es handelt sich hier um ein unmögliches Ereignis.*

Sämtliche Ereignisse eines Zufallsexperimentes werden zu der sogenannten **Potenzmenge $\mathcal{P}(\Omega)$** zusammengefasst. Die Potenzmenge $\mathcal{P}(\Omega)$ heißt **Ereignisraum** und entspricht der Menge aller Teilmengen A von Ω. Besteht der Ergebnisraum Ω aus n Elementen, so besteht der Ereignisraum aus 2^n Elementen. Der Ereignisraum $\mathcal{P}(\Omega)$ enthält sowohl Ω als auch das unmögliche Ereignis $\emptyset$.

Potenzmenge, Ereignisraum

Beispiel 1.3.10:

Der Ergebnisraum eines Münzwurfs besteht aus den zwei Elementen K=„Kopf" und Z=„Zahl". Die Potenzmenge, also der Ereignisraum, besteht daher aus $2^2 = 4$ Elementen.

$$\mathcal{P}(\Omega) = \{K, Z, \{K, Z\}, \emptyset\} = \{K, Z, \Omega, \emptyset\}$$

1.4 Die Wahrscheinlichkeit

1.4.1 Axiomatische Definition der Wahrscheinlichkeit nach Kolmogoroff

Der modernen Wahrscheinlichkeitsrechnung liegt fast ausnahmslos eine abstrakte, axiomatische Definition der Wahrscheinlichkeit zugrunde, die auf den russischen Mathematiker Kolmogoroff zurückgeht. Die axiomatische Definition geht nicht von empirischen Beobachtungen aus, sondern die Wahrscheinlichkeit wird als Zuordnung von reellen Zahlen zu den Ereignissen definiert.

Ausgangspunkt für die axiomatische Definition der Wahrscheinlichkeit ist der Ereignisraum $\mathcal{P}(\Omega)$ eines endlichen Ergebnisraumes Ω. Wie bereits erwähnt, enthält der Ereignisraum sowohl das unmögliche Ereignis $\emptyset$ als auch das sichere Ereignis (also den Ergebnisraum Ω). Außerdem besitzt $\mathcal{P}(\Omega)$ die folgenden Eigenschaften:

(1) Mit je zwei Ereignissen ist auch die Vereinigung und der Durchschnitt der beiden Ereignisse in dem Ereignisraum enthalten.

(2) Mit jedem Ereignis ist auch das dazu komplementäre Ereignis enthalten.

axiomatische Definition der Wahrscheinlichkeit

Gegeben sei der Ereignisraum $\mathcal{P}(\Omega)$ eines endlichen Ergebnisraumes Ω. Eine Funktion $\boldsymbol{P}$, die jedem Ereignis A aus $\mathcal{P}(\Omega)$ eine reelle Zahl $\boldsymbol{P}(A)$ zuordnet, heißt Wahrscheinlichkeitmaß auf $\mathcal{P}(\Omega)$ oder kurz Wahrscheinlichkeit (engl. probability), wenn folgende Eigenschaften erfüllt sind:

(1) $\boldsymbol{P}$ ist nichtnegativ : $\boldsymbol{P}(A) \geq 0$.

(2) $\boldsymbol{P}$ ist additiv : $\boldsymbol{P}(A \cup B) = \boldsymbol{P}(A) + \boldsymbol{P}(B)$ für $A \cap B = \emptyset$.

(3) $\boldsymbol{P}$ ist normiert : $\boldsymbol{P}(\Omega) = 1$ bzw. $0 \leq \boldsymbol{P}(A) \leq 1$.

Die Axiome (1) bis (3) bedeuten:

(1) Jedem Ereignis wird als Wahrscheinlichkeit eine nichtnegative Zahl zugeordnet.

(2) Schließen sich zwei Ereignisse A und B gegenseitig aus, dann ergibt sich die Wahrscheinlichkeit für das zusammengesetzte Ereignis $A \cup B$ als Summe der Wahrscheinlichkeit für das Eintreten des Ereignisses A und der Wahrscheinlichkeit für das Eintreten des Ereignisses B.

(3) Die Wahrscheinlichkeit für das sichere Ereignis ist 1.

Es besteht eine gewisse Analogie zwischen relativen Häufigkeiten (als Dezimalbrüche angegeben) und der Wahrscheinlichkeit. Diese Analogie kann an vielen Stellen zum leichteren Verständnis der Wahrscheinlichkeitsrechnung herangezogen werden. Die oben aufgeführten Axiome der Wahrscheinlichkeit können auf bekannte Eigenschaften der relativen Häufigkeiten übertragen werden.

(1) Relative Häufigkeiten sind immer nichtnegative Zahlen.

(2) Betrachtet werden mehrere Merkmalsausprägungen (A, B, C usw.) und die zugehörigen relativen Häufigkeiten ($f(A), f(B), f(C), ...$). Werden die gegebenen Merkmalsausprägungen, die sämtlich voneinander verschieden sein sollen und sich gegenseitig ausschließen, zu einer Klasse zusammengefasst, dann ergibt sich die relative Häufigkeit der Merkmalsausprägungen, die zu dieser Klasse zusammengefasst werden, als Summe der einzelnen relativen Häufigkeiten.

(3) Liegt ein nicht häufbares Merkmal vor, d.h. die statistischen Einheiten können lediglich eine Merkmalsausprägung annehmen, dann ist die Summe der relativen Häufigkeiten aller Merkmalsausprägungen gleich 1. Relative Häufigkeiten sind also normiert. Eine entsprechende Normierung gilt auch für Wahrscheinlichkeiten. Die Wahrscheinlichkeit für das sichere Ereignis ist immer 1.

In dieser Kurseinheit werden die Ereignisse eines Merkmals X mit i und die Ereignisse eines anderen Merkmals Y mit j indiziert.

Beispiel 1.4.1:

100 *Personen wurden nach ihrer Kinderzahl befragt.*

Kinderzahl x_i	0	1	2	3	4	5	6	>6
$f(x_i)$	0.10	0.20	0.30	0.15	0.10	0.05	0.05	0.05

Die relative Häufigkeit für 1,2 *oder* 3 *Kinder beträgt dann:*

$$\begin{aligned} f(1,2 \text{ oder } 3) &= f(1) + f(2) + f(3) \\ &= 0.2 + 0.3 + 0.15 = 0.65. \end{aligned}$$

Eine entsprechende Additivität gilt auch für Wahrscheinlichkeiten, wenn sich die betrachteten Ereignisse gegenseitig ausschließen.

Mit der axiomatischen Definition der Wahrscheinlichkeit werden keinerlei Aussagen darüber gemacht, wie sich bei einer konkreten Fragestellung die Wahrscheinlichkeiten ergeben. Anliegen der axiomatischen Wahrscheinlichkeitstheorie ist ausschließlich ein geschlossenes, logisch aufeinander aufbauendes Gebilde der Wahrscheinlichkeitsrechnung.

1.4.2 Definition der Wahrscheinlichkeit nach Laplace

Die folgende Definition der Wahrscheinlichkeit geht auf den französischen Naturwissenschaftler Laplace zurück. Nach dessen Definition, die manchmal auch als „klassische Definition" der Wahrscheinlichkeit bezeichnet wird, ist die Wahrscheinlichkeit für das Eintreten des Ereignisses A eines Zufallsexperimentes gleich dem Quotienten aus der Anzahl der für das Eintreten des Ereignisses günstigen Fälle und der Anzahl aller möglichen Fälle.

Wahrscheinlichkeit nach Laplace

Definition der Wahrscheinlichkeit nach Laplace:

Gegeben sei ein endlicher Ergebnisraum Ω. Die Wahrscheinlichkeit für das Eintreten des Ereignisses $A \in \Omega$ ergibt sich als **Quotient** aus der Anzahl der für das Eintreten von A **günstigen Fälle** und der Anzahl **aller möglichen Fälle**. Voraussetzung dabei ist die Gleichwahrscheinlichkeit der vorliegenden Elementarereignisse.

$$\boldsymbol{P}(A) = \frac{\text{Anzahl der für } A \text{ günstigen Fälle}}{\text{Anzahl aller möglichen Fälle}}$$

Beispiel 1.4.2:

a) *In einer Urne befinden sich* 80 *Kugeln, von denen* 20 *rot sind. Werden die Kugeln gut gemischt und wird zufällig eine Kugel entnommen, dann ist die Wahrscheinlichkeit für das Ziehen einer roten Kugel:*
$\boldsymbol{P}$(*„Die gezogene Kugel ist rot“*) $= \frac{20}{80} = \frac{1}{4}$.

b) *Beim Werfen eines Würfels gibt es* 6 *mögliche Augenzahlen. Für das Auftreten einer geraden Augenzahl ergibt sich dann folgende Wahrscheinlichkeit:*
$\boldsymbol{P}$(*„Die Augenzahl ist gerade“*) $= \frac{3}{6} = \frac{1}{2}$.

c) *Wird eine Münze geworfen, dann kann eine der beiden Seiten oben liegen. Die Wahrscheinlichkeit dafür, dass Zahl oben liegt ist dann:*
$\boldsymbol{P}$(*„Zahl“*) $= \frac{1}{2}$.

Den Beispielen ist gemeinsam, dass bei der Durchführung des jeweiligen Zufallsexperimentes die betrachteten Fälle (Elementarereignisse) jeweils mit der gleichen Wahrscheinlichkeit eintreten. Jeder Fall hat bei der Durchführung des Zufallsexperimentes die gleiche Chance, realisiert zu werden. Wegen dieser Eigenschaft ist die Definition nach Laplace für viele Fragestellungen ungeeignet, da sie von der Gleichwahrscheinlichkeit aller Elementarereignisse (den möglichen Fällen) ausgeht und nur für solche Ereignisse Wahrscheinlichkeiten liefert, die sich als Vereinigung von „gleichmöglichen“ Elementarereignissen darstellen lassen.

Beispiel 1.4.3:

Die Problematik der Definition nach Laplace bei Nicht-Vorliegen von gleichwahrscheinlichen Elementarereignissen kann an folgender (etwas absurder) Überlegung klar gemacht werden:

Der Student Paul fährt mit dem Schiff nach Amerika. Da er ängstlich ist, möchte er die Wahrscheinlichkeit dafür berechnen, dass er die Überfahrt heil übersteht.

Es gibt einen günstigen Fall, nämlich unbeschadet in Amerika anzukommen, und zwei mögliche Fälle, nämlich unbeschadet anzukommen oder mit dem Schiff unterzugehen.

Der Student Paul macht deshalb folgenden Ansatz:

$$\boldsymbol{P}(\textit{sichere Ankunft}) = \frac{1}{2} = 0.5 = 50\%.$$

Die Wahrscheinlichkeit ist sicher falsch, da die beiden möglichen Fälle nicht gleichmöglich bzw. gleichwahrscheinlich sind.

Eine besondere Rolle spielt die Definition nach Laplace bei allen Problemen der Wahrscheinlichkeitsrechnung, die in den Bereich der Kombinatorik fallen. Dabei geht es im Allgemeinen um Fragestellungen, bei denen zunächst alle logischen Möglichkeiten verschiedener Anordnungen von Ereignissen (Kombinationen) betrachtet werden. Die Anzahl der Möglichkeiten lässt sich mit den Instrumenten der Kombinatorik berechnen. Dabei muss gewährleistet sein, dass alle Kombinationen die gleiche Chance haben, in die Auswahl zu kommen (Forderung der Gleichwahrscheinlichkeit).

Beispiel 1.4.4:

Beim Werfen einer Münze kann Kopf (K) oder Zahl (Z) oben liegen. Wird die Münze zweimal hintereinander geworfen, dann gibt es 4 gleichmögliche Fälle, da die Reihenfolge von Kopf und Zahl beachtet werden muss:

$$(K, K) \quad (K, Z) \quad (Z, K) \quad (Z, Z)$$

Die Wahrscheinlichkeit dafür, dass zweimal Zahl oben liegt, ist dann

$$\boldsymbol{P}((Z, Z)) = \frac{1}{4}$$

und die Wahrscheinlichkeit dafür, dass einmal Kopf und einmal Zahl oben liegt (unabhängig von der Reihenfolge) lautet:

$$\boldsymbol{P}(\text{„}\textit{Es liegt einmal Zahl und einmal Kopf oben}\text{“}) = \frac{2}{4} = \frac{1}{2}.$$

1.4.3 Kombinatorik

Gegeben sei ein endlicher Ergebnisraum Ω mit $A \in \Omega$. Die für die Laplace-Wahrscheinlichkeit benötigte Anzahl der für A günstigen Fälle und die Anzahl der möglichen Fälle kann mittels kombinatorischer Überlegungen gewonnen werden. Dabei geht es um folgende Fragestellungen:

- Wie viele Möglichkeiten gibt es, n Elemente anzuordnen?
- Wie viele Möglichkeiten gibt es, von n Elementen k auszuwählen?

Für n unterschiedliche Elemente gibt es $n!$ (sprich n-**Fakultät**) Möglichkeiten, die Elemente in einer Reihe anzuordnen.

Fakultät

$$n! = 1 \cdot 2 \cdot ... \cdot n$$

Jede Anordnung von n Elementen heißt **Permutation**.

Permutation

Liegen n Elemente mit m unterschiedlichen Ausprägungen vor, wobei n_i Elemente der Ausprägung i vorliegen $(i = 1, ..., m)$ mit $\sum_{i=1}^{m} n_i = n$, dann gibt es

$$\frac{n!}{n_1!...n_m!}$$

verschiedene Möglichkeiten n Elemente anzuordnen.

Die erste Frage ist somit beantwortet. Die Beantwortung der zweiten Frage hängt von dem angewendeten Ziehungsverfahren ab.

(1) **Modell ohne Zurücklegen**

Aus einer Menge von n verschiedenen Elementen werden k Elemente ohne Zurücklegen gezogen, d.h. jedes Element kann nur einmal gezogen werden.
Zu unterscheiden ist noch, ob die Reihenfolge der Elemente relevant ist oder nicht.

- **Reihenfolge wird berücksichtigt**,
  dann gibt es $\frac{n!}{(n-k)!}$ Möglichkeiten n Elemente anzuordnen.

- **Reihenfolge wird nicht berücksichtigt**,
  dann gibt es $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ Möglichkeiten n Elemente anzuordnen.

Binomialkoeffizient

Der Ausdruck $\binom{n}{k}$ heißt **Binomialkoeffizient**. Es wird $0! = 1$ gesetzt, so dass $\binom{n}{0} = 1$ folgt. Weiter ist $\binom{n}{k} = 0$ für $k > n$.

(2) **Modell mit Zurücklegen**

Die Elemente werden nach der Ziehung wieder zurückgelegt, d.h. es können Elemente mehrfach ausgewählt werden.

- **Reihenfolge wird berücksichtigt**,
  dann gibt es n^k Möglichkeiten n Elemente anzuordnen.
- **Reihenfolge wird nicht berücksichtigt**,
  dann gibt es $\binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-1)!}$ Möglichkeiten n Elemente anzuordnen.

Beispiel 1.4.5:

Modell ohne Zurücklegen

a) *Für einen* 100 *m Sprint gibt es* $\frac{8!}{(8-3)!}$ *mögliche Zieleinläufe. Hier wird die Reihenfolge berücksichtigt, da zwischen* 1., 2. *und* 3. *Platz unterschieden wird.*

b) *Aus* 8 *Personen können* 2 *Personen auf* $\binom{8}{2} = 28$ *Arten ausgewählt werden. Hier wird die Reihenfolge nicht berücksichtigt, da es unwichtig ist, ob Person A oder Person B als erstes ausgewählt wird.*

Modell mit Zurücklegen

a) *Wird ein Würfel* 3 *mal hintereinander geworfen, so gibt es* $6^3 = 216$ *mögliche Ergebnisse.*

b) *Eine Tüte soll mit* 2 *Sorten Pralinen gefüllt werden. Es gibt* $\binom{2+10-1}{10} = 11$ *Möglichkeiten eine Tüte mit* 10 *Pralinen zu füllen.*

1.4.4 Statistische Definition der Wahrscheinlichkeit

Die in diesem Abschnitt behandelte statistische Definition der Wahrscheinlichkeit beruht auf einem Zusammenhang zwischen relativen Häufigkeiten und Wahrscheinlichkeiten.

Betrachtet wird ein Zufallsexperiment, das unter völlig gleichen Bedingungen nacheinander n-mal durchgeführt wird. Nach jeder Durchführung wird die relative Häufigkeit für das Auftreten des Ereignisses A bestimmt. In den ersten Versuchen schwanken die berechneten relativen Häufigkeiten für das Auftreten des Ereignisses A sehr stark. Je größer die Anzahl der Versuche des Zufallsexperimentes ist, desto enger schwanken die relativen Häufigkeiten um einen festen Wert.

Beispiel 1.4.6:

Ein Würfel wurde 200*-mal hintereinander geworfen Nach jedem Durchgang wurde die relative Häufigkeit für das Ereignis A=„Auftreten der Augenzahl* 6*" registriert. Für jeden Durchgang ist die Anzahl n der Würfe (x-Achse) und die zugehörige relative Häufigkeit $f_n(A)$ (y-Achse) in Bild 1.4.1 grafisch dargestellt. Dieser Vorgang wurde* 9 *mal wiederholt.*

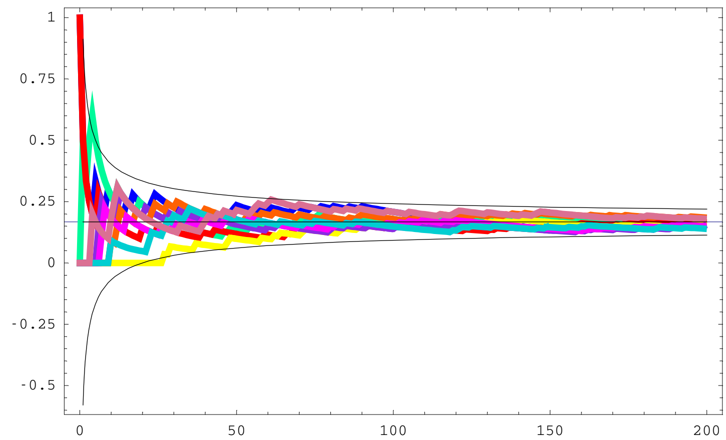


Abbildung 1.4.1: *Relative Häufigkeit für das Auftreten von „Augenzahl* 6*" in Abhängigkeit von der Anzahl der Würfelwürfe*[4]

[4]siehe interaktive Mathematica-Applet „frequenzneu.nbp" auf http://www.fernuni-hagen.de/ls_statistik/forschung/multimedia/eigene.shtml

In Beispiel 1.4.6 schwanken die relativen Häufigkeiten immer weniger um den Wert $\frac{1}{6}$. Je häufiger das Zufallsexperiment durchgeführt wird, desto besser stabilisieren sich die relativen Häufigkeiten. Offensichtlich streben die relativen Häufigkeiten einem „Grenzwert“ zu. Dieser Grenzwert entspricht der Wahrscheinlichkeit für das Ereignis A. Diese Eigenschaft der relativen Häufigkeit führt zu der statistischen Definition der Wahrscheinlichkeit.

statistische Definition der Wahrscheinlichkeit

Nach der statistischen Definition ist die Wahrscheinlichkeit für das Auftreten des Ereignisses A gleich dem Grenzwert der relativen Häufigkeiten, der sich ergibt, wenn das Zufallsexperiment unendlich oft durchgeführt wird:

$$\boldsymbol{P}(A) = \lim_{n\to\infty} f_n(A).$$

Da es in der Wirklichkeit unmöglich ist, ein Zufallsexperiment unendlich oft durchzuführen, ist es ebenso unmöglich, auf die angegebene Art eine Wahrscheinlichkeit zu bestimmen.

empirische Wahrscheinlichkeit

Die statistische Definition der Wahrscheinlichkeit bedeutet, dass über die Berechnung von relativen Häufigkeiten zumindest eine Näherung oder Schätzung der dem Zufallsexperiment zugrunde liegenden Wahrscheinlichkeiten bestimmt werden kann. In dem Fall wird von den sogenannten **empirischen Wahrscheinlichkeiten** gesprochen.

1.4.5 Das Gesetz der großen Zahlen

Im Zusammenhang mit Anwendungen der Statistik wird sehr häufig vom sogenannten Gesetz der großen Zahlen gesprochen, wobei zu beachten ist, dass es verschiedene „Gesetze der großen Zahlen“ gibt.

Die relative Häufigkeit für ein bestimmtes Ereignis eines Zufallexperimentes nähert sich mit zunehmender Anzahl der Wiederholungen des Zufallexperimentes einem bestimmten Wert immer mehr an. Dieser Wert ist im allgemeinen unbekannt, da die relative Häufigkeit mit ihm erst dann übereinstimmt, wenn unendlich viele Versuche durchgeführt werden. Bild 1.4.1 veranschaulicht, dass die relativen Häufigkeiten sich diesem unbekannten Wert, der Wahrscheinlichkeit für das interessierende Ereignis genannt wird, sehr schnell und genau nähern. Dabei wird davon ausgegangen, dass die Wahrscheinlichkeit $\boldsymbol{P}$ für das betrachtete Ereignis eine Konstante ist.

> Die einfachste Form des Gesetzes der großen Zahlen besagt, dass die relative Häufigkeit f für das Auftreten eines Ereignisses sich der theoretischen Wahrscheinlichkeit $\boldsymbol{P}$ für dieses Ereignis nähert, je häufiger das Zufallsexperiment durchgeführt wird.

Gesetz der großen Zahlen

Nach dem Gesetz der großen Zahlen ist dann der Grenzwert der Wahrscheinlichkeit, dass sich die relative Häufigkeit f von der (konstanten) Wahrscheinlichkeit $\boldsymbol{P}$ um mehr als ein beliebig vorgegebenes $\epsilon > 0$ unterscheidet, gleich Null, wenn n gegen unendlich geht.

Das hier skizzierte Gesetz der großen Zahlen hängt eng mit der im vorhergehenden Abschnitt besprochenen statistischen Definition der Wahrscheinlichkeit zusammen.

1.5 Grundregeln der Wahrscheinlichkeitsrechnung

1.5.1 Additionssätze

Das in Abschnitt 1.4.1 bei der axiomatischen Definition der Wahrscheinlichkeit eingeführte 2. Axiom liefert unmittelbar eine Regel für das Rechnen mit Wahrscheinlichkeiten.

Additionssatz für zwei sich gegenseitig ausschließende Ereignisse

Additionssatz für zwei sich gegenseitig ausschließende Ereignisse:

Die Wahrscheinlichkeit für das Ereignis „Vereinigung der Ereignisse A und B“ ist gleich der Summe der Wahrscheinlichkeiten für die Ereignisse A und B, sofern sich die Ereignisse A und B gegenseitig ausschließen.

$$\boldsymbol{P}(A \cup B) = \boldsymbol{P}(A) + \boldsymbol{P}(B), \quad \text{falls } A \cap B = \emptyset.$$

Beispiel 1.5.1:

a) *Beim Werfen eines Würfels gilt $\boldsymbol{P}(2) = \frac{1}{6}$ und $\boldsymbol{P}(4) = \frac{1}{6}$. Die Wahrscheinlichkeit für das Werfen einer „2“ oder „4“ beträgt somit, da sich beide Ereignisse gegenseitig ausschließen,*

$$\boldsymbol{P}(\{2\} \cup \{4\}) = \boldsymbol{P}(2) + \boldsymbol{P}(4) = \tfrac{1}{6} + \tfrac{1}{6} = \tfrac{1}{3}.$$

b) *In einer Urne mit 200 Kugeln befinden sich 40 rote Kugeln und 80 grüne Kugeln. Die Wahrscheinlichkeit für das Ziehen einer roten Kugel (R) beträgt dann $\boldsymbol{P}(R) = \frac{40}{200} = 0.2$. Die Wahrscheinlichkeit für das Ziehen einer grünen Kugel (G) beträgt $\boldsymbol{P}(G) = \frac{80}{200} = 0.4$. Da sich G und R gegenseitig ausschließen, beträgt die Wahrscheinlichkeit für das Ziehen einer grünen oder einer roten Kugel*

$$\boldsymbol{P}(G \cup R) = \boldsymbol{P}(G) + \boldsymbol{P}(R) = 0.4 + 0.2 = 0.6.$$

Dieser Additionssatz lässt sich auf n sich gegenseitig ausschließende Ereignisse erweitern.

Sind $A_1, A_2, ..., A_n$ sich gegenseitig ausschließende Ereignisse eines Zufallsexperimentes ($A_i \cap A_j = \emptyset$ für $i \neq j$), dann gilt

$$\boldsymbol{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \boldsymbol{P}(A_i).$$

Additionssatz für n sich gegenseitig ausschließende Ereignisse

Beispiel 1.5.2:

Gilt beim Werfen eines Würfels $\boldsymbol{P}(2) = \boldsymbol{P}(4) = \boldsymbol{P}(6) = \frac{1}{6}$, so ist $\boldsymbol{P}(\{2\} \cup \{4\} \cup \{6\}) = \boldsymbol{P}(2) + \boldsymbol{P}(4) + \boldsymbol{P}(6) = \frac{1}{2}$.

Schließen sich die Ereignisse nicht gegenseitig aus, dann muss eine andere Regel verwendet werden.

Beispiel 1.5.3:

Gegeben sind zwei bei einem Würfelwurf sich nicht gegenseitig ausschließende Ereignisse A und B.

$A = \{$*die Augenzahl ist kleiner als „4“*$\} = \{1, 2, 3\}$ *mit* $\boldsymbol{P}(A) = \frac{1}{2}$,
$B = \{$*die Augenzahl ist gerade*$\} = \{2, 4, 6\}$ *mit* $\boldsymbol{P}(B) = \frac{1}{2}$.

Für das Ereignis

$A \cup B = \{$*die Augenzahl ist kleiner als „4“* **oder** *gerade*$\} = \{1, 2, 3, 4, 6\}$

ergibt sich die Wahrscheinlichkeit

$$\boldsymbol{P}(A \cup B) = \tfrac{5}{6} \neq \boldsymbol{P}(A) + \boldsymbol{P}(B) = \tfrac{1}{2} + \tfrac{1}{2}.$$

Der Additionssatz für sich gegenseitig ausschließende Ereignisse ist nicht anwendbar, da $\{2\}$ in beiden Ereignissen enthalten ist. Werden die Wahrscheinlichkeiten für A und B einfach addiert, dann wird die Wahrscheinlichkeit für „2“ doppelt gezählt. Wenn also $\boldsymbol{P}(A)$ und $\boldsymbol{P}(B)$ addiert wird, dann muss von der Summe noch $\boldsymbol{P}(2)$ abgezogen werden, um $\boldsymbol{P}(A \cup B)$ zu erhalten. Das ergibt

$$\boldsymbol{P}(A \cup B) = \boldsymbol{P}(A) + \boldsymbol{P}(B) - \boldsymbol{P}(2) = \tfrac{1}{2} + \tfrac{1}{2} - \tfrac{1}{6} = \tfrac{5}{6}.$$

Der in dem Beispiel angesprochenen Sachverhalt kann auch leicht mittels folgendem Venn-Diagramm grafisch verdeutlicht werden.

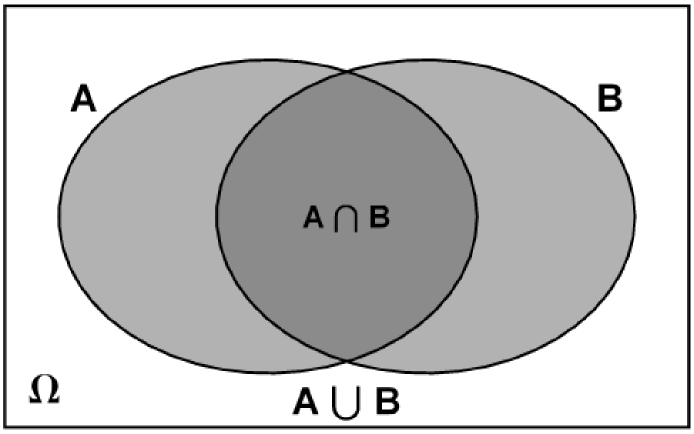


Abbildung 1.5.1: *Vereinigung der Ereignisse A und B*

Soll die Wahrscheinlichkeit für das Ereignis $A \cup B$ bestimmt werden, dann kann $\boldsymbol{P}(A)$ und $\boldsymbol{P}(B)$ addiert werden. Dabei wird dann aber die Fläche $A \cap B$ doppelt gezählt, so dass diese, um ein richtiges Ergebnis zu bekommen, wieder abgezogen werden muss.

Additionssatz für zwei beliebige Ereignisse

Additionssatz für zwei beliebige Ereignisse:

$$\boldsymbol{P}(A \cup B) = \boldsymbol{P}(A) + \boldsymbol{P}(B) - \boldsymbol{P}(A \cap B).$$

Beispiel 1.5.4:

In einer Urne befinden sich 20 *rote und* 30 *grüne Kugeln.* 5 *rote Kugeln und* 10 *grüne Kugeln sind mit einer* 1 *beschriftet. Es sei R das Ereignis „Die gezogene Kugel ist rot“ und E das Ereignis „Die gezogene Kugel ist mit einer* 1 *beschriftet“, dann gilt:*

$\boldsymbol{P}(R) = \frac{20}{50} = 0.4$, $\boldsymbol{P}(E) = \frac{15}{50} = 0.3$ *und* $\boldsymbol{P}(R \cap E) = \frac{5}{50} = 0.1$.

Die Wahrscheinlichkeit dafür, dass die gezogene Kugel rot ist oder mit einer 1 *beschriftet ist, ergibt sich wie folgt:*

$$\boldsymbol{P}(R \cup E) = \boldsymbol{P}(R) + \boldsymbol{P}(E) - \boldsymbol{P}(R \cap E) = 0.4 + 0.3 - 0.1 = 0.6.$$

In Abschnitt 1.3 wurden spezielle Ereignisse betrachtet. Dabei wurde das zu einem Ereignis A komplementäre Ereignis $\overline{A}$ als Ereignis definiert, welches genau dann eintritt, wenn das Ereignis A nicht eintritt. Ein Ereignis und sein komplementäres Ereignis ergeben zusammen immer den vollen Ergebnisraum. Entweder tritt das Ereignis A ein oder es tritt nicht ein (dann tritt aber das komplementäre Ereignis $\overline{A}$ ein). Für A und $\overline{A}$ gilt somit definitionsgemäß

$$A \cup \overline{A} = \Omega.$$

Da sich A und $\overline{A}$ gegenseitig ausschließen und außerdem $\boldsymbol{P}(\Omega) = 1$ gilt, ergibt sich

$$\boldsymbol{P}(A \cup \overline{A}) = \boldsymbol{P}(A) + \boldsymbol{P}(\overline{A}) = \boldsymbol{P}(\Omega) = 1.$$

Aus dieser Gleichung folgt

$$\boldsymbol{P}(A) = 1 - \boldsymbol{P}(\overline{A}).$$

Diese letzte Beziehung kann bei sehr vielen Aufgabenstellungen der Wahrscheinlichkeitsrechnung benutzt werden, um die gesuchte Wahrscheinlichkeit sehr einfach zu bestimmen. Viele Fragestellungen betrachten Ereignisse, die sich aus vielen anderen Ereignissen zusammensetzen. Soll die Wahrscheinlichkeit für ein solches Ereignis A direkt bestimmt werden, dann muss zunächst untersucht werden, wie sich das gesuchte Ereignis A aus den Einzelereignissen zusammensetzt. Anschließend müssen für diese Einzelereignisse die Wahrscheinlichkeiten bestimmt werden, um dann die gewünschte Gesamtwahrscheinlichkeit mit einem ziemlich hohen Rechenaufwand zu ermitteln. Ist die Wahrscheinlichkeit des zu dem gesuchten Ereignis A komplementären Ereignis $\overline{A}$ einfacher zu bestimmen, weil sich $\overline{A}$ zum Beispiel nicht aus so vielen Ereignissen zusammensetzt, dann wird zweckmäßigerweise die Wahrscheinlichkeit für das Ereignis A über die obige Gleichung bestimmt.

Beispiel 1.5.5:

Beim Werfen von zwei Würfeln soll die Wahrscheinlichkeit $\boldsymbol{P}(A)$ für das Ereignis A =„Die Augensumme ist höchstens 11.“ bestimmt werden. Wie Abbildung 1.5.2 zeigt, sind dazu die Wahrscheinlichkeiten von 35 Elementarereignissen zu addieren bzw. können die Wahrscheinlichkeiten für die Augensummen 2, 3, 4, 5,.... 11 addiert werden.

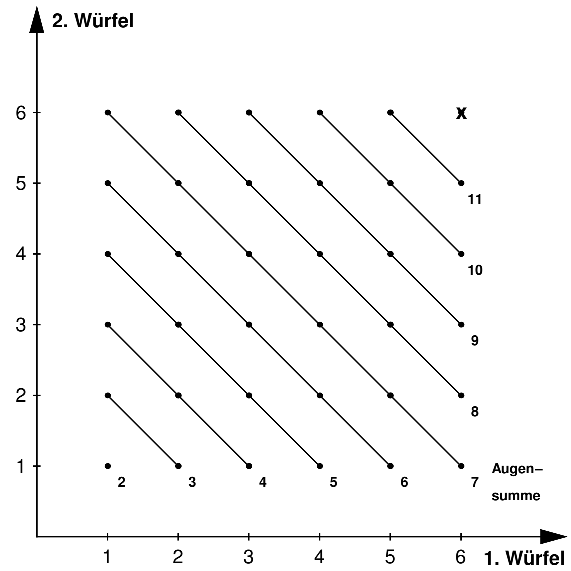


Abbildung 1.5.2: *Werfen von zwei Würfeln*

Die Lösung ergibt sich einfacher, wenn das Komplementärereignis benutzt wird. Das Komplementärereignis zu $A =$ *„Die Augensumme ist höchstens* 11*" ist* $\overline{A} =$ *„Die Augensumme ist gleich* 12*".*

Nun ist $\boldsymbol{P}(\overline{A}) = \frac{1}{36}$ *und damit*

$$\boldsymbol{P}(A) = 1 - \boldsymbol{P}(\overline{A}) = 1 - \frac{1}{36} = \frac{35}{36}.$$

1.5.2 Bedingte Wahrscheinlichkeit

Oft wird das Eintreten von Ereignissen in Abhängigkeit von bestimmten anderen Ereignissen betrachtet.

Beispiel 1.5.6:

In einer Urne liegen 5 *rote und* 3 *grüne Kugeln. Nacheinander werden aus dieser Urne zwei Kugeln entnommen, wobei bei der Entnahme nicht zu erkennen ist, welche Farbe die Kugel hat. Außerdem wird vorausgesetzt, dass vor jedem Zug die Kugeln gemischt werden, so dass das Ziehen der Kugeln einem Zufallsexperiment entspricht mit den beiden möglichen Ergebnissen* $R =$*„Die Farbe der Kugel ist rot" oder*

(G) =„Die Farbe der Kugel ist grün“. Gesucht ist die Wahrscheinlichkeit, beim zweiten Zug eine grüne Kugel aus der Urne zu ziehen.

Vor dem ersten Zug befinden sich 8 Kugeln in der Urne, von denen jede die gleiche Chance hat, gezogen zu werden, so dass die Wahrscheinlichkeitsdefinition nach Laplace anwendbar ist.

$$\boldsymbol{P}(R_1) = \tfrac{5}{8} \text{ und } \boldsymbol{P}(G_1) = \tfrac{3}{8}.$$

Nach dem ersten Zug befinden sich noch 7 Kugeln in der Urne.

*Ist im ersten Zug eine **rote** Kugel gezogen worden, dann sind noch 4 rote und 3 grüne Kugeln in der Urne, so dass für das Ziehen einer grünen Kugel beim zweiten Zug gilt:*

$$\boldsymbol{P}(G_2) = \tfrac{3}{7}.$$

*Wird im ersten Zug eine **grüne** Kugel gezogen, dann befinden sich noch 5 rote und 2 grüne Kugeln in der Urne und für die Wahrscheinlichkeit, im zweiten Zug eine grüne Kugel aus der Urne zu ziehen, ergibt sich:*

$$\boldsymbol{P}(G_2) = \tfrac{2}{7}.$$

Die Wahrscheinlichkeit, im zweiten Zug eine grüne Kugel aus der Urne zu entnehmen, kann nur genau angegeben werden, wenn bekannt ist, welche Farbe die Kugel hat, die im ersten Zug aus der Urne genommen wurde. Das Ergebnis des zweiten Zuges hängt vom Ergebnis des ersten Zuges ab. Das Ergebnis des ersten Zuges ist eine Bedingung für die Wahrscheinlichkeit des zweiten Zuges.

In dem Beispiel bedingt das Ergebnis des ersten Versuches die Wahrscheinlichkeit für das Ergebnis des zweiten Versuches, deshalb wird von bedingter oder konditionaler Wahrscheinlichkeit gesprochen.

bedingte Wahrscheinlichkeit

Unter der **bedingten Wahrscheinlichkeit** $\boldsymbol{P}(B|A)$ (lies: Wahrscheinlichkeit für B gegeben A, oder für B unter der Hypothese bzw. Bedingung A) wird die Wahrscheinlichkeit für das Eintreten des Ereignisses B unter der Voraussetzung, dass das Ereignis A bereits eingetreten ist, verstanden.

$$\boldsymbol{P}(B|A) = \frac{\boldsymbol{P}(A \cap B)}{\boldsymbol{P}(A)}; \quad \boldsymbol{P}(A) > 0.$$

Entsprechend gilt:

$$\boldsymbol{P}(A|B) = \frac{\boldsymbol{P}(A \cap B)}{\boldsymbol{P}(B)}; \quad \boldsymbol{P}(B) > 0.$$

Beispiel 1.5.7:

Aus den Studierenden einer Universität wird zufällig eine Person ausgewählt. Es ist: A = „Es handelt sich um eine Studentin“ und B = „Das Studienfach ist Wirtschaftswissenschaften (WiWi)“. Gesucht ist die Wahrscheinlichkeit dafür, dass die ausgewählte Person WiWi studiert unter der Bedingung, dass es sich um eine Studentin handelt, also $\boldsymbol{P}(B|A)$. Wenn alle Studierenden die gleiche Chance haben, ausgewählt zu werden, dann kann die Wahrscheinlichkeitsdefinition nach Laplace angewendet werden. Die gesuchte Wahrscheinlichkeit ergibt sich, wenn die Anzahl der Studentinnen mit Fach WiWi „$n(A \cap B)$“ zur Anzahl der Studentinnen insgesamt „$n(A)$“ in Beziehung gesetzt wird.

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

Erweitern des Bruches mit $\frac{1}{n}$ ergibt

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

1.5.3 Unabhängige Ereignisse

Beispiel 1.5.6 hat gezeigt, dass die Wahrscheinlichkeit, im zweiten Zug eine grüne Kugel aus der Urne zu ziehen, davon abhängt, welche Farbe die zuerst gezogene Kugel hat. Das Beispiel legt die Vermutung nahe, dass es bei Zufallsexperimenten Ereignisse geben kann, die voneinander abhängig sind und andererseits auch voneinander unabhängige Ereignisse vorkommen können.

Beispiel 1.5.8:

Wie in Beispiel 1.5.6 sei eine Urne mit 3 *grünen und* 5 *roten Kugeln gegeben. Seien* G_1 *bzw.* G_2 *die Ereignisse „Grüne Kugel im ersten" bzw. „Grüne Kugel im zweiten Zug".*

Wird jetzt davon ausgegangen, dass vor dem Ziehen der zweiten Kugel die erste gezogene Kugel wieder zurückgelegt wird, dann befinden sich vor dem Ziehen der zweiten Kugel wieder 8 *Kugeln, und zwar* 3 *grüne und* 5 *rote in der Urne. Für die Wahrscheinlichkeit, im zweiten Zug eine grüne Kugel zu ziehen ergibt sich dann*

$$\boldsymbol{P}(G_2) = \boldsymbol{P}(G_2|G_1) = \boldsymbol{P}(G_2|\overline{G}_1) = \tfrac{3}{8}.$$

Die Wahrscheinlichkeit im zweiten Zug eine grüne Kugel zu ziehen hängt jetzt nicht mehr davon ab, welche Farbe die im ersten Zug gezogene Kugel hatte.
Die Abhängigkeit bzw. Unabhängigkeit der beiden Ereignisse G_2 *und* G_1 *wird dadurch beeinflusst, dass einmal „ohne Zurücklegen" und das andere Mal „mit Zurücklegen" aus der Urne gezogen wird.*

Die Ereignisse A und B sind genau dann **(stochastisch) unabhängig**, wenn gilt

$$\boldsymbol{P}(B|A) = \boldsymbol{P}(B|\overline{A}) = \boldsymbol{P}(B) \quad \text{oder} \quad \boldsymbol{P}(A|B) = \boldsymbol{P}(A|\overline{B}) = \boldsymbol{P}(A).$$

Gilt $\boldsymbol{P}(B|A) \neq \boldsymbol{P}(B|\overline{A})$ bzw. $\boldsymbol{P}(A|B) \neq \boldsymbol{P}(A|\overline{B})$, so sind die Ereignisse **(stochastisch) abhängig**.

stochastisch unabhängige/ abhängige Ereignisse

Zwei Ereignisse A und B sind genau dann voneinander abhängig, wenn die Wahrscheinlichkeit für das Auftreten des Ereignisses B davon abhängt, ob das Ereignis A aufgetreten ist oder nicht. Die Abhängigkeit bzw. Unabhängigkeit zweier Ereignisse und deren Komplemente kann

Vierfeldertafel

mitunter sehr leicht mit Hilfe einer sogenannten **Vierfeldertafel** veranschaulicht werden.

Allgemeine Form einer Vierfeldertafel:

	B	$\overline{B}$	$\sum$
A	$\boldsymbol{P}(A \cap B)$	$\boldsymbol{P}(A \cap \overline{B})$	$\boldsymbol{P}(A)$
$\overline{A}$	$\boldsymbol{P}(\overline{A} \cap B)$	$\boldsymbol{P}(\overline{A} \cap \overline{B})$	$\boldsymbol{P}(\overline{A})$
$\sum$	$\boldsymbol{P}(B)$	$\boldsymbol{P}(\overline{B})$	1

Beispiel 1.5.9:

In einer Urne befinden sich 20 *rote und* 30 *grüne Kugeln. Davon sind* 5 *rote und* 10 *grüne Kugeln mit einer* 1 *beschriftet.*
Es sei R das Ereignis „Die gezogene Kugel ist rot“, G das Ereignis „Die gezogene Kugel ist grün“ und E das Ereignis „Die gezogene Kugel ist mit einer 1 *beschriftet“.*

	E	$\overline{E}$	$\sum$
R	0.1	0.3	0.4
$G = \overline{R}$	0.2	0.4	0.6
$\sum$	0.3	0.7	1

Nach der Definition der bedingten Wahrscheinlichkeit gilt:

$$\boldsymbol{P}(E|R) = \frac{\boldsymbol{P}(E \cap R)}{\boldsymbol{P}(R)} = \frac{0.1}{0.4} = 0.25.$$

Ferner gilt:

$$\boldsymbol{P}(E|\overline{R}) = \frac{\boldsymbol{P}(E \cap \overline{R})}{\boldsymbol{P}(\overline{R})} = \frac{0.2}{0.6} = 0.33.$$

Daraus ergibt sich die Beziehung

$$\boldsymbol{P}(E|R) \neq \boldsymbol{P}(E|\overline{R}) \neq \boldsymbol{P}(E) = 0.3.$$

Definitionsgemäß sind damit die beiden Ereignisse E und R abhängig.

Die Unabhängigkeit bzw. Abhängigkeit von Ereignissen spielt für die Statistik eine wichtige Rolle. In zahlreichen Verfahren aus dem Bereich der Stichprobentheorie und anderen Anwendungen der Statistik wird bei den betrachteten Ereignissen bzw. Zufallsexperimenten Unabhängigkeit vorausgesetzt.

1.5.4 Multiplikationssätze

Aus der Definition der bedingten Wahrscheinlichkeit folgt der

Multiplikationssatz für beliebige Ereignisse:

$$\boldsymbol{P}(A \cap B) = \boldsymbol{P}(B|A) \cdot \boldsymbol{P}(A),$$

d.h., die Wahrscheinlichkeit $\boldsymbol{P}(A \cap B)$, dass sowohl Ereignis A als auch Ereignis B eintritt, ist gleich der bedingten Wahrscheinlichkeit für das Eintreten des Ereignisses B, bei gegebenem Ereignis A, multipliziert mit der Wahrscheinlichkeit für das Eintreten des Ereignisses A. Entsprechend gilt

$$\boldsymbol{P}(A \cap B) = \boldsymbol{P}(A|B) \cdot \boldsymbol{P}(B).$$

Multiplikationssatz für beliebige Ereignisse

Beispiel 1.5.10:

Nach Beispiel 1.5.6 ist

$$\boldsymbol{P}(G_2|G_1) = \tfrac{2}{7} \quad \textit{und} \quad \boldsymbol{P}(G_1) = \tfrac{3}{8}.$$

Die Wahrscheinlichkeit, sowohl im ersten als auch im zweiten Zug eine grüne Kugel zu ziehen, beträgt dann

$$\boldsymbol{P}(G_2 \cap G_1) = \boldsymbol{P}(G_2|G_1) \cdot \boldsymbol{P}(G_1) = \tfrac{2}{7} \cdot \tfrac{3}{8} = \tfrac{3}{28}.$$

Der folgende Multiplikationssatz für unabhängige Ereignisse (s. Analogie zu KE 1 Abschnitt 3.2.2, zweidimensionale Verteilung empirisch unabhängiger Merkmale) findet häufig Anwendung, da in sehr vielen Fällen die Wahrscheinlichkeit für das gemeinsame Auftreten von Ereignissen gesucht ist und dabei ausgegangen werden kann, dass die Ereignisse unabhängig voneinander sind.

Multiplikationssatz für unabhängige Ereignisse

Multiplikationssatz für unabhängige Ereignisse:

Für unabhängige Ereignisse gilt die Beziehung $\boldsymbol{P}(A|B) = \boldsymbol{P}(A)$ bzw. $\boldsymbol{P}(B|A) = \boldsymbol{P}(B)$, so dass sich der allgemeine Multiplikationssatz wie folgt vereinfacht:

$$\boldsymbol{P}(A \cap B) = \boldsymbol{P}(A) \cdot \boldsymbol{P}(B).$$

Die Wahrscheinlichkeit für das gemeinsame Auftreten der Ereignisse A und B ist gleich dem Produkt der Wahrscheinlichkeiten für das Eintreten des Ereignisses A und für das Eintreten des Ereignisses B.

Beispiel 1.5.11:

a) *Beim Werfen von zwei idealen Würfeln sind die Ergebnisse der beiden Würfel unabhängig voneinander. Die Wahrscheinlichkeit, mit dem ersten Würfel eine* 2 *und mit dem zweiten Würfel eine* 6 *zu werfen, beträgt dann*

$$\boldsymbol{P}(\{2\} \cap \{6\}) = \boldsymbol{P}(2) \cdot \boldsymbol{P}(6) = \tfrac{1}{6} \cdot \tfrac{1}{6} = \tfrac{1}{36}.$$

b) *Beim mehrmaligen Werfen einer Münze sind die einzelnen Würfe unabhängig voneinander. Die Wahrscheinlichkeit, viermal hintereinander Zahl zu werfen, beträgt dann*

$$\boldsymbol{P}(Z_1 \cap Z_2 \cap Z_3 \cap Z_4) = \boldsymbol{P}(Z_1) \cdot \boldsymbol{P}(Z_2) \cdot \boldsymbol{P}(Z_3) \cdot \boldsymbol{P}(Z_4) = \tfrac{1}{2} \cdot \tfrac{1}{2} \cdot \tfrac{1}{2} \cdot \tfrac{1}{2} = \tfrac{1}{16}.$$

Das in dem Teil b) dieses Beispiels angesprochene Problem wird sehr häufig in völlig missverständlicher Weise gedeutet. Es wird nämlich immer wieder die folgende Ansicht vertreten: Wenn mit einer Münze mehrmals geworfen wird und es ist 5- oder 8mal hintereinander Zahl gekommen, dann wird vermutet, dass nun mit einer sehr hohen Wahrscheinlichkeit Kopf kommt, denn dass so häufig hintereinander Zahl kommt, ist beinahe unmöglich. In dieser Überlegung steckt folgender Trugschluss: Wenn mit einer Münze geworfen wird, dann ist in der Tat die Wahrscheinlichkeit, das 6- oder 9mal hintereinander nur Zahl kommt, sehr klein, nämlich $1/64$ bzw. $1/512$. Wenn aber bereits 5- bzw. 8mal

Zahl hintereinander erschienen ist und die Würfe erfolgen unabhängig voneinander, dann ist bei dem folgenden Wurf die Wahrscheinlichkeit, dass Zahl oben liegt, ebenso 1/2 wie bei jedem anderen Wurf auch.

Ein ähnlicher Trugschluss findet sich unter Spielern, die gern Roulette spielen. Hier wird immer wieder geglaubt, wenn 8- oder 10mal hintereinander rot bzw. schwarz aufgetreten ist, müsse mit sehr hoher Wahrscheinlichkeit die andere Farbe kommen. Dahinter steckt der gleiche Fehler wie beim Münzwurf.

Zum Abschluss dieses Abschnittes wird noch ein Beispiel angegeben, das besonders deutlich zeigt, dass mit der Anschauung und mit Vermutungen Wahrscheinlichkeiten falsch abgeschätzt werden.

Beispiel 1.5.12:

In einem Raum befinden sich n Personen. Wie groß ist die Wahrscheinlichkeit dafür, dass wenigstens 2 *am gleichen Tag Geburtstag haben (Ereignis A)? Ein Jahr werde mit* 365 *Tagen gerechnet.*

Zunächst wird die Wahrscheinlichkeit dafür bestimmt, dass alle Personen an verschiedenen Tagen Geburtstag haben (Ereignis $\overline{A}$). Da das Jahr 365 *Tage hat, gibt es insgesamt* 365^n *Möglichkeiten der Geburtstagszusammenstellung (Variation mit Berücksichtigung der Anordnung und mit Wiederholung). Davon sind bei*

$$\frac{365!}{(365-n)!} = 365 \cdot 364 \cdot ... \cdot (365 - n + 1)$$

Möglichkeiten die Geburtstage verschieden (ohne Wiederholung). Also gilt

$$\boldsymbol{P}(\overline{A}) = \frac{365 \cdot 364 \cdot ... \cdot (365 - n + 1)}{365^n}$$

und

$$\boldsymbol{P}(A) = 1 - \boldsymbol{P}(\overline{A}) = 1 - \frac{365 \cdot 364 \cdot ... \cdot (365 - n + 1)}{365^n}.$$

Für $n = 1, 2, 3, ..., 96$ *sind die Ergebnisse im folgenden zusammengestellt. Es ist zu erkennen, dass bereits bei* 23 *Personen* $\boldsymbol{P}(A) > 0.5$ *gilt.*

n	$\boldsymbol{P}(A)$	n	$\boldsymbol{P}(A)$	n	$\boldsymbol{P}(A)$
1	0	33	0.77497185	65	0.99768310
2	0.00273972	34	0.79531686	66	0.99809570
3	0.00820416	35	0.81438323	67	0.99844004
4	0.01635591	36	0.83218210	68	0.99872639
5	0.02713557	37	0.84873400	69	0.99896366
6	0.04046248	38	0.86406782	70	0.99915957
7	0.05623570	39	0.87821966	71	0.99932075
8	0.07433529	40	0.89123181	72	0.99945288
9	0.09462383	41	0.90315161	73	0.99956080
10	0.11694817	42	0.91403047	74	0.99964864
11	0.14114137	43	0.92392285	75	0.99971987
12	0.16702478	44	0.93288536	76	0.99977743
13	0.19441027	45	0.94097590	77	0.99982377
14	0.22310251	46	0.94825284	78	0.99986095
15	0.25290132	47	0.95477440	79	0.99989066
16	0.28360400	48	0.96059797	80	0.99991433
17	0.31500766	49	0.96577960	81	0.99993310
18	0.34691141	50	0.97037358	82	0.99994795
19	0.37911852	51	0.97443199	83	0.99995964
20	0.41143838	52	0.97800450	84	0.99996882
21	0.44368833	53	0.98113811	85	0.99997599
22	0.47569530	54	0.98387696	86	0.99998158
23	0.50729723	55	0.98626228	87	0.99998592
24	0.53834425	56	0.98833235	88	0.99998928
25	0.56869970	57	0.99012245	89	0.99999186
26	0.59824082	58	0.99166497	90	0.99999384
27	0.62685928	59	0.99298944	91	0.99999536
28	0.65446147	60	0.99412266	92	0.99999652
29	0.68096853	61	0.99508879	93	0.99999739
30	0.70631624	62	0.99590957	94	0.99999806
31	0.73045463	63	0.99660438	95	0.99999856
32	0.75334752	64	0.99719047	96	0.99999893

***Legende** zur Tabelle:*

n = *Anzahl der Personen,*

$\boldsymbol{P}(A)$ = *Wahrscheinlichkeit dafür, dass wenigstens 2 von n Personen am gleichen Tag Geburtstag haben.*

1.5.5 Das Theorem von Bayes

Für manche Anwendungen der Wahrscheinlichkeitsrechnung spielt eine Regel eine Rolle, die nach dem englischen Statistiker Thomas Bayes (1702-1761) benannt ist. Dazu wird zunächst der **Satz von der totalen Wahrscheinlichkeit** hergeleitet:

Sei Ω ein Ergebnisraum, der sich als Vereinigung von paarweise ausschließenden Ereignissen $A_1, A_2, ..., A_n$ darstellen lässt. Für die Ereignisse $A_i \subset \Omega$ $(i = 1, ..., n)$, gilt somit

$$A_i \cap A_j = \emptyset \quad \text{für } i \neq j, \quad \text{mit } \bigcup_{i=1}^{n} A_i = \Omega.$$

Bild 1.5.3 gibt eine grafische Darstellung des Ergebnisraumes an. Sei nun $B \subseteq \Omega$ ein Ereignis mit $\boldsymbol{P}(B) > 0$. Dann gilt:

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup ... \cup (A_n \cap B).$$

Da die A_i sich paarweise ausschließen, gilt das auch für die $(A_i \cap B)$, und damit

$$\boldsymbol{P}(B) = \boldsymbol{P}(A_1 \cap B) + \boldsymbol{P}(A_2 \cap B) + ... + \boldsymbol{P}(A_n \cap B) = \sum_{i=1}^{n} \boldsymbol{P}(A_i \cap B).$$

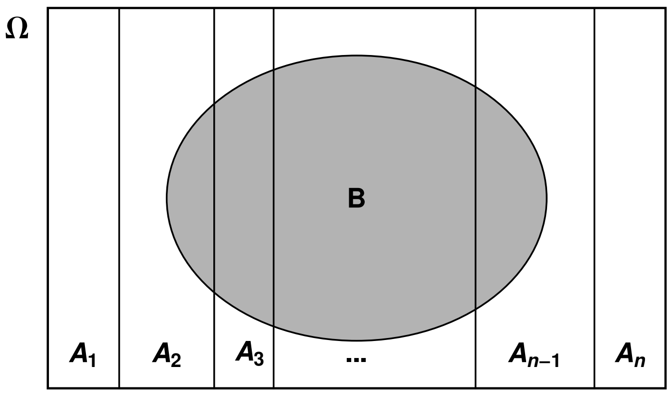


Abbildung 1.5.3: *Darstellung von Ω zum Satz der totalen Wahrscheinlichkeit*

Aus dem Multiplikationssatz für beliebige Ereignisse folgt

$$\boldsymbol{P}(A_i \cap B) = \boldsymbol{P}(B|A_i) \cdot \boldsymbol{P}(A_i).$$

Dies in die obige Summe eingesetzt ergibt den sogenannten Satz von der totalen Wahrscheinlichkeit.

Satz von der totalen Wahrscheinlichkeit

Es seien $B, A_1, A_2, ..., A_n$ Ereignisse aus einem Ergebnisraum Ω mit

(1) $\boldsymbol{P}(B) > 0$,
(2) $\bigcup\limits_{i=1}^{n} A_i = \Omega$,
(3) $A_i \cap A_j = \emptyset \quad$ für $i \neq j$.

Dann gilt der **Satz von der totalen Wahrscheinlichkeit**:

$$\boldsymbol{P}(B) \;=\; \sum_{i=1}^{n} \boldsymbol{P}(B|A_i) \cdot \boldsymbol{P}(A_i)$$

Beispiel 1.5.13:

Von den Studenten des FB Wirtschaftswissenschaften einer Universität seien 40% im ersten, 25% im zweiten, 15% im dritten und 20% im vierten Semester. An der Vorlesung Statistik I nehmen 80% der Studenten im ersten, 20% der Studenten im zweiten, 10% der Studenten im dritten und 10% der Studenten im vierten Semester teil. Aus den Studenten des FB Wirtschaftswissenschaften soll ein Student zufällig ausgewählt werden. Wie groß ist die Wahrscheinlichkeit dafür, dass dieser Student an der Statistik-Vorlesung teilnimmt?

Es werden folgende Ereignisse definiert:

B *: Hörer der Statistik-Vorlesung,*
A_1 *: Student aus dem ersten Semester,*
A_2 *: Student aus dem zweiten Semester,*
A_3 *: Student aus dem dritten Semester,*
A_4 *: Student aus dem vierten Semester.*

Da nur Studenten aus den ersten vier Semestern betrachtet werden, ist

$$A_1 \cup A_2 \cup A_3 \cup A_4 = \Omega \quad \textit{und } A_i \cap A_j = \emptyset \quad \textit{für } i \neq j.$$

Aus dem Satz der totalen Wahrscheinlichkeit ergibt sich dann

$$\begin{aligned}\boldsymbol{P}(B) &= \sum_{i=1}^{4} \boldsymbol{P}(B|A_i) \cdot \boldsymbol{P}(A_i) \\ &= 0.8 \cdot 0.4 + 0.2 \cdot 0.25 + 0.1 \cdot 0.15 + 0.1 \cdot 0.2 \\ &= 0.32 + 0.05 + 0.015 + 0.02 = 0.405.\end{aligned}$$

Mit einer Wahrscheinlichkeit von 0.405 *ist der zufällig ausgewählte Student des FB Wirtschaftswissenschaften Hörer der Statistik-Vorlesung.*

Aus der Definition der bedingten Wahrscheinlichkeit ergibt sich

$$\boldsymbol{P}(A_j|B) = \frac{\boldsymbol{P}(A_j \cap B)}{\boldsymbol{P}(B)} \quad \text{sowie } \boldsymbol{P}(A_j \cap B) = \boldsymbol{P}(B|A_j) \cdot \boldsymbol{P}(A_j).$$

Wird in dem Bruch für $\boldsymbol{P}(B)$ der Ausdruck aus dem Satz der totalen Wahrscheinlichkeit und für $\boldsymbol{P}(A_j \cap B)$ der obenstehende Ausdruck eingesetzt, so ergibt sich das Theorem von Bayes.

Theorem von Bayes

Es seien $B, A_1, A_2, ..., A_n$ Ereignisse aus einem Ergebnisraum Ω mit

(1) $\boldsymbol{P}(B) > 0$,
(2) $\bigcup_{i=1}^{n} A_i = \Omega$,
(3) $A_i \cap A_j = \emptyset$ für $i \neq j$.

Dann gilt das **Theorem von Bayes**:

$$\boldsymbol{P}(A_j|B) = \frac{\boldsymbol{P}(B|A_j) \cdot \boldsymbol{P}(A_j)}{\sum_{i=1}^{n} \boldsymbol{P}(B|A_i) \cdot \boldsymbol{P}(A_i)}.$$

Beispiel 1.5.14:

Es wird an das Beispiel 1.5.13 zur totalen Wahrscheinlichkeit angeknüpft. Gesucht ist die Wahrscheinlichkeit $\boldsymbol{P}(A_1|B)$, dass ein zufällig ausgewählter Hörer der Statistik-Vorlesung aus dem ersten Semester ist. Aus dem Theorem von Bayes ergibt sich:

$$\boldsymbol{P}(A_1|B) = \boldsymbol{P}(B|A_1) \cdot \frac{\boldsymbol{P}(A_1)}{\sum_{i=1}^{4} \boldsymbol{P}(B|A_i) \cdot \boldsymbol{P}(A_i)}.$$

Es ist $\boldsymbol{P}(A_1) = 0.4$ *und* $\boldsymbol{P}(B|A_1) = 0.8$.

Letzteres ist die Wahrscheinlichkeit dafür, dass ein Student aus dem ersten Semester die Statistik-Vorlesung besucht. Für den Nenner wurde oben

$$\sum_{i=1}^{4} \boldsymbol{P}(B|A_i) \cdot \boldsymbol{P}(A_i) = 0.405$$

berechnet. Es ist also

$$\boldsymbol{P}(A_1|B) = \frac{0.4 \cdot 0.8}{0.405} = 0.79.$$

Das Theorem von Bayes findet bei Entscheidungen unter Unsicherheit bzw. in Risikosituationen Anwendung, und zwar speziell zur Abschätzung von Wahrscheinlichkeiten. Für das Theorem von Bayes liegt dann oft folgende Situation vor:

Es gibt mehrere sich gegenseitig ausschließende Zustände oder Alternativen $A_1, A_2, ..., A_n$ und Schätzungen der Wahrscheinlichkeiten $\boldsymbol{P}(A_i)$ über das Eintreten der Zustände bzw. Alternativen. Die $\boldsymbol{P}(A_i)$ heißen **a priori Wahrscheinlichkeiten**, da sie die Wahrscheinlichkeiten für das Eintreten der Ereignisse A_i **vor** der Kenntnis eines Ereignisses B angeben. Es werden dann über Realisationen von Zufallsexperimenten Wahrscheinlichkeiten $\boldsymbol{P}(B|A_j)$ ermittelt. Das Eintreten des Ereignisses B dient unter Verwendung des Theorems von Bayes zur Berechnung der **a posteriori Wahrscheinlichkeiten** $\boldsymbol{P}(A_j|B)$, welches die Wahrscheinlichkeit für das Eintreten des Ereignisses A_j **nach** Kenntnis von B angibt. Aufgrund dieser Zusatzinformation wird somit eine Verbesserung gegenüber den a priori Wahrscheinlichkeiten erzielt.

Beispiel 1.5.15*:*

Ein Unternehmen möchte ein neues Produkt auf den Markt bringen. Dazu wird intern eine Qualitätsprüfung durchgeführt. Von den bisherigen Produkten des Unternehmens ist bekannt, dass 75% *die Qualitätsprüfung bestehen. Aufgrund von langjähriger Erfahrung ist weiter bekannt, dass*

60% der Produkte des Unternehmens, welche die Qualitätsprüfung bestehen, auf dem Markt Erfolg haben. Weiter geht das Unternehmen davon aus, dass 20% der Produkte, welche die Qualitätsprüfung nicht bestanden haben, gute Absatzchancen hätten. Mit A sei das Ereignis „Gute Absatzchancen“ und mit Q das Ereignis „Bestehen der internen Qualitätsprüfung“ bezeichnet.

Wie groß ist die Wahrscheinlichkeit, dass ein Produkt, welches gute Absatzchancen hat, die interne Qualitätsprüfung besteht?

$$\boldsymbol{P}(Q) = 0.75 \qquad \boldsymbol{P}(A|Q) = 0.6 \qquad \boldsymbol{P}(A|\overline{Q}) = 0.2$$

$$\begin{aligned} \boldsymbol{P}(Q|A) &= \frac{\boldsymbol{P}(A|Q)\cdot\boldsymbol{P}(Q)}{\boldsymbol{P}(A|Q)\cdot\boldsymbol{P}(Q)+\boldsymbol{P}(A|\overline{Q})\cdot\boldsymbol{P}(\overline{Q})} \\ &= \frac{0.6\cdot 0.75}{0.6\cdot 0.75+0.2\cdot 0.25} = 0.9 \end{aligned}$$

Dieses Ergebnis bestätigt die Vermutung des Unternehmens, dass ein Produkt, welches gute Absatzchancen hat, mit einer hohen Wahrscheinlichkeit die interne Qualitätsprüfung besteht.

Zu beachten ist, dass die a-posteriori Wahrscheinlichkeit von der Angabe der a-priori Wahrscheinlichkeit beeinträchtigt wird und in manchen Fällen sogar eine extreme Beeinflussung vorliegen kann.

Beispiel 1.5.16:

Wird im vorherigen Beispiel die a-priori Wahrscheinlichkeit mit $\boldsymbol{P}(Q) = 0.25$ angegeben, so ergibt sich für die Wahrscheinlichkeit, dass ein Produkt, welches gute Absatzchancen hat, die interne Qualitätsprüfung besteht, lediglich ein Wert von 0.5.

$$\boldsymbol{P}(Q) = 0.25 \qquad \boldsymbol{P}(A|Q) = 0.6 \qquad \boldsymbol{P}(A|\overline{Q}) = 0.2$$

$$\boldsymbol{P}(Q|A) = \frac{0.6\cdot 0.25}{0.6\cdot 0.25+0.2\cdot 0.75} = 0.5$$

Beispiel 1.5.17:

Eine Villa ist durch eine Alarmanlage gesichert. Wird dort eingebrochen, gibt diese mit einer Wahrscheinlichkeit von 0.99 Alarm. Mit einer Wahrscheinlichkeit von 0.01 tritt ein Fehlalarm auf. Die Wahrscheinlichkeit *für einen Einbruch wird bei* 0.001 *bzw.* 0.05 *vermutet. Das Ereignis* E *gibt das Ereignis „Es wird eingebrochen" an und mit* A *wird das Ereignis „Die Alarmanlage gibt Alarm" bezeichnet. Gesucht ist die Wahrscheinlichkeit, dass bei Auftreten eines Alarms tatsächlich ein Einbruch vorliegt.*

Fall 1:

$$\boldsymbol{P}(E) = \mathbf{0.001} \qquad \boldsymbol{P}(A|E) = 0.99 \qquad \boldsymbol{P}(A|\overline{E}) = 0.01$$

$$\begin{aligned} \boldsymbol{P}(E|A) &= \frac{\boldsymbol{P}(A|E) \cdot \boldsymbol{P}(E)}{\boldsymbol{P}(A|E) \cdot \boldsymbol{P}(E) + \boldsymbol{P}(A|\overline{E}) \cdot \boldsymbol{P}(\overline{E})} \\ &= \frac{0.99 \cdot 0.001}{0.99 \cdot 0.001 + 0.01 \cdot 0.999} = \mathbf{0.09} \end{aligned}$$

Fall 2:

$$\boldsymbol{P}(E) = \mathbf{0.05} \qquad \boldsymbol{P}(A|E) = 0.99 \qquad \boldsymbol{P}(A|\overline{E}) = 0.01$$

$$\boldsymbol{P}(E|A) = \frac{0.99 \cdot 0.05}{0.99 \cdot 0.05 + 0.01 \cdot 0.95} = \mathbf{0.84}$$

Auffallend ist, dass im Fall 1 $\boldsymbol{P}(E|A)$ *recht gering ist. Dieser Sachverhalt kann anhand einer* ***Vierfeldertafel*** *verdeutlicht werden.*

	E	$\overline{E}$	$\sum$
A	0.00099	0.00999	0.01098
$\overline{A}$	0.00001	0.98901	0.98902
$\sum$	0.001	0.999	1

Da die Wahrscheinlichkeit eines Einbruchs gering ist, geht die Alarmanlage viel öfter versehentlich an ($\boldsymbol{P}(A \cap \overline{E}) = 0.00999$) *als im Falle eines Einbruchs* ($\boldsymbol{P}(A \cap E) = 0.00099$). *In* 100000 *Fällen geht die Alarmanlage* 999 *mal versehentlich an und* 99 *mal, wenn tatsächlich eingebrochen wird. Der Anteil der Alarme, in denen tatsächlich ein Einbruch vorliegt, beträgt dann* $\frac{99}{1098} = 0.09$ $(= \frac{\boldsymbol{P}(A \cap E)}{\boldsymbol{P}(A \cap E) + \boldsymbol{P}(A \cap \overline{E})} = \boldsymbol{P}(E|A))$. *Dieses Ergebnis liefert auch die Formel von Bayes (Hinweis:* $\boldsymbol{P}(A \cap E) = \boldsymbol{P}(A|E) \cdot \boldsymbol{P}(E)$). *Im Gegensatz dazu geht die Alarmanlage mit einer hohen Wahrscheinlichkeit tatsächlich nicht an, wenn kein Einbruch stattfindet* $\boldsymbol{P}(\overline{A} \cap \overline{E}) = 0.98901$.

2 Zufallsvariablen und Wahrscheinlichkeitsverteilungen

2.1 Zufallsvariablen

In vielen Fällen sind nicht die detailliert angegebenen Ergebnisse eines Zufallsexperimentes von Interesse, sondern den Ergebnissen werden reelle Zahlen zugeordnet. Zu beachten ist dabei, dass die zugeordnete reelle Zahl durch den Ausgang des Zufallsexperimentes bestimmt wird und somit zufällig ist.

Beispiel 2.1.1:

a) Eine Münze wird zehnmal geworfen. Die genaue Folge von Kopf und Zahl ist dann häufig unwichtig. Vielmehr wird gefragt, wie oft Zahl (oder Kopf) aufgetreten ist.

b) Beim Werfen von mehreren Würfeln interessiert häufig nicht die Augenzahl der einzelnen Würfel, sondern nur die Augensumme.

Durch den Übergang zu reellen Zahlen können Wahrscheinlichkeiten und Kennwerte der deskriptiven Statistik (z.B. Lagemaße, Streuungsmaße) von verschiedenen Experimenten besser berechnet und besser miteinander verglichen werden.

Jedem Element ω des Ergebnisraumes Ω wird durch eine Funktion $X(\omega)$ eine reelle Zahl zugeordnet:

$$X : \omega \to X(\omega) \in \mathbb{R} \qquad (\mathbb{R}\text{: Menge der reellen Zahlen}).$$

Die Funktion $\boldsymbol{X}$ wird als **Zufallsvariable** oder auch als zufällige Variable bezeichnet. Die Variable X ist also eine Größe, die beim Auftreten jedes zufälligen Ereignisses ω einen davon abhängigen reellen Wert $X(\omega)$ annimmt, und somit vom Zufall abhängt. Die einzelnen Werte einer Zufallsvariablen heißen Realisationen und die Menge aller Realisationen geben den Definitionsbereich an.

Zufallsvariable

Im folgenden werden **Zufallsvariablen mit einem großen Buchstaben bezeichnet**. Der entsprechende **kleine Buchstabe** steht für die **Realisationen** der Zufallsvariablen.

Beispiel 2.1.2:

Für einen Münzwurf kann X die Menge der Ergebnisse $\{Kopf, Zahl\}$ *auf die Menge der Realisationen* $\{0, 1\}$ *abbilden:*

$$X(\omega) = \begin{cases} 0, & \text{wenn } \omega = \text{Kopf}, \\ 1, & \text{wenn } \omega = \text{Zahl}. \end{cases}$$

diskrete Zufallsvariable

Eine Zufallsvariable, die nur einzelne (höchstens abzählbar viele) Werte annehmen kann, wird als **diskret** bezeichnet.

Beispiel 2.1.3:

Eine diskrete Zufallsvariable liegt beim Würfelwurf vor, indem die Augenzahl beim Werfen eines Würfels betrachtet wird. Die Zufallsvariable X kann nur die Werte 1, 2, 3, 4, 5 *und* 6 *annehmen.*

stetige Zufallsvariable

Kann eine Zufallsvariable jeden Wert der reellen Zahlen annehmen oder kann sie innerhalb eines Intervalls der reellen Zahlen jeden Wert annehmen, so wird von einer **stetigen** Zufallsvariablen gesprochen.

Beispiel 2.1.4:

Wird die zufällig schwankende Lebensdauer von Glühlampen gemessen, liegt eine stetige Zufallsvariable vor. Die Lebensdauer kann in dem in Frage kommenden Bereich jeden beliebigen Wert annehmen.

Zufallsvariablen treten in fast allen Bereichen der Wirtschaftswissenschaften auf. Dabei ist in vielen Fragestellungen nicht sofort erkennbar, dass es sich um ein Zufallsexperiment handelt. In vielen Fällen handelt es sich vor allem um Zufallsvariablen, weil aus unterschiedlichen Gründen, alle Einflüsse einer bestimmten Größe nicht ermittelt und in einen Zusammenhang eingeordnet werden können.

Beispiel 2.1.5:

a) Die von einem Gut in einer Periode nachgefragte Menge hängt von verschiedenen Einflussgrößen ab. Zu diesen Einflüssen auf die Nachfragemenge gehören z.B. der Preis des Gutes, die Preise konkurrierender Güter, das verfügbare Einkommen der nachfragenden Haushalte. Damit sind aber nicht alle Einflüsse angegeben, und häufig werden bei der Analyse der Nachfrage eines Gutes noch nicht einmal diese Einflussgrößen alle berücksichtigt. Im einfachsten Fall *wird die nachgefragte Menge nur zu dem Preis des betreffenden Gutes in Beziehung gesetzt und somit ergibt sich dann die Nachfragefunktion. Die Vielzahl der übrigen Einflussgrößen führt dazu, dass die tatsächliche Nachfrage Schwankungen unterliegt, und somit zumindest mit einem einfachen Modellansatz nicht mehr erklärt werden kann. Zum Teil wirken diese Einflüsse sehr unregelmäßig, so dass die tatsächliche Nachfrage Schwankungen unterliegt, die keinerlei Gesetzmäßigkeiten erkennen lassen. Dabei wird von Zufallsschwankungen gesprochen. Die nachgefragte Menge ist damit eine Zufallsvariable.*

b) In der Produktionstheorie werden Produktionsvorgänge durch Produktionsfunktionen beschrieben. Die produzierte Menge wird dabei in Beziehung gesetzt zu den eingesetzten Mengen der Produktionsfaktoren. Selbst wenn es gelingt, die technischen Eigenheiten des Produktionsprozesses in einer Produktionsfunktion einzufangen, wird durch die Produktionsfunktion in vielen Fällen dennoch keine eindeutige Abhängigkeit zwischen produzierter Menge und Einsatzmengen der Produktionsfaktoren hergestellt werden können. In einem Produktionsprozess wirken im Allgemeinen zahlreiche unkontrollierbare Einflüsse mit. Dazu können z.B. Luftfeuchtigkeit, Luftbewegungen, Staubpartikel in der Luft, Schwankungen in den Materialeigenschaften usw. gehören. Das alles führt dazu, dass das Produktionsergebnis zumindest in einem abgegrenzten Intervall zufällig schwankt. Die produzierte Menge ist bei gegebenen Mengen der Produktionsfaktoren eine Zufallsvariable.

2.2 Wahrscheinlichkeitsfunktion, Dichtefunktion und Verteilungsfunktion

2.2.1 Diskrete Zufallsvariablen

Wahrscheinlichkeitsverteilung

Für diskrete Zufallsvariablen ergeben die Werte der Zufallsvariablen und die ihnen zugeordneten Wahrscheinlichkeiten zusammen die **Wahrscheinlichkeitsverteilung**. Zwischen einer Häufigkeitsverteilung im Bereich der deskriptiven Statistik und einer Wahrscheinlichkeitsverteilung besteht eine starke Analogie.

Ist X eine diskrete Zufallsvariable, die die Werte $x_1, ..., x_i, ..., x_n$ mit den Wahrscheinlichkeiten $\boldsymbol{P}(X = x_i) = \boldsymbol{P}(x_i)$ annimmt $(i = 1, ..., n)$, dann kann für diese Wahrscheinlichkeit auch $f_X(x_i)$ geschrieben werden.

$$\boxed{f_X(x_i) = \boldsymbol{P}(X = x_i) = \boldsymbol{P}(x_i)}$$

Die Wahrscheinlichkeitsverteilung der Zufallsvariablen kann in tabellarischer Form angegeben werden.

x_i	x_1	x_2	...	x_n
$f_X(x_i)$	$f_X(x_1)$	$f_X(x_2)$	...	$f_X(x_n)$

Beispiel 2.2.1:

Beim Werfen von zwei Münzen ergibt sich für die Wahrscheinlichkeitsverteilung der Anzahl X der Würfe, bei denen Zahl oben gelegen hat, folgende Werte.

x_i	0	1	2
$f_X(x_i)$	0.25	0.5	0.25

In manchen Fällen können die Wahrscheinlichkeiten $f_X(x_i)$ auch in Form einer funktionalen Abhängigkeit dargestellt werden.

Beispiel 2.2.2:

a) In einer Urne befinden sich 100 *Kugeln, die mit den Zahlen von* 1 *bis* 100 *beschriftet sind. Wird eine Kugel zufällig gezogen, dann beträgt die Wahrscheinlichkeit, eine Kugel mit der Zahl i* $(i = 1, ..., 100)$ *zu ziehen,*

$$f_X(i) = \frac{1}{100}.$$

b) Wird eine Münze n-mal geworfen, dann ist die Wahrscheinlichkeit dabei x-mal Zahl zu werfen gegeben durch

$$f_X(x) = \binom{n}{x}\left(\frac{1}{2}\right)^n; \quad x = 0, ..., n.$$

Die tabellarisch oder allgemein angegebene Zuordnungsvorschrift, durch die jedem Wert der **diskreten** Zufallsvariablen die entsprechende Wahrscheinlichkeit zugeordnet wird, heißt **Wahrscheinlichkeitsfunktion $f_X(x)$**.

Wahrscheinlichkeitsfunktion

Aus der Definition der Wahrscheinlichkeit folgt, dass die Wahrscheinlichkeitsfunktion die folgenden beiden Eigenschaften besitzt:

$$0 \leq f_X(x_i) \leq 1 \qquad \text{und} \qquad \sum_{i=1}^{n} f_X(x_i) = 1.$$

Die Wahrscheinlichkeitsverteilung einer diskreten Zufallsvariablen kann mit der Häufigkeitsverteilung für relative Häufigkeiten eines diskreten Merkmals verglichen werden. Die beiden angegebenen Eigenschaften für die Wahrscheinlichkeitsfunktion sind Eigenschaften, die auch für relative Häufigkeiten zutreffen.

Die grafische Darstellung einer Wahrscheinlichkeitsfunktion $f_X(x)$ einer diskreten Zufallsvariablen X kann in Form eines Stab- oder Säulendiagramms erfolgen (s. Abbildung 2.2.1).

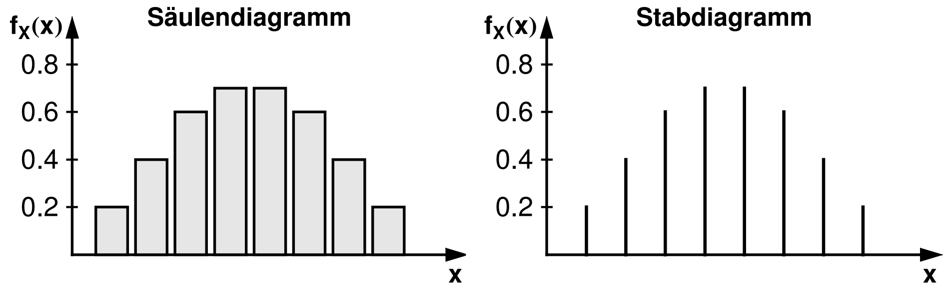


Abbildung 2.2.1: *Säulen- und Stabdiagramm zur Darstellung einer Wahrscheinlichkeitsfunktion*

In der deskriptiven Statistik wurde neben der Häufigkeitsverteilung auch die Verteilung der Summenhäufigkeiten zur Beschreibung der Verteilung eines Merkmals eingeführt. Die Summenhäufigkeitsverteilung in der deskriptiven Statistik entspricht der Verteilungsfunktion einer Zufallsvariablen.

Verteilungsfunktion

Die **Verteilungsfunktion $\boldsymbol{F_X(x)}$** einer Zufallsvariablen X ist definiert durch:

$$F_X(x) = \boldsymbol{P}(X \leq x) = \boldsymbol{P}(-\infty < X \leq x).$$

Für eine **diskrete** Zufallsvariable ist die Verteilungsfunktion somit gegeben durch:

$$F_X(x) = \boldsymbol{P}(X \leq x) = \sum_{x_i \leq x} f_X(x_i).$$

Die **Verteilungsfunktion** F_X einer Zufallsvariablen X besitzt folgende **Eigenschaften:**

(1) F_X ist monoton steigend.

Diese Eigenschaft kann leicht plausibel gemacht werden. Die Verteilungsfunktion entsteht durch sukzessives Aufaddieren der Wahrscheinlichkeiten. Wahrscheinlichkeiten können aber keine negativen Werte annehmen. Das bedeutet, dass die Verteilungsfunktion niemals fallen kann. Diese Eigenschaft ist auch schon von der Summenhäufigkeitsverteilung aus der deskriptiven Statistik bekannt.

(2) F_X ist rechtsseitig stetig,

d.h. es gilt $\lim\limits_{c \to 0} F_X(x + c) = F_X(x)$.

Diese Eigenschaft ergibt sich aus der Tatsache, dass die Verteilungsfunktion definiert ist als $F_X(x) = \boldsymbol{P}(X \leq x)$. Wird die Verteilungsfunktion durch $\boldsymbol{P}(X < x)$ definiert, dann heißt sie linksseitig stetig.

Anschaulich bedeutet die rechtsseitige Stetigkeit:

Hat die Verteilungsfunktion an einer Stelle x einen Sprung (Unstetigkeitsstelle), dann nimmt die Verteilungsfunktion an dieser Stelle den Wert an, der zu der oberen Sprunggrenze gehört. In grafischen Darstellungen wird dies verdeutlicht, indem die entsprechende Sprunggrenze hervorgehoben wird. (Vgl. dazu die Verteilungsfunktion in Abbildung 2.2.2)

(3) $\lim\limits_{x \to -\infty} F_X(x) = 0$.

Diese Eigenschaft bedeutet, dass die Verteilungsfunktion gegen 0 geht, wenn x immer kleiner wird.

(4) $\lim\limits_{x \to \infty} F_X(x) = 1$.

Wenn x immer größer wird, geht der Wert der Verteilungsfunktion schließlich gegen 1.

Die angegebenen Eigenschaften einer Verteilungsfunktion gelten auch bei stetigen Zufallsvariablen, auf die im nächsten Abschnitt eingegangen wird. Es ist zu beachten, dass die Verteilungsfunktion grundsätzlich für die ganze Menge der reellen Zahlen definiert ist. Zu einer korrekten Beschreibung einer Verteilungsfunktion gehört deshalb die Angabe der Funktion für den gesamten Definitionsbereich. Bei diskreten Zufallsgrößen führt das dazu, dass üblicherweise eine intervallweise Beschreibung erfolgen muss.

Beispiel 2.2.3:

Beim Werfen eines Würfels ergibt sich als Wahrscheinlichkeitsverteilung:

x_i	$x_1 = 1$	$x_2 = 2$	$x_3 = 3$
$f_X(x_i)$	$f_X(x_1) = \frac{1}{6}$	$f_X(x_2) = \frac{1}{6}$	$f_X(x_3) = \frac{1}{6}$
	$x_4 = 4$	$x_5 = 5$	$x_6 = 6$
	$f_X(x_4) = \frac{1}{6}$	$f_X(x_5) = \frac{1}{6}$	$f_X(x_6) = \frac{1}{6}$

Da die Verteilungsfunktion für die einzelnen Intervalle verschiedene Werte annimmt, muss eine intervallweise Beschreibung erfolgen.

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

Abbildung 2.2.2 zeigt eine grafische Darstellung von Wahrscheinlichkeitsfunktion $f_X(x)$ und Verteilungsfunktion $F_X(x)$ der Zufallsvariablen „Augenzahl X beim Würfeln mit einem Würfel“.

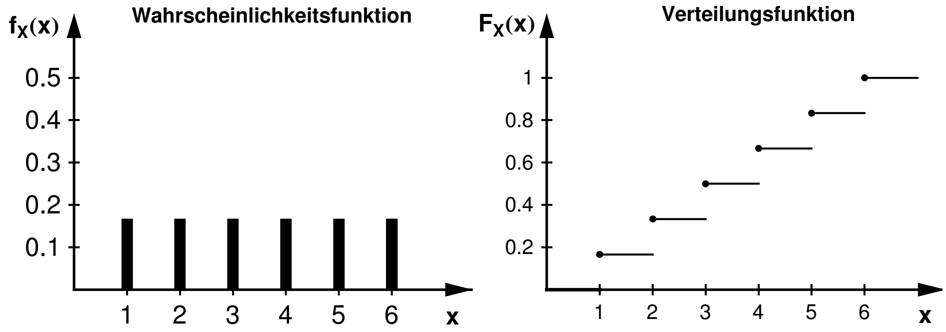


Abbildung 2.2.2: *Wahrscheinlichkeitsfunktion und Verteilungsfunktion eines Würfelwurfs*

Die grafische Darstellung der Verteilungsfunktion einer diskreten Zufallsvariablen ergibt immer eine Treppenfunktion.

2.2.2 Stetige Zufallsvariablen

Eine stetige Zufallsvariable kann innerhalb eines reellen Intervalls jeden beliebigen Wert annehmen. Von daher ist die Wahrscheinlichkeit, dass eine stetige Zufallsvariable exakt einen bestimmten Wert annimmt, gleich Null. Wahrscheinlichkeiten können für stetige Zufallsvariablen somit nur für Intervalle berechnet werden. Die Wahrscheinlichkeitsverteilung einer stetigen Zufallsvariablen kann approximativ mittels eines Histogramms dargestellt werden, wobei der Flächeninhalt eines vorliegenden Rechtecks proportional zur Wahrscheinlichkeit ist, dass die Zufallsvariable in das entsprechende Intervall fällt. Für immer kleiner werdende Intervalle wird ersichtlich, dass die Wahrscheinlichkeitsverteilung einer stetigen Zufallsvariablen durch eine glatte Kurve, der sogenannten **Dichtefunktion**, dargestellt wird (s. Abbildung 2.2.3).

Die Wahrscheinlichkeit, dass die Zufallsvariable in ein beliebiges Intervall fällt, wird mit der entsprechenden Fläche unterhalb der glatten Kurve angegeben. Diese Fläche wird über das Integral berechnet. D.h. anstelle der Summe der Rechteckflächen wird das Integral verwendet, welches als unendliche Summe von vielen kleinen Rechteckflächen aufgefasst werden kann.

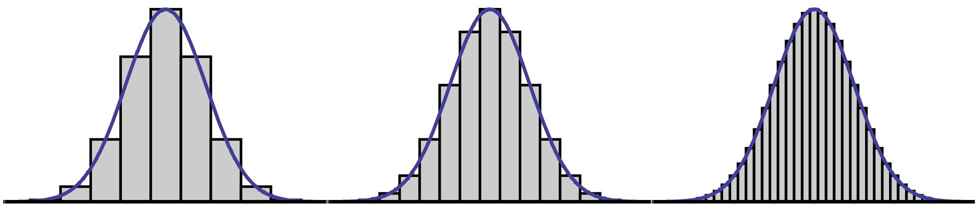


Abbildung 2.2.3: *Approximation der Dichtefunktion einer stetigen Zufalssvariablen durch Histogramme*

Für eine **Dichtefunktion $\boldsymbol{f_X(x)}$** einer **stetigen** Zufallsvariablen X gilt:

Dichtefunktion

$$\begin{aligned} \int_{-\infty}^{\infty} f_X(x)dx &= 1 && \text{(Normierungseigenschaft).} \\ f_X(x) &\geq 0 \quad \text{für alle } x \in \mathbb{R} && \text{(Nicht-Negativitätseigenschaft).} \end{aligned}$$

Die Wahrscheinlichkeit, dass die stetige Zufallsvariable X zwischen zwei reellen Zahlen a und b liegt, wird wie folgt angegeben:

$$\boldsymbol{P}(a \leq X \leq b) = \boldsymbol{P}(a < X \leq b) = \boldsymbol{P}(a \leq X < b) = \int_a^b f_X(x)dx.$$

Die Wahrscheinlichkeit $\boldsymbol{P}(a \leq X \leq b)$ entspricht der Fläche unter der Dichtefunktion $f_X(x)$ von dem Punkt a bis zum Punkt b. Die in Abbildung 2.2.4 dargestellte Dichtefunktion zeigt, dass die Wahrscheinlichkeit, dass die Zufallsvariable einen Wert in einem Intervall am Rand annimmt geringer ist als in einem Intervall im mittleren Bereich, d.h. die Fläche unterhalb der Kurve über ein Intervall im mittleren Bereich ist größer als die Fläche unterhalb der Kurve über ein gleichbreites Intervall im Randbereich. Die Werte liegen somit *dichter* im mittleren Bereich. Daher der Name Dichtefunktion.

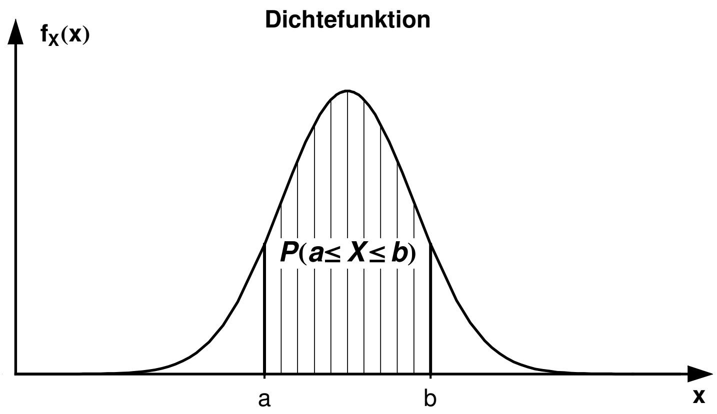


Abbildung 2.2.4: *Dichtefunktion einer stetigen Zufallsvariablen*

Aus der Interpretation der Wahrscheinlichkeit als Fläche über einem Intervall unterhalb der Dichtefunktion folgt direkt, dass die Punktwahrscheinlichkeit, dass eine Zufallsvariable einen bestimmten, vorgegebenen Wert (Punkt) x_0 annimmt, gleich Null ist.

$$\boldsymbol{P}(X = x_0) = \boldsymbol{P}(x_0 \leq X \leq x_0) = \int_{x_0}^{x_0} f_X(x)dx = 0$$

Ein einzelner Wert stellt nämlich ein „Intervall der Länge 0“ dar. Die von der Dichtefunktion und diesem „Intervall“ eingeschlossene Fläche hat daher die Größe 0. Aufgrund der Tatsache, dass die Punktwahrscheinlichkeit einer stetigen Zufallsvariablen den Wert 0 annimmt, gilt die Gleichheit der Beziehung $\boldsymbol{P}(a \leq X \leq b) = \boldsymbol{P}(a < X \leq b) = \boldsymbol{P}(a \leq X < b)$.

Beispiel 2.2.4:

Wird die Geschwindigkeit eines PKW im Stadtverkehr als Zufallsvariable aufgefasst, so ist die Wahrscheinlichkeit, dass der Tachometer eine Geschwindigkeit von exakt 30.0 *km/h anzeigt, gleich Null. Dennoch ist das Ereignis „Die Geschwindigkeitsmessung ergibt* 30.0 *km/h“ nicht* *unmöglich. Beispielsweise muss jeder PKW, der schneller als* 30.0 *km/h fährt, beim Beschleunigen diese Geschwindigkeit irgendwann erreicht und überschritten haben.*

Diese Ausführungen zeigen, dass aus der Tatsache, dass die Wahrscheinlichkeit für ein bestimmtes Ereignis gleich Null ist, nicht auf die Unmöglichkeit dieses Ereignisses geschlossen werden darf. Andererseits ist die Wahrscheinlichkeit für das „unmögliche Ereignis“ gleich 0.

Ist die Dichtefunktion nur in einem endlichen Intervall positiv,

$$f^*_X(x) > 0 \qquad \text{für} \qquad x_u \le x \le x_o,$$

so lautet die Schreibweise

$$f_X(x) = \begin{cases} f^*_X(x) & \text{für } x_u \le x \le x_o \\ 0 & \text{sonst.} \end{cases}$$

Es ist dann

$$\int_{x_u}^{x_o} f^*_X(x) dx = 1.$$

Auch für stetige Zufallsvariablen existiert eine Verteilungsfunktion. Die Verteilungsfunktion ist hier grundsätzlich so definiert, wie für eine diskrete Zufallsvariable, d.h.

$$F_X(x) = \boldsymbol{P}(X \le x).$$

Für eine stetige Zufallsvariable kann jedoch die Verteilungsfunktion nicht mehr durch Summation ermittelt werden, sondern sie wird durch ein Integral mit variabler oberer Grenze x bestimmt.

Verteilungsfunktion

Die **Verteilungsfunktion $\boldsymbol{F_X(x)}$** einer **stetigen** Zufallsvariablen X besitzt die gleichen Eigenschaften wie die einer diskreten Zufallsvariablen. Sie lautet:

$$F_X(x) = \int_{-\infty}^{x} f_X(\xi)\ d\xi.$$

Das ξ (sprich: xi) ist das kleine x des griechischen Alphabets. Zur Bestimmung der Verteilungsfunktion wird die Stammfunktion der Dichtefunktion gebildet. Ist $F_X(x)$ differenzierbar, so gilt $F'_X(x) = f_X(x)$.

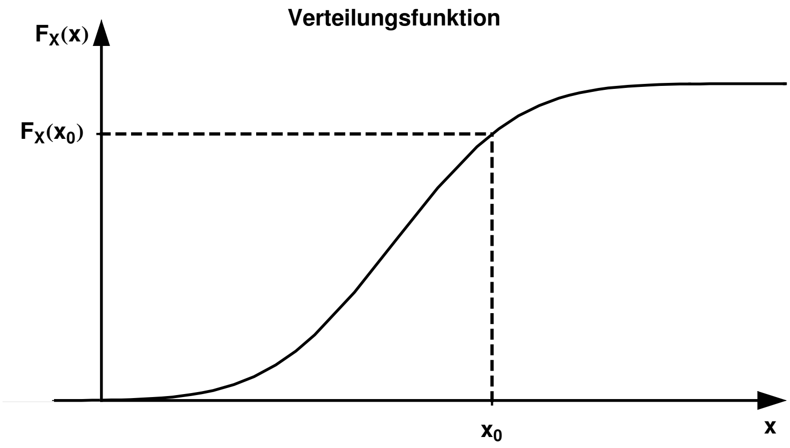


Abbildung 2.2.5: *Verteilungsfunktion einer stetigen Zufallsvariablen*

Der Wert $F_X(x_0)$ entspricht der Wahrscheinlichkeit $\boldsymbol{P}(X \leq x_0)$ und lässt sich als Fläche unter der Dichtefunktion $f_X(x)$ bis zum Punkt x_0 interpretieren (s. Abbildung 2.2.6).

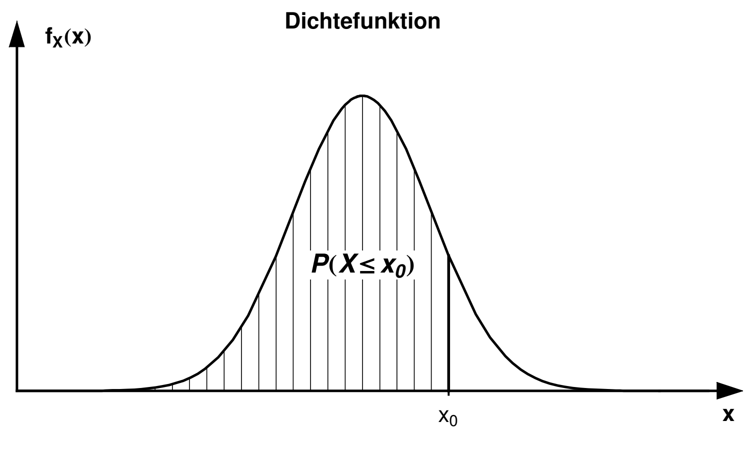


Abbildung 2.2.6: *Darstellung von $F_X(x_0) = \boldsymbol{P}(X \leq x_0)$ mittels der Dichtefunktion*

Für die Bestimmung von Verteilungsfunktionen ist zu beachten, dass die Grenzen so angegeben werden, dass es sich um eine rechtsstetige Verteilungsfunktion handelt. Darauf ist bei Dichtefunktionen mit Sprungstellen zu achten.

Beispiel 2.2.5:

Rechteckverteilung (stetige Gleichverteilung)

Es sei die Dichtefunktion

$$f_X(x) = \begin{cases} \frac{1}{2} & \text{für } 3 \leq x \leq 5 \\ 0 & \text{sonst} \end{cases}$$

gegeben. Als Verteilungsfunktion ergibt sich für $3 \leq x < 5$

$$F_X(x) = \int_{-\infty}^{x} f_X(\xi)\ d\xi = \int_{3}^{x} \frac{1}{2}\ d\xi = \left[\frac{1}{2}\xi\right]_3^x = \frac{1}{2}x - \frac{3}{2}.$$

(Es ist $\int_{-\infty}^{3} f_X(\xi)\ d\xi = 0$*, da für* $x < 3$ $f_X(x) = 0$ *gilt.)*

Für $x < 3$ *ist* $F_X(x) = 0$ *und für* $5 \leq x$ *ist* $F_X(x) = 1$*.*

Die Verteilungsfunktion lautet somit

$$F_X(x) = \begin{cases} 0 & \text{für} \quad x < 3 \\ \frac{1}{2}x - \frac{3}{2} & \text{für } 3 \leq x < 5 \\ 1 & \text{für } 5 \leq x. \end{cases}$$

Da die Dichtefunktion der Rechteckverteilung im Definitionsbereich keine Sprungstellen aufweist, kann die Verteilungsfunktion unter dem Aspekt der Rechtsstetigkeit äquivalent mittels folgender Alternativen angegeben werden:

$$F_X(x) = \begin{cases} 0 & \text{für} \quad x \leq 3 \\ \frac{1}{2}x - \frac{3}{2} & \text{für } 3 < x < 5 \\ 1 & \text{für } 5 \leq x \end{cases} \qquad F_X(x) = \begin{cases} 0 & \text{für} \quad x \leq 3 \\ \frac{1}{2}x - \frac{3}{2} & \text{für } 3 < x \leq 5 \\ 1 & \text{für } 5 < x \end{cases}$$

$$F_X(x) = \begin{cases} 0 & \text{für} \quad x \leq 3 \\ \frac{1}{2}x - \frac{3}{2} & \text{für } 3 \leq x \leq 5 \\ 1 & \text{für } 5 \leq x \end{cases}$$

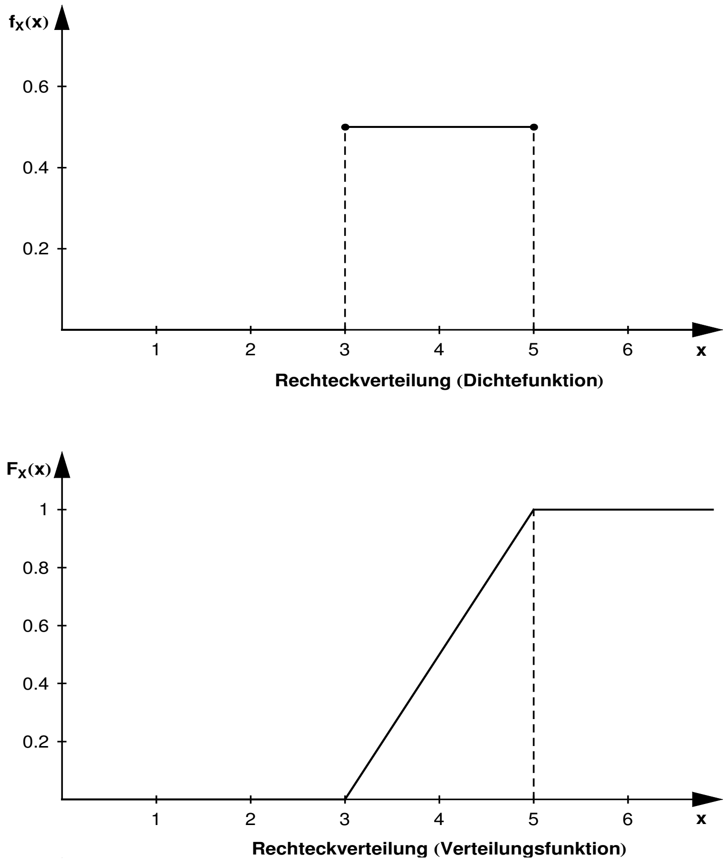


Abbildung 2.2.7: *Dichte- und Verteilungsfunktion zu Beispiel 2.2.5*

Beispiel 2.2.6:

Dreiecksverteilung

$$f_X(x) = \begin{cases} 0.25x - 0.5 & \text{für } 2 \leq x \leq 4 \\ -0.25x + 1.5 & \text{für } 4 \leq x \leq 6 \\ 0 & \text{sonst.} \end{cases}$$

$$F_X(x) = \begin{cases} 0 & \text{für } \quad x < 2 \\ 0.125x^2 - 0.5x + 0.5 & \text{für } 2 \leq x < 4 \\ -0.125x^2 + 1.5x - 3.5 & \text{für } 4 \leq x < 6 \\ 1 & \text{für } 6 \leq x. \end{cases}$$

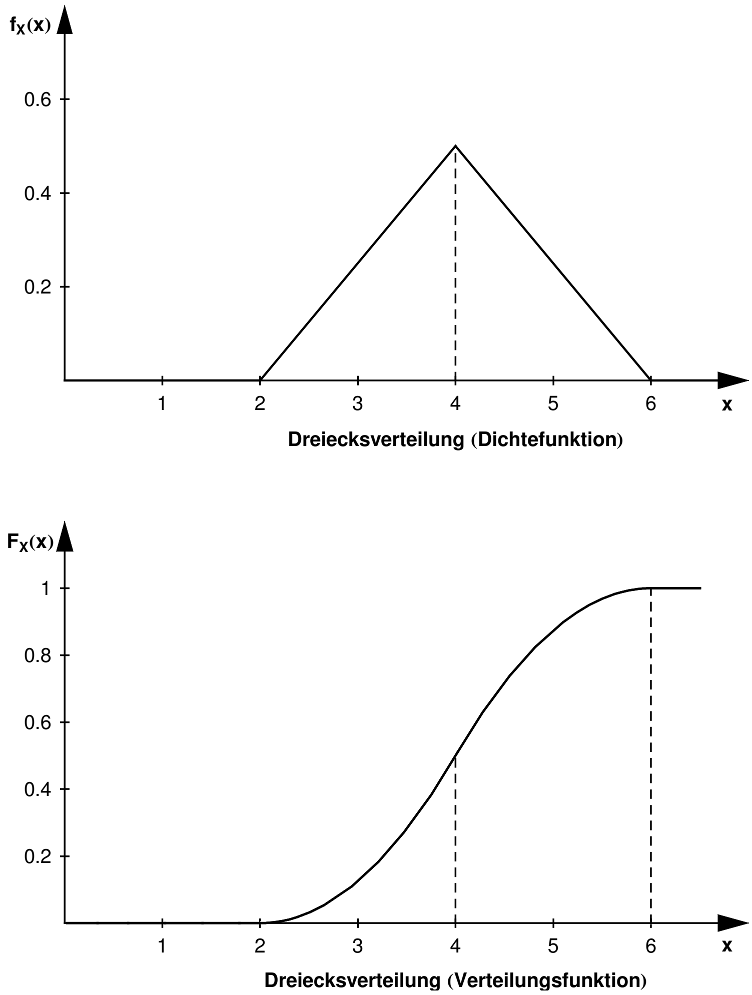


Abbildung 2.2.8: *Dichte- und Verteilungsfunktion zu Beispiel 2.2.6*

Mit der Angabe der Dichte- **oder** der Verteilungsfunktion kann auf zwei verschiedenen Wegen die Wahrscheinlichkeit bestimmt werden, dass die stetige Zufallsvariable X einen Wert in einem bestimmten Intervall $[a;b]$ annimmt.

$$\begin{aligned} \boldsymbol{P}(a \leq X \leq b) &= \int_a^b f_X(x)dx = \int_{-\infty}^b f_X(x)dx - \int_{-\infty}^a f_X(x)dx \\ &= F_X(b) - F_X(a) \end{aligned}$$

Beispiel 2.2.7:

$$f_X(x) = \begin{cases} 2x-2 & \text{für } 1 \le x \le 2 \\ 0 & \text{sonst} \end{cases}$$

$$F_X(x) = \begin{cases} 0 & \text{für } \quad x < 1 \\ x^2-2x+1 & \text{für } 1 \le x < 2 \\ 1 & \text{für } 2 \le x \end{cases}$$

$$\begin{aligned} \boldsymbol{P}(1 \le x \le 2) &= \int_1^2 (2x-2)dx = \Big[x^2-2x\Big]_1^2 \\ &= 4-4-1+2 = 1 \end{aligned}$$

$$\boldsymbol{P}(1 \le x \le 2) = F_X(2) - F_X(1) = 1-(1-2+1) = 1$$

$$\boldsymbol{P}(1.5 \le x \le 2) = \Big[x^2-2x\Big]_{1.5}^2 = 4-4-\frac{9}{4}+3 = 0.75$$

$$\boldsymbol{P}(1.5 \le x \le 2) = F_X(2) - F_X(1.5) = 1-\left(\frac{9}{4}-3+1\right) = 0.75$$

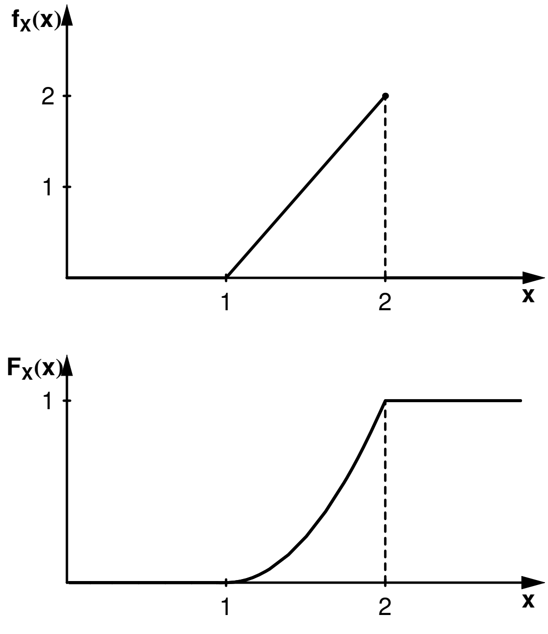


Abbildung 2.2.9: *Dichte- und Verteilungsfunktion zu Beispiel 2.2.7*

Ist eine Dichte nur für $x_u \leq X \leq x_o$ positiv (in diesem Bereich wird die Dichte mit $f_X^*(x)$ bezeichnet), so gelten für die Verteilungsfunktion nachstehende Überlegungen:

$x < x_u$:

$$F_X(x) = \int_{-\infty}^{x} f_X(\xi)\; d\xi = \int_{-\infty}^{x} 0\; d\xi = 0$$

$x_u \leq x < x_o$:

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} f_X(\xi)\; d\xi = \int_{-\infty}^{x_u} 0\; d\xi + \int_{x_u}^{x} f_X^*(\xi)\; d\xi \\ &= 0 + \int_{x_u}^{x} f_X^*(\xi)\; d\xi = \int_{x_u}^{x} f_X^*(\xi)\; d\xi \end{aligned}$$

$x_0 \leq x$:

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} f_X(\xi)\; d\xi = \int_{-\infty}^{\infty} f_X(\xi)\; d\xi - \int_{x}^{\infty} 0\; d\xi \\ &= 1 - 0 = 1 \end{aligned}$$

Für eine Dichtefunktion, die nur zwischen den Werten x_u und x_o von 0 verschieden und im übrigen Bereich 0 ist, gilt also:

- Für $x < x_u$ nimmt die Verteilungsfunktion den Wert 0 an.
- Für das Intervall mit den Grenzen x_u und x_o wird die Verteilungsfunktion als Integral der Dichtefunktion bestimmt und zwar mit x_u als unterer Grenze und x als variabler oberer Grenze.
- Ist $x \geq x_o$, so nimmt die Verteilungsfunktion den Wert 1 an.

Anmerkung:

Für Dichte- bzw. Wahrscheinlichkeits- und Verteilungsfunktionen sind auch die Bezeichnungen $f(x)$ bzw. $f(x_i)$ und $F(x)$ üblich, falls Verwechslungen ausgeschlossen sind. Die Indizierung mit „X“ kann dann entfallen. Auf der Homepage des Lehrstuhls befinden sich unter http://www.fernuni-hagen.de/ls_statistik/forschung/multimedia/ Mathematica-Applets, die einen Vergleich zwischen der empirischen und theoretischen Verteilung bzw. die Konvergenz der empirischen Verteilungsfunktion zeigen.

2.3 Parameter von Wahrscheinlichkeitsverteilungen

Zur Charakterisierung von Häufigkeitsverteilungen wurden in der deskriptiven Statistik Lage- und Streuungsmaße bestimmt. In ähnlicher Weise können für Wahrscheinlichkeitsverteilungen Lage- und Streuungsparameter bestimmt werden.

2.3.1 Erwartungswert

Als wichtigstes Lagemaß der empirischen Häufigkeitsverteilung eines quantitativen Merkmals wurde das arithmetische Mittel angegeben. Der Erwartungswert $E(X)$ einer Zufallsvariablen X ist jener Wert, der sich bei *genügend häufiger* Wiederholung des Zufallsexperimentes als Mittelwert der Ergebnisse ergeben würde. Im Vergleich zum arithmetischen Mittel, welches die tatsächliche Lage der Häufigkeitsverteilung beschreibt, ist der Erwartungswert eine theoretische Größe. Der Erwartungswert ist als Parameter zur Bestimmung der Lage einer Verteilung formal direkt mit dem arithmetischen Mittel vergleichbar. Dies ist aus der angegebenen Berechnungsformel ersichtlich.

Erwartungswert

Erwartungswert einer diskreten Zufallsvariablen:

$$E(X) = \sum_{i=1}^{n} x_i f_X(x_i)$$

Erwartungswert einer stetigen Zufallsvariablen:

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$$

Anstelle von $E(X)$ wird der Erwartungswert einer Zufallsvariablen X auch mit μ_X (oder auch nur μ) bezeichnet.

Für diskrete Zufallsvariablen berechnet sich der Erwartungswert als Summe der Produkte aus den möglichen Realisationen des Zufallsexperiments und der dazugehörigen Wahrscheinlichkeit. Der Erwartungswert kann einen Wert annehmen, der keinem möglichen Ergebnis des zugrundeliegenden Zufallsexperiments entspricht.

Beispiel 2.3.1:

Es wird mit einem Würfel gewürfelt. Die Wahrscheinlichkeit für die Augenzahl i beträgt dann $f_X(i) = \frac{1}{6}$ *für* $i = 1, ..., 6$.

Der Erwartungswert beträgt

$$\begin{aligned} E(X) &= \sum_{i=1}^{6} i f_X(i) \\ &= 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5. \end{aligned}$$

Hier beträgt der Erwartungswert 3.5, *was keinem möglichen Ergebnis beim Würfelwurf entspricht.*

Beispiel 2.3.2:

Die Dichtefunktion der stetigen Zufallsvariable X wird mit

$$f_X(x) = \begin{cases} 0.125x - 0.25 & \text{für } 2 \leq x \leq 6 \\ 0 & \text{sonst} \end{cases}$$

angegeben. Diese Dichtefunktion ist in der folgenden Abbildung grafisch dargestellt.

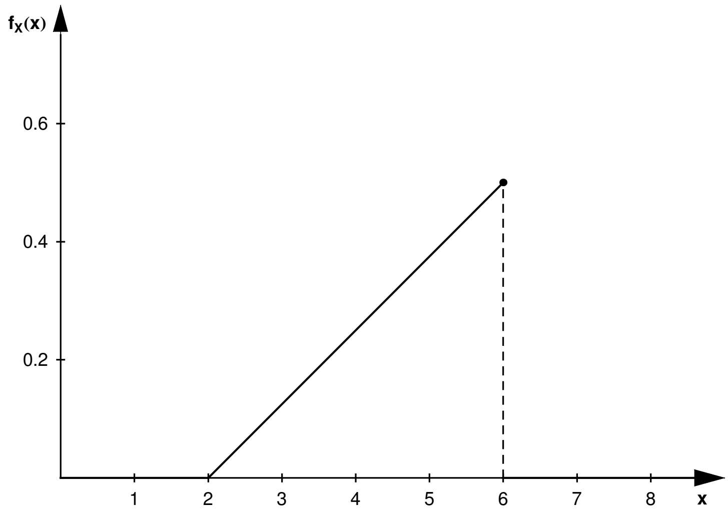


Abbildung 2.3.1: *Dichtefunktion zu Beispiel 2.3.2*

Für den Erwartungswert ergibt sich:

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f_X(x)dx = \int_2^6 x(0.125x - 0.25)dx \\ &= \int_2^6 (0.125x^2 - 0.25x)dx \\ &= \left[0.125 \cdot \frac{1}{3}x^3 - 0.25 \cdot \frac{1}{2}x^2\right]_2^6 \\ &= 9 - 4.5 - \frac{1}{3} + 0.5 = 4\frac{2}{3} = 4.\overline{6}. \end{aligned}$$

Der Erwartungswert spielt in den Anwendungen der Wahrscheinlichkeitsrechnung innerhalb der Wirtschaftswissenschaften eine sehr wichtige Rolle. Wenn in Entscheidungsproblemen die Zielgröße eine Zufallsvariable ist, dann entsteht grundsätzlich das Problem eines eindeutigen Entscheidungskriteriums. Anstelle der Zufallsvariablen kann der Erwartungswert der Zufallsvariablen für die Herleitung eines Entscheidungskriteriums verwendet werden.

2.3.2 Standardabweichung und Varianz

Als wichtigstes Streuungsmaß für die Häufigkeitsverteilung eines quantitativen Merkmals wurden in der deskriptiven Statistik die Varianz und die Standardabweichung eingeführt. Diese beiden Maße werden innerhalb der Wahrscheinlichkeitsrechnung als Parameter für die Streuung der Wahrscheinlichkeitsverteilung einer Zufallsvariablen verwendet. Die beiden Streuungsparameter sind ähnlich definiert wie in der deskriptiven Statistik.

Varianz

Varianz einer diskreten Zufallsvariablen:

$$\mathrm{Var}(X) = \sum_{i=1}^{n} [x_i - E(X)]^2 f_X(x_i) = \sum_{i=1}^{n} x_i^2 f_X(x_i) - [E(X)]^2$$

Varianz einer stetigen Zufallsvariablen:

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

Wird anstelle der Zufallsvariablen X die Zufallsvariable $[X - E(X)]^2 = (X - \mu)^2$ betrachtet, so entspricht die Varianz der erwarteten quadratischen Abweichung der Zufallsvariablen X vom Erwartungswert.

Anstelle von $\mathrm{Var}(X)$ wird die Varianz einer Zufallsvariablen X oft mit σ_X^2 (oder σ^2) bezeichnet.

Standardabweichung

Die **Standardabweichung** ist auch hier als Quadratwurzel aus der Varianz definiert.

$$\sigma_X = \sqrt{\mathrm{Var}(X)}$$

Anmerkung:

Die Existenz des Mittelwertes und der Varianz kann sowohl bei stetigen Zufallsvariablen als auch bei den diskreten Zufallsvariablen, die (abzählbar) unendlich viele Werte annehmen können, im Allgemeinen nicht garantiert werden. Es gibt wichtige Beispiele von Wahrscheinlichkeitsverteilungen, die einen oder beide Parameter nicht besitzen. Das gleiche gilt natürlich sinngemäß für die im nächsten Abschnitt dargestellten allgemeineren Momente.

Beispiel 2.3.3:

Für das Würfelbeispiel (vgl. Beispiel 2.3.1) ergibt sich:

$$\begin{aligned}
\mathrm{Var}(X) &= \sum_{i=1}^{6}(i-3.5)^2\frac{1}{6} \\
&= 2.5^2\cdot\frac{1}{6} + 1.5^2\cdot\frac{1}{6} + 0.5^2\cdot\frac{1}{6} + 0.5^2\cdot\frac{1}{6} + 1.5^2\cdot\frac{1}{6} + 2.5^2\cdot\frac{1}{6} \\
&= \frac{1}{6}(6.25 + 2.25 + 0.25 + 0.25 + 2.25 + 6.25) \\
&= \frac{1}{6}\cdot 17.5 = 2.91\overline{6}, \\
\sigma_X &= \sqrt{2.916} = 1.708.
\end{aligned}$$

Beispiel 2.3.4:
Für ein stetiges Merkmal X mit der Dichtefunktion

$$f_X(x) = \begin{cases} \frac{1}{2} & \text{für } 3 \leq x \leq 5 \\ 0 & \text{sonst} \end{cases}$$

berechnet sich die Varianz und Standardabweichung zu

$$\begin{aligned} \mathrm{Var}(X) &= \int_3^5 (x-4)^2 \frac{1}{2} dx = \frac{1}{2} \int_3^5 (x^2 - 8x + 16) dx \\ &= \frac{1}{2} \left[\frac{1}{3} x^3 - 4x^2 + 16x \right]_3^5 \\ &= \frac{1}{2} \left[\frac{1}{3} \cdot 125 - 4 \cdot 25 + 16 \cdot 5 - \frac{1}{3} \cdot 27 + 4 \cdot 9 - 16 \cdot 3 \right] = \frac{1}{3}, \\ \sigma_X &= 0.58. \end{aligned}$$

2.3.3 Momente

Momente

Erwartungswert und Varianz sind Spezialfälle einer allgemeinen Klasse von Parametern zur Charakterisierung einer Wahrscheinlichkeitsverteilung durch eine Kenngröße. Es handelt sich hierbei um die **Momente** einer Zufallsvariablen. Unterschieden wird in

- Momente um 0 und
- Momente in Bezug auf einen Parameter a, wobei vor allem die **zentralen Momente**, die auf den Erwartungswert der Zufallsvariablen bezogen werden, interessieren.

k-tes Moment um Null

Das **k-te Moment um Null μ_k'** ist wie folgt definiert:

$$\begin{aligned} \mu_k' = E(X^k) &= \sum_{i=1}^{n} x_i^k f_X(x_i) \text{ im diskreten Fall bzw.} \\ \mu_k' = E(X^k) &= \int_{-\infty}^{\infty} x^k f_X(x) dx \text{ im stetigen Fall.} \end{aligned}$$

Der Erwartungswert ist also das erste Moment um Null.

Für das **zentrale Moment $\boldsymbol{k}$-ter Ordnung $\boldsymbol{\mu_k}$** gilt mit $E(X) = \mu$:

zentrales Moment k-ter Ordnung

$$\begin{aligned}
\mu_k = E\big[(X-\mu)^k\big] &= \sum_{i=1}^{n}(x_i-\mu)^k f_X(x_i) \text{ bzw.} \\
\mu_k = E\big[(X-\mu)^k\big] &= \int_{-\infty}^{\infty}(x-\mu)^k f_X(x)dx.
\end{aligned}$$

Die **Varianz** ist das **zentrale Moment zweiter Ordnung**. Für die Varianz gilt der Steinersche Verschiebungssatz:

zentrales Moment zweiter Ordnung (Varianz)

$$\begin{aligned}
\mathrm{Var}(X) &= E[(X-E[X])^2] = E[(X-\mu)^2] \\
&= E(X^2 - 2X\mu + \mu^2) = E(X^2) - 2\mu\mu + \mu^2 \\
&= E(X^2) - \mu^2 \\
&= E(X^2) - [E(X)]^2.
\end{aligned}$$

Mit Hilfe von Momenten können verschiedene Eigenschaften von eingipfligen Verteilungen wie z.B. die **Schiefe** (engl.: Skewness) und die **Wölbung (Kurtosis)** untersucht werden (s. empirische Schiefe, Wölbung, Kurseinheit 1). Diese Eigenschaften werden zur Charakterisierung einer Verteilung benutzt.

Schiefe
Wölbung (Kurtosis)

Ein Parameter für die Symmetrie einer Verteilung ist der **Momentenkoeffizient der Schiefe $\boldsymbol{\gamma_3}$**, welcher aus dem dritten zentralen Moment abgeleitet wird.

Momentenkoeffizient der Schiefe

$$\gamma_3 = \frac{E([X-E(X)]^3)}{\mathrm{Var}(X)^{\frac{3}{2}}} = \frac{E[(X-\mu)^3]}{\sigma^3}$$

Für symmetrische Verteilungen gilt $\gamma_3 = 0$. Bei linkssteilen/rechtssteilen Verteilungen nimmt γ_3 einen positiven/negativen Wert an. Weiter ist γ_3 invariant gegenüber linearer Transformationen, d.h. γ_3 der Zufallsvariabeln X entspricht γ_3 der Zufallsvariablen $aX + b$, $a > 0$.

Die Flachheit einer Verteilung, d.h. ob eine Verteilung flach oder steil verläuft, wird mittels der Wölbung beurteilt.

Kurtosis

Ein Parameter für die Wölbung einer Verteilung ist die **Kurtosis $\boldsymbol{\gamma_4}$**, welche aus dem vierten zentralen Moment abgeleitet wird. Die Kurtosis berechnet sich zu

$$\gamma_4 = \frac{E([X - E(X)]^4)}{\mathrm{Var}(X)^2} = \frac{E[(X-\mu)^4]}{\sigma^4}.$$

Kurtosis-Exzess

Die Kurtosis einer Normalverteilung (Normalverteilung s. Abschnitt 3.3) nimmt den Wert 3 an, so dass anstelle der Kurtosis oft der **Kurtosis-Exzess $\boldsymbol{\tilde{\gamma}_4}$** betrachtet wird.

$$\tilde{\gamma}_4 = \gamma_4 - 3 = \frac{E[(X-\mu)^4]}{\sigma^4} - 3$$

Der Kurtosis-Exzess gibt an, inwieweit sich die Wölbung der vorliegenden Verteilung von der einer Normalverteilung unterscheidet. Für die Normalverteilung gilt $\tilde{\gamma}_4 = 0$, so dass bei Verteilungen mit einem Kurtosis-Exzess gleich Null von normalgipfligen (mesokurtischen) Verteilungen gesprochen wird. Ist $\tilde{\gamma}_4 > 0$, so liegt eine hochgipflige (leptokurtische) Verteilung vor, während für flachgipflige (platykurtische) Verteilungen $\tilde{\gamma}_4 < 0$ gilt. Kurtosis und Kurtosis-Exzess sind invariant gegenüber linerarer Tranformationen der Form $aX + b$, $a \neq 0$.

Beispiel 2.3.5:

Gegeben sei die Dichtefunktion

$$f_X(x) = \begin{cases} \frac{1}{2} & \text{für } 2 \leq x \leq 4 \\ 0 & \text{sonst} \end{cases}$$

mit $E(X) = 3$ und $\mathrm{Var(X)} = \frac{1}{3}$. Für die Schiefe und den Kurtosis-Exzess gilt:

$$\begin{aligned}
\gamma_3 &= \frac{E[(X-3)^3]}{\sqrt{\frac{1}{27}}} = \sqrt{27}\int_2^4 \frac{1}{2}(x-3)^3\,dx \\
&= \frac{\sqrt{27}}{8}\Big[(x-3)^4\Big]_2^4 = \frac{\sqrt{27}}{8}(1-1) = 0, \\
\tilde{\gamma}_4 &= \frac{E[(X-3)^4]}{\frac{1}{9}} - 3 = 9\int_2^4 \frac{1}{2}(x-3)^4\,dx - 3 \\
&= \frac{9}{10}\Big[(x-3)^5\Big]_2^4 - 3 = \frac{9}{10}(1+1) - 3 = -\frac{6}{5}.
\end{aligned}$$

2.4 Funktionen von Zufallsvariablen

2.4.1 Wahrscheinlichkeitsverteilung einer Funktion von Zufallsvariablen

In vielen Anwendungsfällen der Wahrscheinlichkeitsrechnung treten Zufallsvariablen als Argumente von Funktionen auf. Das ist z.B. immer dann der Fall, wenn eine Größe Y eine Funktion einer anderen Größe X ist und X eine Zufallsvariable ist.

Beispiel 2.4.1:

Der in einer Periode erzielte Gewinn eines produzierten Gutes hängt davon ab, wie viele Mengeneinheiten verkauft werden. Der Absatz eines Gutes ist aber in sehr vielen Fällen eine Zufallsvariable. Der Gewinn ist also eine Funktion der Menge, die selbst wiederum eine Zufallsvariable ist.

Wenn bei einer Funktion $Y = g(X)$ die unabhängige Variable X eine Zufallsvariable ist, dann ist auch die abhängige Variable Y wieder eine Zufallsvariable. Die Wahrscheinlichkeitsverteilung der Zufallsvariablen X bestimmt dann über die Funktion $Y = g(X)$ die Wahrscheinlichkeitsverteilung der Zufallsvariablen Y.

Beispiel 2.4.2:

Gegeben sei eine diskrete Zufallsvariable X mit der folgenden Wahrscheinlichkeitsverteilung:

x_i	-1	0	1	2	3
$f_X(x_i)$	0.2	0.2	0.2	0.2	0.2

Die Größe Y wird in Abhängigkeit von der Zufallsvariablen X wie folgt angegeben:

$$Y = X^2 - 2X + 3.$$

Für die Wahrscheinlichkeitsverteilung der Zufallsvariablen Y ergibt sich:

y_i	$y_1 = 1 + 2 + 3 = 6$	$y_2 = 0 + 0 + 3 = 3$	$y_3 = 1 - 2 + 3 = 2$
$f_Y(y_i)$	0.2	0.2	0.2
	$y_4 = 4 - 4 + 3 = 3$	$y_5 = 9 - 6 + 3 = 6$	
	0.2	0.2	

Werden gleiche y_i-Werte zusammengefasst, lautet die Wahrscheinlichkeitsverteilung für Y:

y_i	2	3	6
$f_Y(y_i)$	0.2	0.4	0.4

Ist die Zufallsvariable X stetig, dann kann die Wahrscheinlichkeitsverteilung von Y nicht in so einfacher Weise wie in dem eben angeführten Beispiel ermittelt werden.

lineare Transformation

Linearkombination

Liegt eine lineare Funktion **einer** Zufallsvariablen vor, also $Y = aX + b$, so wird diese auch als **lineare Transformation** der Zufallsvariablen X bezeichnet. Fehlt bei einer linearen Funktion mehrerer Zufallsvariablen das absolute Glied, dann wird von einer **Linearkombination** der Zufallsvariablen gesprochen. Sind z.B. X_1, X_2 und X_3 Zufallsvariablen, dann ist

$$Y = a_1X_1 + a_2X_2 + a_3X_3$$

eine Linearkombination dieser Zufallsvariablen.

2.4.2 Erwartungswert und Varianz einer Funktion von Zufallsvariablen

Betrachtet wird zunächst die **lineare Transformation einer Zufallsvariablen $\boldsymbol{X}$** mit $Y = aX + b$ bei gegebenem Erwartungswert $E(X)$.

Erwartungswert einer linearen Transformation

Für den **Erwartungswert der linearen Transformation** einer Zufallsvariablen X in der Form $\boldsymbol{Y = aX + b}$ gilt:

$$E(Y) = E(aX + b) = aE(X) + b.$$

Dieser Erwartungswert ist gleich dem linear transformierten Erwartungswert der Zufallsvariablen.

Für die **Varianz der linearen Transformation** einer Zufallsvariablen X in der Form $\boldsymbol{Y = aX + b}$ ergibt sich:

$$\begin{aligned}
\operatorname{Var}(Y) &= \operatorname{Var}(aX+b) = E[(aX+b-E[aX+b])^2] \\
&= E[(aX+b-aE(X)-b)^2] = E[a^2(X-E[X])^2] \\
&= a^2E[(X-E[X])^2] = a^2E[(X-\mu)^2] \\
&= a^2\operatorname{Var}(X).
\end{aligned}$$

Varianz einer linearen Transformation

Hier ist zu beachten, dass das absolute Glied der linearen Transformationsgleichung wegfällt, und dass außerdem der Faktor von X bei der Varianz nicht a, sondern a^2 lautet.

Beispiel 2.4.3:

Die Zufallsvariable X habe den Erwartungswert $E(X) = 5$ und die Varianz $\operatorname{Var}(X) = 4$. Die Zufallsvariable $Y = 3X + 7$ hat dann den Erwartungswert $E(Y) = 3 \cdot 5 + 7 = 22$ und die Varianz $\operatorname{Var}(X) = 3^2 \cdot 4 = 36$.

Betrachtet wird nun eine **lineare Funktion mehrerer Zufallsvariablen**:

lineare Funktion

$$Y = a_1X_1 + a_2X_2 + a_3X_3 + ... + a_nX_n + b.$$

Alle X_i $(i = 1, ..., n)$ sind Zufallsvariablen, deren Erwartungswert bekannt ist.

Für den **Erwartungswert einer linearen Funktion** von Zufallsvariablen $\boldsymbol{Y = a_1X_1 + a_2X_2 + a_3X_3 + ... + a_nX_n + b}$ gilt:

$$E(Y) = a_1E(X_1) + a_2E(X_2) + ... + a_nE(X_n) + b.$$

Erwartungswert einer linearen Funktion

Um eine einfache Beziehung für die Varianz der Zufallsvariablen Y anzugeben, wird zunächst vorausgesetzt, dass die Zufallsvariablen X_i paarweise voneinander **stochastisch unabhängig** sind, d.h. für die zweidimensionale Dichtefunktion gilt $f_{X_iX_j} = f_{X_i}f_{X_j}$, für alle $i \neq j$

stochastisch unabhängig

$(i, j = 1, ..., n)$. Der Begriff der stochastischen Unabhängigkeit wird näher in Abschnitt 2.5.3 erläutert.

Varianz einer linearen Funktion

Für paarweise stochastisch unabhängige Zufallsvariablen X_i $(i = 1, ..., n)$ und deren lineare Transformationen $Y_i = a_i X_i + b_i$ gilt für die **Varianz einer linearen Funktion** der Zufallsvariablen Y_i in der Form $\boldsymbol{Y = \sum_{i=1}^{n} Y_i}$:

$$\text{Var}(Y) = \sum_{i=1}^{n} a_i^2 \text{Var}(X_i).$$

Hinweis:

In Abschnitt 2.5.4 wird $\text{Var}(\sum_{i=1}^{n} Y_i)$ unter der Voraussetzung angegeben, dass stochastische Abhängigkeit der X_i vorliegt.

2.4.3 Standardisierung von Zufallsvariablen

Ein Spezialfall der linearen Transformation einer Zufallsvariablen ist die Standardisierung (z-Transformation) von Zufallsvariablen, die analog wie die Standardisierung von Daten (s. Kurseinheit 1, Abschnitt 2.3.4) durchgeführt wird.

Standardisierung einer Zufallsvariablen

Ist eine Zufallsvariable X mit Erwartungswert $E(X) = \mu$ und Varianz $\text{Var}(X) = \sigma^2$ gegeben, so lautet die **standardisierte Zufallsvariable $\boldsymbol{Z}$** wie folgt:

$$Z = \frac{X - E(X)}{\sqrt{\text{Var}(X)}} = \frac{X - \mu}{\sigma}.$$

Der Erwartungswert einer standardisierten Zufafallsvariablen ist 0 und ihre Varianz ist 1. Diese Eigenschaft wird genutzt, um Normalverteilungen mit beliebigem Erwartungswert und beliebiger Varianz in eine Standardnormalverteilung, welche tabelliert ist, zu transformieren.

2.5 Zweidimensionale diskrete Wahrscheinlichkeitsverteilungen

In den vorhergehenden Abschnitten wurden jeweils einzelne Zufallsvariablen, sogenannte eindimensionale Zufallsvariablen, betrachtet. Ist die gemeinsame Verteilung mehrerer Zufallsvariablen von Interessen wird von **mehrdimensionalen Zufallsvariablen** oder einem **Zufallsvektor** gesprochen. Zur Veranschaulichung der Aussagen über zweidimensionale Wahrscheinlichkeitsverteilungen können die Ausführungen über zweidimensionale Häufigkeitsverteilungen im Bereich der deskriptiven Statistik herangezogem werden. Viele der dort gemachten Ausführungen lassen sich sinngemäß auf zweidimensionale diskrete Wahrscheinlichkeitsverteilungen übertragen. Die folgenden Aussagen lassen sich auf stetige Zufallsvariablen übertragen, indem die Summen durch Integrale ersetzt werden.

mehrdimensionale Zufallsvariablen
Zufallsvektor

2.5.1 Wahrscheinlichkeits- und Dichtefunktion zweidimensionaler Zufallsvariablen

Eine **diskrete zweidimensionale** Zufallsvariable wird beschrieben durch ihre Wahrscheinlichkeitsfunktion $f_{XY}(x_i, y_j)$.

$$f_{XY}(x_i, y_j) = \boldsymbol{P}(X = x_i, Y = y_j) \qquad i = 1, ..., m \qquad j = 1, ..., r$$

Die Wahrscheinlichkeitsfunktion wird meistens in Tabellenform dargestellt.

Beispiel 2.5.1:

Von 50 Studienanfängern wurde eine zweidimensionale Häufigkeitsverteilung der Merkmale Mathematik- und Englischnote ermittelt. Die relativen Häufigkeiten entsprechen genau den Wahrscheinlichkeiten für das gemeinsame Auftreten der Merkmalsausprägungen (x_i, y_i) beim zufälligen „Ziehen" eines Studienanfängers. Für die Wahrscheinlichkeitsverteilung von Mathematik- und Englischnote beim Experiment „Zufällige Entnahme eines Studienbewerbers" ergibt sich folgende Tabelle:

x_i		y_j Englischnote					$\sum$
		1	2	3	4	5	
Mathe-matik-note	1	0.02	0.04	0.02	-	-	0.08
	2	0.04	0.08	0.08	0.02	-	0.22
	3	-	0.08	0.2	0.16	0.04	0.48
	4	0.04	0.04	-	0.06	0.02	0.16
	5	-	-	0.04	0.02	-	0.06
	$\sum$	0.1	0.24	0.34	0.23	0.06	1

Analog zu den Ausführungen über Häufigkeitsverteilungen zweier Merkmale wird für die Wahrscheinlichkeitsverteilung verlangt:

$$f_{XY}(x_i, y_j) \geq 0 \qquad \text{und} \qquad \sum_{i=1}^{m} \sum_{j=1}^{r} f_{XY}(x_i, y_j) = 1.$$

2.5.2 Randverteilungen und bedingte Verteilungen

Liegt eine zweidimensionale Wahrscheinlichkeitsverteilung vor, wobei das Interesse lediglich auf einer Zufallsvariablen beruht, dann wird wie auch in der deskriptiven Statistik, die marginale Verteilung (Randverteilung) betrachtet. Die jeweils andere Zufallsvariable bleibt dann unberücksichtigt.

Beispiel 2.5.2:

In einer Urne befinden sich 2 *rote und* 4 *grüne Kugeln, die die Zahlen* 1 *und* 2 *bzw.* 1, 2, 3 *und* 4 *tragen.*

Die Wahrscheinlichkeitsverteilung der zweidimensionalen Zufallsvariablen $(X, Y) =$ *(Farbe, Zahl) sieht dann für die einmalige Entnahme einer Kugel folgendermaßen aus:*

x_i	y_j: $y_1 = 1$	$y_2 = 2$	$y_3 = 3$	$y_4 = 4$
$x_1 =$ rot	$\frac{1}{6}$	$\frac{1}{6}$	-	-
$x_2 =$ grün	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Sind nur die roten Kugeln von Interesse, so gibt die Summe aller Wahrscheinlichkeiten in der Zeile „rot“ an, wie groß die Wahrscheinlichkeit $f_X(x_1) = f_{XY}(X = x_1, Y$ *beliebig) ist.*

$$\begin{aligned} f_X(x_1) &= f_{XY}(x_1, y_1) + f_{XY}(x_1, y_2) + f_{XY}(x_1, y_3) + f_{XY}(x_1, y_4) \\ &= \frac{1}{6} + \frac{1}{6} + 0 + 0 = \frac{2}{6} = \frac{1}{3} \end{aligned}$$

bzw.

$$f_X(x_1) = \sum_{j=1}^{r} f_{XY}(x_1, y_j).$$

Für die Wahrscheinlichkeit, eine grüne Kugel mit beliebiger Zahl zu ziehen, gilt:

$$f_X(x_2) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{4}{6} = \frac{2}{3}.$$

Entsprechend werden die Wahrscheinlichkeiten $f_Y(y_1)$, $f_Y(y_2)$,... *berechnet und trägt diese direkt in der Wahrscheinlichkeitstabelle am Ende der Zeilen und Spalten ein. Es ergeben sich so die* ***Rand-*** *bzw.* ***Marginalverteilungen*** $\boldsymbol{f_X(x_i)}$ *bzw.* $\boldsymbol{f_Y(y_j)}$.

x_i	y_j: y_1	y_2	y_3	y_4	$\boldsymbol{f_X(x_i)}$
x_1	$\frac{1}{6}$	$\frac{1}{6}$	-	-	$\boldsymbol{\frac{1}{3}}$
x_2	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\boldsymbol{\frac{2}{3}}$
$\boldsymbol{f_Y(y_j)}$	$\boldsymbol{\frac{1}{3}}$	$\boldsymbol{\frac{1}{3}}$	$\boldsymbol{\frac{1}{6}}$	$\boldsymbol{\frac{1}{6}}$	**1**

Für eine diskrete zweidimensionale Zufallsvariable (X, Y) heißt die sich ergebende Verteilung von X bzw. Y, welche die jeweils andere Zufallsvariable unberücksichtigt lässt, **Randverteilung (marginale Verteilung)**.

Es ist:

$$\begin{aligned} \boldsymbol{f_X(x_i)} &= \sum_{j=1}^{r} f_{XY}(x_i, y_j) \quad \text{Randverteilung von } X, \\ \boldsymbol{f_Y(y_j)} &= \sum_{i=1}^{m} f_{XY}(x_i, y_j) \quad \text{Randverteilung von } Y. \end{aligned}$$

Randverteilung (marginale Verteilung)

Wahrscheinlichkeitstabelle

Ähnlich wie bei Häufigkeitsverteilungen diskreter, zweidimensionaler Merkmale hat die **Wahrscheinlichkeitstabelle** folgendes Aussehen:

x_i	y_j			$f_X(x_i)$
	y_1	$\ldots$	y_r	
x_1	$f_{XY}(x_1, y_1)$	$\ldots$	$f_{XY}(x_1, y_r)$	$\sum_{j=1}^{r} f_{XY}(x_1, y_j)$
$\vdots$	$\vdots$		$\vdots$	$\vdots$
x_m	$f_{XY}(x_m, y_1)$	$\ldots$	$f_{XY}(x_m, y_r)$	$\sum_{j=1}^{r} f_{XY}(x_m, y_j)$
$f_Y(y_j)$	$\sum_{i=1}^{m} f_{XY}(x_i, y_1)$	$\ldots$	$\sum_{i=1}^{m} f_{XY}(x_i, y_r)$	1

bedingte Verteilung

Wird die Verteilung einer Zufallsvariablen betrachtet, wenn für die andere Zufallsvariable ein bestimmter Wert vorgegeben wird, so ergibt sich die **bedingte Verteilung** von X für gegebenes Y bzw. von Y für gegebenes X.

Beispiel 2.5.3:

In Beispiel 2.5.1 wurde die Wahrscheinlichkeitsverteilung von Mathematik- (X) und Englischnoten (Y) für das einmalige „Ziehen“ eines Studienbewerbers angegeben.

x_i		y_j Englischnote					$\sum$
		1	2	3	4	5	
Mathematiknote	1	0.02	0.04	0.02	-	-	0.08
	2	0.04	0.08	0.08	0.02	-	0.22
	3	-	0.08	0.2	0.16	0.04	0.48
	4	0.04	0.04	-	0.06	0.02	0.16
	5	-	-	0.04	0.02	-	0.06
	$\sum$	0.1	0.24	0.34	0.23	0.06	1

*Gezogen wird ein Studienbewerber, der nach seiner Mathematiknote befragt wird. Das Ergebnis ist beispielsweise „Mathematiknote $X = 3$“. Gesucht ist nun die Wahrscheinlichkeit, dass dieser Studienbewerber die Englischnote $Y = 4$ hat, d.h. gesucht ist die **bedingte Wahrscheinlichkeit***

$$\boldsymbol{P}(Y = 4|X = 3).$$

Es gilt

$$\boldsymbol{P}(Y = 4|X = 3) = \frac{\boldsymbol{P}(Y = 4, X = 3)}{\boldsymbol{P}(X = 3)}.$$

Dabei ist

$$\begin{aligned} \boldsymbol{P}(Y=4, X=3) &= f_{XY}(X=3, Y=4), \\ \boldsymbol{P}(X=3) &= f_X(X=3). \end{aligned}$$

Mit

$$f_{XY}(3,4) = 0.16 \qquad \textit{und} \qquad f_X(3) = 0.48$$

ergibt sich für die gesuchte Wahrscheinlichkeit:

$$\boldsymbol{P}(Y=4|X=3) = f_Y(Y=4|X=3) = \frac{0.16}{0.48} = \frac{1}{3}.$$

Der Ausdruck $f_Y(Y=y_j|X=x_i)$ bedeutet, dass für jeden Wert von Y die Wahrscheinlichkeit unter der Bedingung $X=x_i$ bestimmbar ist. Es ergibt sich die bedingte Verteilung von Y für $X=x_i$.

*Die bedingte Verteilung $f_Y(y_j|x_i)$ ergibt sich, indem **zeilenweise** die angegebenen Wahrscheinlichkeiten durch den Wert der zugehörigen Randverteilung dividiert werden.*

$\boldsymbol{f_Y(y_j\|x_i)}$						
x_i	y_j					$\sum$
	1	2	3	4	5	
1	0.25	0.5	0.25	-	-	1
2	0.182	0.364	0.364	0.09	-	1
3	-	0.167	0.417	0.333	0.083	1
4	0.25	0.25	-	0.375	0.125	1
5	-	-	0.667	0.333	-	1

*Die bedingte Verteilung $f_X(x_i|y_j)$ ergibt sich analog, indem die gegebenen Wahrscheinlichkeiten **spaltenweise** durch den Wert der zugehörigen Randverteilung dividiert werden.*

$\boldsymbol{f_X(x_i\|y_j)}$					
x_i	y_j				
	1	2	3	4	5
1	0.2	0.167	0.059	-	-
2	0.4	0.333	0.235	0.077	-
3	-	0.333	0.588	0.615	0.667
4	0.4	0.167	-	0.231	0.333
5	-	-	0.118	0.077	-
$\sum$	1	1	1	1	1

Die Ergebnisse aus Beispiel 2.5.3 lassen sich folgendermaßen zusammenfassen:

bedingte Verteilung

Für die **bedingte Verteilung** einer zweidimensionalen **diskreten** Zufallsvariablen gilt:

$$f_X(x_i|y_j) = \frac{f_{XY}(x_i, y_j)}{f_Y(y_j)}, \quad \sum_{i=1}^{m} f_X(x_i|y_j) = 1, \quad f_Y(y_j) > 0$$

bzw.

$$f_Y(y_j|x_i) = \frac{f_{XY}(x_i, y_j)}{f_X(x_i)}, \quad \sum_{j=1}^{r} f_Y(y_j|x_i) = 1, \quad f_X(x_i) > 0.$$

2.5.3 Parameter zweidimensionaler Verteilungen

Für zweidimensionale Verteilungen können zunächst folgende Parameter bestimmt werden:

- Erwartungswerte und Varianzen bzw. Standardabweichungen der **Randverteilungen**.

- Erwartungswerte und Varianzen der **bedingten Verteilungen**.

Zur Bestimmung dieser Parameter kann, da es sich um Parameter eindimensionaler Wahrscheinlichkeitsverteilungen handelt, auf Abschnitt 2.3 dieser Kurseinheit verwiesen werden.

Beispiel 2.5.4:

Es wird auf das Beispiel 2.5.3 zurückgegriffen.

x_i		y_j					
		Englischnote					
		1	2	3	4	5	$\sum$
Mathematiknote	1	0.02	0.04	0.02	-	-	0.08
	2	0.04	0.08	0.08	0.02	-	0.22
	3	-	0.08	0.2	0.16	0.04	0.48
	4	0.04	0.04	-	0.06	0.02	0.16
	5	-	-	0.04	0.02	-	0.06
	$\sum$	0.1	0.24	0.34	0.26	0.06	1

Für den Erwartungswert bei Betrachtung der Mathematiknote eines Bewerbers (beim „einmaligen Ziehen“) ergibt sich:

$$E(X) = 0.08 \cdot 1 + 0.22 \cdot 2 + 0.48 \cdot 3 + 0.16 \cdot 4 + 0.06 \cdot 5 = 2.9.$$

Wird die Englischnote allein betrachtet, gilt:

$$E(Y) = 0.1 \cdot 1 + 0.24 \cdot 2 + 0.34 \cdot 3 + 0.26 \cdot 4 + 0.06 \cdot 5 = 2.94.$$

$E(X)$ und $E(Y)$ sind die Erwartungswerte der Randverteilungen.

$$\begin{aligned}
\mathrm{Var}(X) &= \sum f_X(x_i) \cdot x_i^2 - [E(X)]^2 \\
&= 0.08 \cdot 1^2 + 0.22 \cdot 2^2 + 0.48 \cdot 3^2 + 0.16 \cdot 4^2 + 0.06 \cdot 5^2 - 2.9^2 \\
&= 9.34 - 8.41 = 0.93 \\
\mathrm{Var}(Y) &= 0.1 \cdot 1^2 + 0.24 \cdot 2^2 + 0.34 \cdot 3^2 + 0.26 \cdot 4^2 + 0.06 \cdot 5^2 - 2.94^2 \\
&= 9.78 - 8.64 = 1.14
\end{aligned}$$

Die Erwartungswerte und Varianzen der bedingten Verteilungen werden am einfachsten mit Hilfe der tabellierten bedingten Verteilung aus Beispiel 2.5.3 bestimmt. Ist beispielsweise bereits die Mathematiknote eines Bewerbers bekannt ($x_i = 3$), so kann der bedingte Erwartungswert der Englischnote bestimmt werden:

$$E(Y|x_i = 3) = 0.167 \cdot 2 + 0.417 \cdot 3 + 0.333 \cdot 4 + 0.083 \cdot 5 = 3.33.$$

Umgekehrt ist der bedingte Erwartungswert für die Mathematiknote eines Bewerbers, dessen Englischnote mit $y_j = 1$ bekannt ist,

$$E(X|y_j = 1) = 0.2 \cdot 1 + 0.4 \cdot 2 + 0.4 \cdot 4 = 2.6.$$

Die bedingten Varianzen berechnen sich entsprechend.

Ein spezieller Parameter für zweidimensionale Wahrscheinlichkeitsverteilungen ist die sogenannte **Kovarianz** mit

Kovarianz

$$\mathrm{Cov}(X, Y) = E([X - E(X)][Y - E(Y)]) = \sigma_{XY}.$$

Kovarianz diskreter Zufallsvariablen

$$\begin{aligned}
\mathrm{Cov}(X, Y) &= \sum_{i=1}^{m} \sum_{j=1}^{r} [x_i - E(X)][y_j - E(Y)] f_{XY}(x_i, y_j) \\
&= \sum_{i=1}^{m} \sum_{j=1}^{r} x_i y_j f_{XY}(x_i, y_j) - E(X)E(Y) \\
&= E(XY) - E(X)E(Y)
\end{aligned}$$

Beispiel 2.5.5:

Gegeben sei die in der folgenden Tabelle dargestellte zweidimensionale Wahrscheinlichkeitsverteilung mit den zugehörigen Randverteilungen.

X	Y		
	$y_1 = 1$	$y_2 = 2$	$f_X(x_i)$
$x_1 = 2$	$f_{XY}(x_1, y_1) = 0.4$	$f_{XY}(x_1, y_2) = 0.3$	$f_X(x_1) = 0.7$
$x_2 = 4$	$f_{XY}(x_2, y_1) = 0.2$	$f_{XY}(x_2, y_2) = 0.1$	$f_X(x_2) = 0.3$
$f_Y(y_i)$	$f_Y(y_1) = 0.6$	$f_Y(y_2) = 0.4$	1

Es ist

$$\begin{aligned} E(X) &= 0.7 \cdot 2 + 0.3 \cdot 4 = 2.6, \\ E(Y) &= 0.6 \cdot 1 + 0.4 \cdot 2 = 1.4. \end{aligned}$$

Die Kovarianz lautet

$$\begin{aligned} \operatorname{Cov}(X, Y) &= (2 - 2.6) \cdot (1 - 1.4) \cdot 0.4 + (2 - 2.6) \cdot (2 - 1.4) \cdot 0.3 \\ &\quad + (4 - 2.6) \cdot (1 - 1.4) \cdot 0.2 + (4 - 2.6) \cdot (2 - 1.4) \cdot 0.1 \\ &= 0.6 \cdot 0.4 \cdot 0.4 - 0.6 \cdot 0.6 \cdot 0.3 - 1.4 \cdot 0.4 \cdot 0.2 + 1.4 \cdot 0.6 \cdot 0.1 \\ &= 0.096 - 0.108 - 0.112 + 0.084 = -0.04. \end{aligned}$$

stochastisch unabhängig

Zwei diskrete Zufallsvariablen X und Y sind genau dann **stochastisch unabhängig**, wenn für die gemeinsame Verteilung der Zufallsvariablen gilt:

$$f_{XY}(x_i, y_j) = f_X(x_i) f_Y(y_j) \quad \text{für alle } i \text{ und } j,$$

Im anderen Fall werden die Zufallsvariablen als **stochastisch abhängig** bezeichnet.

stochastisch abhängig

Für die Untersuchung der stochastischen Abhängigkeit bzw. Unabhängigkeit von Zufallsvariablen werden zunächst die Randverteilungen bestimmt und anschließend miteinander multipliziert (s. Analogie zu KE 1 Abschnitt 3.2, unabhängige und abhängige Merkmale). Ergibt sich dabei die ursprüngliche Verteilung, liegt Unabhängigkeit vor.

Beispiel 2.5.6:

Für die Verteilung aus Beispiel 2.5.5 ergibt sich

$$f_{XY}(x_1, y_1) = 0.4 \quad \textit{und} \quad f_X(x_1) f_Y(y_1) = 0.7 \cdot 0.6 = 0.42.$$

Die Zufallsvariablen X und Y sind stochastisch abhängig.

Sinngemäß entsprechen diese Ausführungen denen über die Abhängigkeit und Unabhängigkeit von Merkmalen in der deskriptiven Statistik. Der Unabhängigkeit von Merkmalen entspricht die stochastische Unabhängigkeit von Zufallsvariablen.

> Für unabhängige Zufallsvariablen gilt: $\text{Cov}(X, Y) = 0$.
> Die **Umkehrung** des Satzes **gilt nicht**.

Nimmt die Kovarianz $\text{Cov}(X, Y)$ den Wert 0 an, so sind die Zufallsvariablen X und Y nicht unbedingt unabhängig.

Regeln zur Berechnung der Kovarianz

> Allgemein gilt für zwei Zufallsvariablen X und Y:
>
> $$\begin{aligned} \text{Cov}(X, X) &= \text{Var}(X), \\ \text{Cov}(X, Y) &= E(XY) - E(X)E(Y), \\ \text{Cov}(X, Y) &= \text{Cov}(Y, X), \\ \text{Cov}(\tilde{X}, \tilde{Y}) &= a_1 a_2 \text{Cov}(X, Y) \\ \text{mit} \quad & \tilde{X} = a_1 X + b_1, \quad \tilde{Y} = a_2 Y + b_2. \end{aligned}$$

2.5.4 Varianz einer linearen Funktion

Nach Einführung der Kovarianz und den Regeln zur Berechnung der Kovarianz kann nun allgemein die Varianz $\text{Var}(Y)$ der linearen Funktion Y mit Y als Summe der linearen Transformationen $aX_i + b_i$ der Zufallsvariablen X_i

$$Y = \sum_{i=1}^{n} a_i X_i + \sum_{i=1}^{n} b_i = \sum_{i=1}^{n} a_i X_i + b \quad \text{mit} \quad b = \sum_{i=1}^{n} b_i$$

berechnet werden, wobei die Annahme der paarweisen Unabhängigkeit der X_i $(i = 1, ..., n)$ nicht mehr vorausgesetzt wird (s. Abschnitt 2.4.2).

$$\begin{aligned} \text{Var}(Y) &= E([Y - E(Y)]^2) \\ &= E\left(\left[\sum_{i=1}^{n} a_i X_i + b - \sum_{i=1}^{n} a_i E(X_i) - b\right]^2\right) \\ &= E\left(\left[\sum_{i=1}^{n} a_i [X_i - E(X_i)]\right]^2\right) \end{aligned}$$

$$
\begin{aligned}
&= E\left(\sum_{i=1}^{n} a_i^2 [X_i - E(X_i)]^2 + \sum_{i \neq j} a_j a_j [X_i - E(X_i)][X_j - E(X_j)]\right) \\
&= \sum_{i=1}^{n} a_i^2 \mathrm{Var}(X_i) + \sum_{i \neq j} a_i a_j \mathrm{Cov}(X_i, X_j) \\
&= \sum_{i=1}^{n} a_i^2 \mathrm{Var}(X_i) + 2 \sum_{i<j} a_i a_j \mathrm{Cov}(X_i, X_j).
\end{aligned}
$$

Die Summe $\sum_{i \neq j}$ bezeichnet dabei die Summe über alle i und j mit $i \neq j$. Da gilt $\mathrm{Cov}(X_i, Y_j) = \mathrm{Cov}(Y_j, X_i)$, d.h. jede Kovarianz wird doppelt gezählt, kann $\sum_{i \neq j}$ durch $2\sum_{i<j}$ ersetzt werden.

Varianz einer linearen Funktion

Für die lineare Funktion $\boldsymbol{Y = \sum_{i=1}^{n} Y_i}$ linearer Transformationen $Y_i = a_i X_i + b_i$ von **stochastisch abhängigen** Zufallsvariablen X_i $(i = 1, ..., n)$ gilt:

$$
\mathrm{Var}\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} a_i^2 \mathrm{Var}(X_i) + 2 \sum_{i<j} a_i a_j \mathrm{Cov}(X_i, X_j).
$$

Beispiel 2.5.7:

Gegeben sind die zwei Zufallsvariablen X und Y aus Beispiel 2.5.5 mit $E(X) = 1.4$, $E(Y) = 2.6$ und $\mathrm{Cov}(X, Y) = -0.04$. Gesucht wird die Varianz der Zufallsvariablen $Z = 4X + 3Y - 2$. Dazu werden zunächst $\mathrm{Var}(X)$ und $\mathrm{Var}(Y)$ mittels der Formel $\mathrm{Var}(X) = E(X^2) - [E(X)]^2$ berechnet.

$$
\begin{aligned}
E(X^2) &= 0.7 \cdot 2^2 + 0.3 \cdot 4^2 = 7.6 \\
E(Y^2) &= 0.6 \cdot 1^2 + 0.4 \cdot 2^2 = 2.2 \\
\mathrm{Var}(X) &= 7.6 - 6.76 = 0.84 \\
\mathrm{Var}(Y) &= 2.2 - 1.96 = 0.24
\end{aligned}
$$

Mit diesen Angaben berechnet sich die Varianz von Z zu:

$$
\begin{aligned}
\mathrm{Var}(Z) &= 4^2 \mathrm{Var}(X) + 3^2 \mathrm{Var}(Y) + 4 \cdot 3 \mathrm{Cov}(X, Y) + 3 \cdot 4 \mathrm{Cov}(Y, X) \\
&= 4^2 \mathrm{Var}(X) + 3^2 \mathrm{Var}(Y) + 2 \cdot 4 \cdot 3 \mathrm{Cov}(X, Y) \\
&= 16 \mathrm{Var}(X) + 9 \mathrm{Var}(Y) + 24 \mathrm{Cov}(X, Y) \\
&= 16 \cdot 0.84 + 9 \cdot 0.24 + 24 \cdot (-0.04) = 14.64.
\end{aligned}
$$

3 Spezielle Wahrscheinlichkeitsverteilungen

In den vorangegangenen Abschnitten wurden Zufallsvariablen, Wahrscheinlichkeitsverteilungen von Zufallsvariablen und Parameter von Wahrscheinlichkeitsverteilungen (z.B. Erwartungswert und Varianz) im Allgemeinen betrachtet. Für viele verschiedene Anwendungen der Wahrscheinlichkeitsrechnung spielen einige spezielle Wahrscheinlichkeitsverteilungen eine besondere Rolle.

Im Folgenden werden einige wichtige Wahrscheinlichkeitsverteilungen eingeführt. Da in bestimmten Fällen die Anwendung einzelner Wahrscheinlichkeitsverteilungen recht aufwendig ist, wird oft die Möglichkeit zur Annäherung (Approximation) einer Wahrscheinlichkeitsverteilung durch eine andere genutzt (s. Abschnitt 3.4).

3.1 Gleichverteilung

Eine der einfachsten Wahrscheinlichkeitsverteilungen ist die Gleichverteilung einer **diskreten** Zufallsvariablen.

Ist X eine diskrete Zufallsvariable, welche die Werte x_i ($i = 1, 2, ..., n$) mit positiven Wahrscheinlichkeiten $f_X(x_i)$ annimmt, dann heißt X **diskret gleichverteilt**, wenn jeder Wert mit der gleichen Wahrscheinlichkeit angenommen wird. Für n Beobachtungen wird die Wahrscheinlichkeitsfunktion der **diskreten Gleichverteilung** wie folgt angegeben:

diskrete Gleichverteilung

$$f_X(x_i) = \frac{1}{n}, \quad \text{d.h.} \sum_{i=1}^{n} f_X(x_i) = \sum_{i=1}^{n} \frac{1}{n} = 1.$$

Wahrscheinlichkeitsfunktion

Liegen die Beobachtungen in aufsteigender Reihenfolge vor, d.h. $x_1 \leq x_2 \leq ... \leq x_n$, so gilt für die Verteilungsfunktion $F_X(x) = \boldsymbol{P}(X \leq x) = \sum_{x_i \leq x} f_X(x_i)$:

$$F_X(x) = \begin{cases} 0 & \text{für} \quad x < x_1 \\ \frac{i}{n} & \text{für } x_i \leq x < x_{i+1}; \ i = 1, ..., n-1 \\ 1 & \text{für } x_n \leq x. \end{cases}$$

Verteilungsfunktion

Eine diskrete Gleichverteilung liegt z.B. beim Würfelwurf vor, wobei alle Augenzahlen mit der Wahrscheinlichkeit $\frac{1}{6}$ auftreten. Abbildung 3.1.1 zeigt die theoretische Wahrscheinlichkeitsverteilung und die Verteilungsfunktion eines Würfelwurfs.

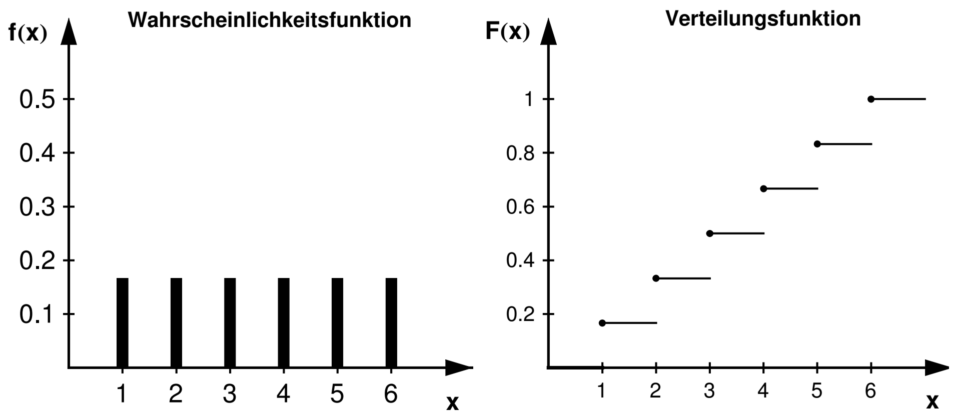


Abbildung 3.1.1: *Diskrete Gleichverteilung (Würfelwurf)*

Im Fall $x_i = i$ können die Parameter Erwartungswert $E(X)$ und Varianz $\mathrm{Var}(X)$ der diskreten Gleichverteilung direkt angegeben werden.

Erwartungswert und Varianz der diskreten Gleichverteilung im Fall $x_i = i$

$$
\begin{aligned}
E(X) &= \sum_{i=1}^{n} x_i f_X(x_i) = \frac{1}{n}\sum_{i=1}^{n} x_i \\
&= \frac{1}{n}\sum_{i=1}^{n} i = \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2} \\
\mathrm{Var}(X) &= E(X^2) - (E(X))^2 = \frac{1}{n}\sum_{i=1}^{n} i^2 - \left(\frac{n+1}{2}\right)^2 \\
&= \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 \\
&= \frac{1}{12}(n^2 - 1)
\end{aligned}
$$

Ist X eine **stetige Zufallsvariable**, so heißt X stetig gleichverteilt (rechteckverteilt), wenn die Dichtefunktion in einem bestimmten Intervall eine Konstante ist.

Lautet das Intervall $a \leq x \leq b$, so hat die Dichtefunktion den Wert $\frac{1}{b-a}$, denn die Bedingung

$$\int_{-\infty}^{\infty} f_X(x)dx = 1$$

muss erfüllt sein.

$$\int_{-\infty}^{\infty} f_X(x)dx = \int_a^b \frac{1}{b-a}dx = \left[\frac{x}{b-a}\right]_a^b = \frac{b-a}{b-a} = 1.$$

Für die **stetige Gleichverteilung (Rechteckverteilung)** über dem Intervall $[a;b]$ ist die Dichtefunktion gegeben durch:

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & \text{für } a \leq x \leq b \\ 0 & \text{sonst.} \end{cases}$$

Für die Verteilungsfunktion einer stetigen Gleichverteilung ergibt sich für $a \leq x < b$:

$$\begin{aligned} F_X(x) &= \int_a^x \frac{1}{b-a}\, d\xi = \int_a^x \frac{1}{b-a}\, d\xi = \left[\frac{\xi}{b-a}\right]_a^x \\ &= \frac{x}{b-a} - \frac{a}{b-a} = \frac{x-a}{b-a}. \end{aligned}$$

stetige Gleichverteilung

Dichtefunktion

Verteilungsfunktion

Unter der Berücksichtigung von $F_X(x) = 0$ für $x < a$ und $F_X(x) = 1$ für $x \geq b$ gilt somit:

$$F_X(x) = \begin{cases} 0 & \text{für} \quad x < a \\ \dfrac{x-a}{b-a} & \text{für } a \leq x < b \\ 1 & \text{für } b \leq x. \end{cases}$$

Abbildung 3.1.2 zeigt die Dichte- und die Verteilungsfunktion einer stetigen Gleichverteilung.

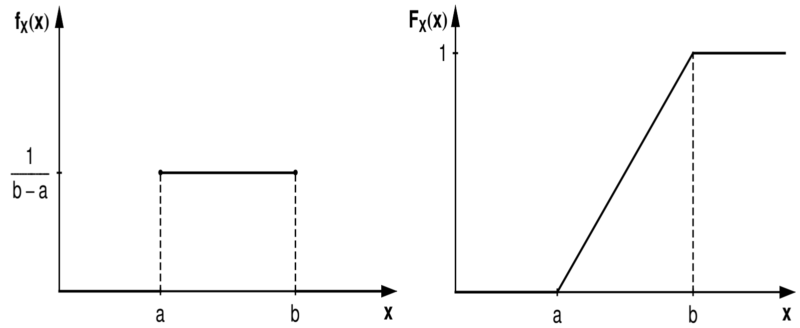


Abbildung 3.1.2: *Dichte und Verteilungsfunktion einer stetigen Gleichverteilung*

Für die stetige Gleichverteilung ergeben sich die Parameter:

Erwartungswert und Varianz der stetigen Gleichverteilung

$$E(X) = \frac{b+a}{2} \quad \text{und} \quad \mathrm{Var}(X) = \frac{(b-a)^2}{12}.$$

$$\begin{aligned}
E(X) &= \int_a^b \frac{1}{b-a} x \, dx = \frac{1}{b-a}\left[\frac{1}{2}x^2\right]_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{b+a}{2}, \\
\mathrm{Var}(X) &= \int_a^b \frac{1}{b-a} x^2 \, dx - \left(\frac{b+a}{2}\right)^2 = \frac{1}{b-a}\left[\frac{1}{3}x^3\right]_a^b - \left(\frac{b+a}{2}\right)^2 \\
&= \frac{b^3-a^3}{3(b-a)} - \frac{(b+a)^2}{4} \\
&= \frac{b^2+ab+a^2}{3} - \frac{b^2+2ab+a^2}{4} \\
&= \frac{b^2-2ab+a^2}{12} = \frac{(b-a)^2}{12}.
\end{aligned}$$

3.2 Binomialverteilung

Eine der wichtigsten diskreten Wahrscheinlichkeitsverteilung ist die **Binomialverteilung**. Zur Herleitung der Binomialverteilung werden einige Grundlagen aus der Kombinatorik benötigt, siehe Abschnitt 1.4.3.

Betrachtet wird ein bestimmtes Zufallsexperiment, das sogenannte **Bernoulli-Experiment**. Es liegen lediglich zwei mögliche Ereignisse vor, die Ereignisse A und $\overline{A}$, die mit den Wahrscheinlichkeiten $\boldsymbol{P}(A) = \pi$ und $\boldsymbol{P}(\overline{A}) = 1 - \boldsymbol{P}(A) = 1 - \pi$ auftreten. Bei wiederholter Durchführung des Zufallsexperiments seien die einzelnen Ereignisse voneinander unabhängig.

Bernoulli-Experiment

Beispiel 3.2.1:

a) Beim Werfen mit einer Münze können die Ergebnisse „Kopf" oder „Zahl" auftreten. Ist A das Ereignis „Zahl", so gilt für eine ideale Münze $\boldsymbol{P}(A) = 0.5$.

b) Aus einer Warenlieferung werden zufällig Stücke entnommen. Es können die beiden Ereignisse „das Stück ist einwandfrei" oder „das Stück ist Ausschuss" auftreten.

Es soll nun die Wahrscheinlichkeit dafür bestimmt werden, dass bei n unabhängigen Wiederholungen des Bernoulli-Experimentes x-mal das Ereignis A und damit $(n-x)$-mal das Ereignis $\overline{A}$ eintritt. Eine mögliche Anordnung einer solchen Ereignisfolge von Bernoulli-Experimenten lautet:

$$\underbrace{A, A, ..., A, A}_{x-mal}, \underbrace{\overline{A}, \overline{A}, ..., \overline{A}, \overline{A}}_{(n-x)-mal}.$$

Aufgrund der vorausgesetzten Unabhängigkeit kann der Multiplikationssatz für stochastisch unabhängige Ereignisse angewendet werden.

$$\begin{aligned} \boldsymbol{P}(x\text{-mal } A;\, (n-x)\text{-mal } \overline{A}) &= \underbrace{\boldsymbol{P}(A) \cdot ... \cdot \boldsymbol{P}(A)}_{x-mal} \cdot \underbrace{\boldsymbol{P}(\overline{A}) \cdot ... \cdot \boldsymbol{P}(\overline{A})}_{(n-x)-mal} \\ &= [\boldsymbol{P}(A)]^x \cdot [\boldsymbol{P}(\overline{A})]^{n-x} = \pi^x (1-\pi)^{n-x} \end{aligned}$$

Das Ereignis „x-mal das Ereignis A und $(n-x)$-mal das Ereignis $\overline{A}$" tritt auch ein, wenn eine andere Reihenfolge der einzelnen Ereignisse vorliegt.

Nach den Regeln der Kombinatorik gibt es genau $\binom{n}{x}$ solch möglicher Ereignisfolgen (Anzahl der Kombinationen ohne Berücksichtigung der Reihenfolge und ohne Wiederholungen; für das Ereignis A werden aus n Platzziffern x herausgezogen).

Binomialverteilung

Zugrundegelegt wird ein Bernoulli-Experiment mit den beiden möglichen Ereignissen A und $\overline{A}$ und den Wahrscheinlichkeiten $\boldsymbol{P}(A) = \pi$ und $\boldsymbol{P}(\overline{A}) = 1 - \pi$. Die **Binomialverteilung** gibt bei n-maliger unabhängiger Wiederholung des Experimentes die Wahrscheinlichkeit an, dass genau x-mal das Ereignis A eintritt.

Wahrscheinlichkeitsfunktion

$$B(x|n,\pi) = f_X(x) = \binom{n}{x}\pi^x(1-\pi)^{n-x} \qquad x = 0, 1, ..., n$$

Die Verteilungsfunktion $F_X(x) = \boldsymbol{P}(X \leq x)$ ergibt sich durch Summation:

Verteilungsfunktion

$$F_X(x) = \sum_{k=0}^{x}\binom{n}{k}\pi^k(1-\pi)^{n-k}, \qquad x = 0, 1, ..., n.$$

Die Zufallsvariable X heißt binomialverteilt mit den Parametern n und π. Die Schreibweise lautet $\boldsymbol{X \sim B(n, \pi)}$.

Beispiel 3.2.2:

Bei 4maligem Werfen einer Münze gibt es 6 verschiedene Ergebnisfolgen bzw. Anordnungen, die 2mal das Ergebnis „Zahl“ aufweisen können, nämlich:

$$\begin{array}{ccc} (K,K,Z,Z) & (K,Z,K,Z) & (Z,K,K,Z) \\ (K,Z,Z,K) & (Z,K,Z,K) & (Z,Z,K,K) \end{array}$$

Es gilt $\boldsymbol{P}(Z) = \boldsymbol{P}(K) = 0.5$.

Erfolgen die Münzwürfe unabhängig voneinander, dann tritt jede der angegebenen Ergebnisfolgen mit der Wahrscheinlichkeit

$$0.5^4 = 0.0625 = \frac{1}{16}$$

auf. Da es 6 verschiedene Ergebnisfolgen gibt, bei denen „zweimal Zahl“ oben liegt, gilt dann:

$$\boldsymbol{P}(\text{„zweimal Zahl“}) = 6 \cdot \frac{1}{16} = \frac{3}{8}.$$

Abbildung 3.2.1 zeigt Wahrscheinlichkeits- und Verteilungsfunktion einer $B(4, 0.5)$-verteilten Zufallsvariablen.

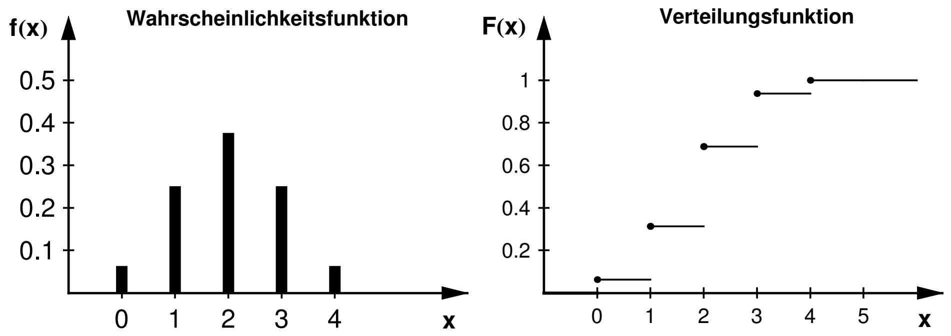


Abbildung 3.2.1: *Binomialverteilung mit $n = 4$ und $\pi = 0.5$*

Beispiel 3.2.3:

Es wird mit 4 Münzen, deren Vorderseiten eine 1 und deren Rückseiten eine 0 aufweisen geworfen. Für die Wahrscheinlichkeitsverteilung der Ergebnissumme ergibt sich dann aus der Binomialverteilung folgende Wahrscheinlichkeitsfunktion:

x_i	0	1	2	3	4
$f_X(x_i)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

Die Verteilungsfunktion lautet somit:

$$F_X(x) = \begin{cases} 0 & \text{für} \quad x < 0 \\ \frac{1}{16} & \text{für } 0 \leq x < 1 \\ \frac{5}{16} & \text{für } 1 \leq x < 2 \\ \frac{11}{16} & \text{für } 2 \leq x < 3 \\ \frac{15}{16} & \text{für } 3 \leq x < 4 \\ 1 & \text{für } 4 \leq x. \end{cases}$$

Mittels folgender Überlegung lässt sich der Erwartungswert einer binomialverteilten Zufallsvariablen auf einfache Weise bestimmen. Die einmalige Durchführung des zugrundegelegten Bernoulli-Experimentes kann durch eine Zufallsvariable X_i mit den beiden möglichen Werten 1 (Ereignis A tritt ein; „Erfolg") und 0 (Ereignis A tritt nicht ein; „Misserfolg") beschrieben werden. Es ist $\boldsymbol{P}(1) = \pi$ und $\boldsymbol{P}(0) = 1 - \pi$ sowie $E(X_i) = 1 \cdot \pi + 0 \cdot (1 - \pi) = \pi$. Wird ein Folge von n unabhängigen Bernoulli-Experimenten durchgeführt, so ergibt sich die Zufallsvariable X, welche das Ereignis „Anzahl des Ereignisses A" beschreibt, durch Addition der einzelnen unabhängigen Zufallsvariablen X_i, d.h. $X = \sum_{i=1}^{n} X_i$. Somit gilt:

Erwartungswert der Binomialverteilung

Der **Erwartungswert der Binomialverteilung** ergibt sich zu:

$$E(X) = \sum_{i=1}^{n} E(X_i) = n\pi.$$

Durch entsprechende Überlegungen ist auch die Varianz leicht zu ermitteln. Für das einzelne Experiment gilt:

$$\begin{aligned}
\mathrm{Var}(X_i) &= (0-\pi)^2 \cdot (1-\pi) + (1-\pi)^2 \cdot \pi \\
&= \pi^2 - \pi^3 + \pi - 2\pi^2 + \pi^3 = \pi - \pi^2 \\
&= \pi(1-\pi).
\end{aligned}$$

Varianz der Binomialverteilung

Wegen der Unabhängigkeit der X_i gilt dann für die **Varianz der Binomialverteilung**:

$$\mathrm{Var}(X) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = n\pi(1-\pi).$$

Beispiel 3.2.4:

Erwartungswert und Varianz für den in Beispiel 3.2.3 dargestellten Münzwurf ergeben sich zu

$$\begin{aligned}
E(X) &= n\pi = 4 \cdot 0.5 = 2, \\
\mathrm{Var}(X) &= n\pi(1-\pi) = 4 \cdot 0.5 \cdot 0.5 = 1.
\end{aligned}$$

Die Binomialverteilung besitzt eine wichtige, als **Reproduktivität** bezeichnete Eigenschaft:

Sind die unabhängigen Zufallsvariablen X und Y binomialverteilt mit $X \sim B(n, \pi)$ und $Y \sim B(m, \pi)$, so ist auch die Zufallsvariable $X + Y$ binomialverteilt, und zwar $\boldsymbol{X + Y \sim B(n + m, \pi)}$. Zu beachten ist, dass für beide ursprüngliche Binomialverteilungen die Wahrscheinlichkeit π übereinstimmen muss.

Reproduktivität der Binomialverteilung

Beispiel 3.2.5:

Mit einer idealen Münze werden 3 Wurfserien zu je 4 Würfen durchgeführt. Wie groß ist die Wahrscheinlichkeit dafür, dass bei den 3 Serien insgesamt viermal das Ergebnis „Zahl" erscheint?

Unter Beachtung der Reproduktivität von Binomialverteilungen ist dann $n = 12$ *und*

$$f_X(4) = \binom{12}{4}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^8 = \binom{12}{4} \cdot \frac{1}{2^{12}} = \frac{495}{4096} = 0.1208.$$

Tabelle der Binomialverteilung:

Die Tabelle der Binomialverteilung des Gesamtglossars enthält für $n = 1$ bis $n = 20$ die Werte der Wahrscheinlichkeitsfunktion für $\pi = 0.05,\ \pi = 0.1,\ \pi = 0.15, ...,\ \pi = 0.5$. Für Werte $\pi > 0.5$ ergeben sich die Werte der Wahrscheinlichkeitsfunktion aus der Tabelle für $\pi^* = 1 - \pi$, indem anstelle von x für die Zufallsvariable die Werte $n - x$ verwendet werden.

Beispiel 3.2.6:

In einer Urne befinden sich zu 70% rote Kugeln. 12 Kugeln werden nacheinander so herausgegriffen, dass vor jedem Zug die vorher gezogene Kugel wieder zurückgelegt wird. Wie groß ist die Wahrscheinlichkeit, 8 rote Kugeln zu ziehen?

Es ist $\pi = 0.7,\ n = 12,\ x = 8$.

$$B(x|n, \pi) = B(n - x|n, 1 - \pi)$$

Es wird für $n = 12, x' = n - x = 4$ *und* $\pi' = 1 - \pi = 0.3$ *in der Tabelle nachgeschlagen.*

$$B(8|12, 0.7) = B(12 - 8|12, (1 - 0.7)) = B(4|12, 0.3) = 0.2311$$

3.3 Normalverteilung

3.3.1 Definition der Normalverteilung

Die **Normalverteilung** ist die wichtigste stetige Verteilung. Sie spielt bei nahezu allen Anwendungen der Statistik eine große Rolle.

Normalverteilung

Dichtefunktion

Die Dichtefunktion der **Normalverteilung** lautet:

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

Zu der Dichtefunktion $f_X(x)$ existiert keine elementare Stammfunktion, so dass die **Verteilungsfunktion** der Normalverteilung nicht mehr mit Hilfe elementarer Funktionen darstellbar ist. Die Werte der Verteilungsfunktion können mittels numerischer Integrationsverfahren oder spezieller Tabellen (s. Gesamtglossar) angegeben werden. Die Parameter der Normalverteilung lauten:

Erwartungswert und Varianz

$$E(X) = \mu \quad \text{und} \quad \mathrm{Var}(X) = \sigma^2.$$

Eine normalverteilte Zufallsvariable X wird als $N(\mu, \sigma^2)$-verteilt bezeichnet. Die Schreibweise lautet $\boldsymbol{X \sim N(\mu, \sigma^2)}$.

Erwartungswert und Varianz bzw. Standardabweichung der Normalverteilung lassen sich also unmittelbar aus der Dichtefunktion ablesen.

Aus der Dichtefunktion ergibt sich, dass die Normalverteilung in einem konkreten Fall durch die Angabe von μ und σ^2 jeweils spezifiziert werden muss. Es gibt also nicht nur eine Normalverteilung, sondern eine ganze Klasse von Normalverteilungen. Die Dichtefunktion der Normalverteilung hat folgende typische Gestalt:[5]

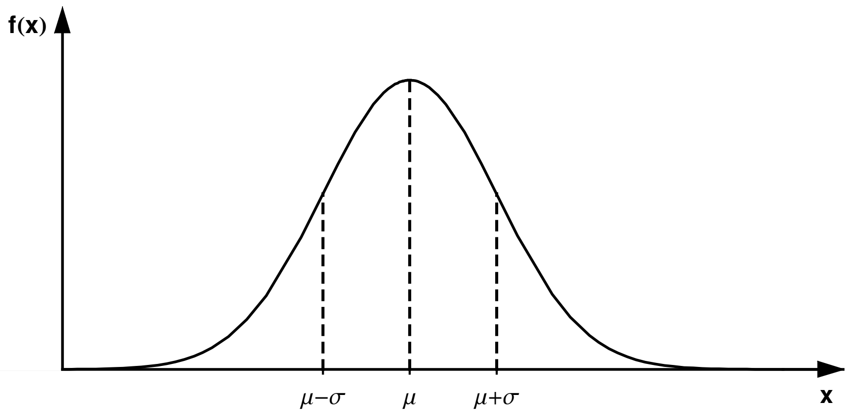


Abbildung 3.3.1: *Dichtefunktion der Normalverteilung*

[5]siehe interaktive Mathematica-Applets auf der Homepage des Lehrstuhls http://www.fernuni-hagen.de/ls_statistik/forschung/multimedia/

Die Dichtefunktion ist symmetrisch und hat ihren Gipfel bei $x = \mu$. An den Stellen $x = \mu - \sigma$ und $x = \mu + \sigma$ befinden sich Wendepunkte. Aufgrund der Symmetrie gilt, dass Erwartungswert, Median und Modalwert übereinstimmen. In der Abbildung 3.3.2 sind Normalverteilungen für verschiedene Werte von μ und σ^2 dargestellt.

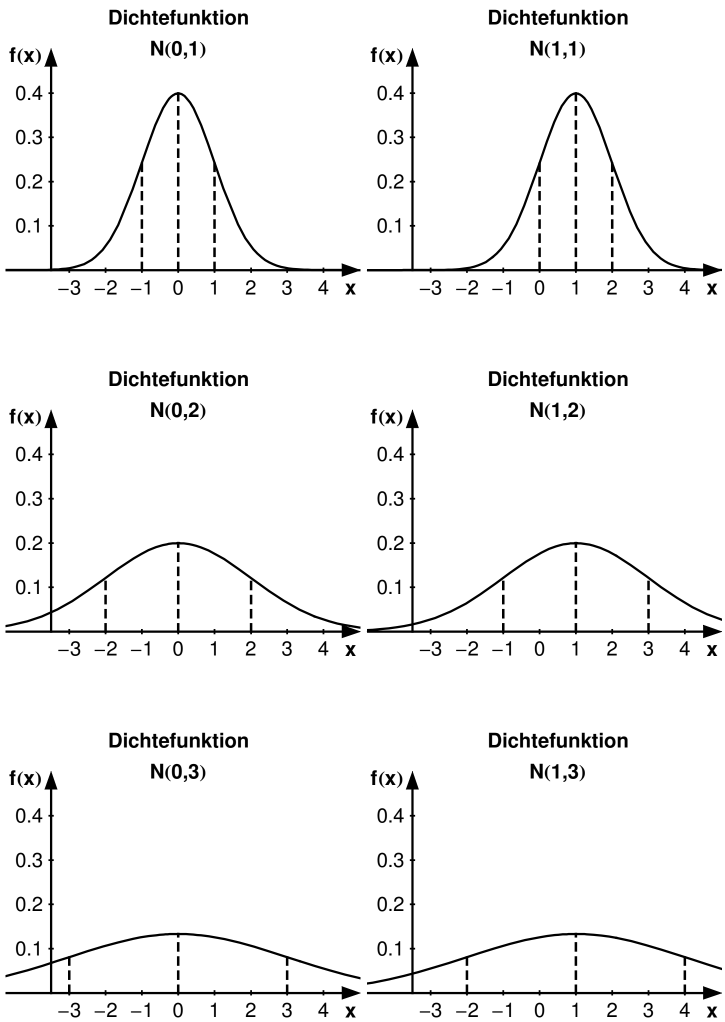


Abbildung 3.3.2: *Verschiedene Normalverteilungen*

> Ist eine Zufallsvariable normalverteilt mit dem Erwartungswert 0 und der Varianz 1, so liegt eine **Standardnormalverteilung** vor. In diesem Zusammenhang wir die Zufallsvariable mit Z bezeichnet, d.h. $\boldsymbol{Z \sim N(0,1)}$.

Standardnormalverteilung

3.3.2 Lineare Transformation einer Normalverteilung

In Abschnitt 2.4.2 wurden lineare Transformationen von Zufallsvariablen betrachtet. Die Normalverteilung besitzt in diesem Zusammenhang die Eigenschaft:

lineare Transformation einer Normalverteilung

> Jede lineare Transformation $\boldsymbol{Y = aX+b}$ einer normalverteilten Zufallsvariablen X mit dem Erwartungswert μ und der Varianz σ^2 ergibt eine **normalverteilte Zufallsvariable** $\boldsymbol{Y}$.

Die Parameter $E(Y)$ und $\mathrm{Var}(Y)$ ergeben sich aus den in Abschnitt 2.4.2 eingeführten Regeln zu

$$\begin{aligned} E(Y) &= aE(X) + b = a\mu + b, \\ \mathrm{Var}(Y) &= a^2\mathrm{Var}(X) = a^2\sigma^2. \end{aligned}$$

Durch eine geeignete Wahl von a und b lässt sich nun erreichen, dass durch die lineare Transformation eine $N(\mu, \sigma^2)$-verteilte Zufallsvariable X in eine $N(0, 1)$-verteilte, also eine standardnormalverteilte Zufallsvariable Z transformiert werden kann.

Es muss dann gelten:

$$\begin{aligned} E(Z) &= 0 = a\mu + b, \\ \mathrm{Var}(Z) &= 1 = a^2\sigma^2. \end{aligned}$$

Aus diesen Gleichungen ergibt sich

$$a = \frac{1}{\sigma} \quad \text{und} \quad b = -\frac{\mu}{\sigma}.$$

Jede $N(\mu, \sigma^2)$-verteilte Zufallsvariable X geht also durch die lineare Transformation

$$Z = \frac{1}{\sigma}X - \frac{\mu}{\sigma} = \frac{X - \mu}{\sigma}$$

in die $N(0, 1)$-verteilte Zufallsvariable Z über bzw. durch Standardisierung einer normalverteilten Zufallsvariablen X ergibt sich die standardnormalverteilte Zufallsvariable Z (s. auch Abschnitt 2.4.3).

Zur Bestimmung von Wahrscheinlichkeiten über bestimmte Intervalle werden, soweit diese nicht direkt numerisch berechnet werden, von daher nur die Tabellen der Standardnormalverteilung benötigt, aus denen die Wahrscheinlichkeiten $F_Z(z) = P(Z \leq z)$ abgelesen werden können.

3.3.3 Bestimmung von Wahrscheinlichkeiten

Grafisch gesehen ist die Wahrscheinlichkeit dafür, dass X zwischen x_1 und x_2 liegt, durch die Fläche unterhalb der Dichtefunktion der Normalverteilung zwischen x_1 und x_2 gegeben.

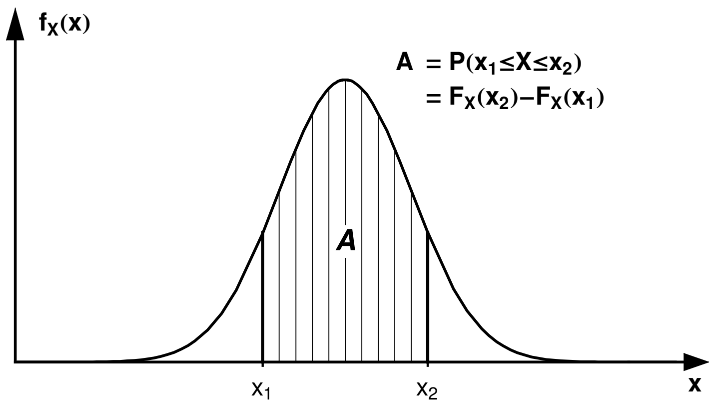


Abbildung 3.3.3: *Grafische Veranschaulichung von* $\boldsymbol{P}(x_1 \leq X \leq x_2)$

In Abschnitt 2.2.2 wurde bereits gezeigt, dass sich diese Wahrscheinlichkeit wie folgt angeben läßt:

$$\begin{aligned}\boldsymbol{P}(x_1 \leq X \leq x_2) &= \int_{x_1}^{x_2} f_X(x)dx = \int_{-\infty}^{x_2} f_X(x)dx - \int_{-\infty}^{x_1} f_X(x)dx \\ &= F_X(x_2) - F_X(x_1).\end{aligned}$$

Die Ungleichung

$$(1) \quad x_1 \leq X \leq x_2$$

kann nun schrittweise umgeformt werden in

$$x_1 - \mu \leq X - \mu \leq x_2 - \mu$$

und schließlich zu

$$(2) \quad \frac{x_i - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{x_2 - \mu}{\sigma}.$$

In der Mitte dieser Ungleichung steht nun die standardnormalverteilte Zufallsvariable Z mit

$$Z = \frac{X - \mu}{\sigma}.$$

Da die Ungleichungen (1) und (2) äquivalent sind, ist auch die Wahrscheinlichkeit, dass X zwischen x_1 und x_2 liegt, genauso groß wie die, dass Z zwischen

$$\frac{x_1 - \mu}{\sigma} = z_1 \quad \text{und} \quad \frac{x_2 - \mu}{\sigma} = z_2$$

liegt. Jedes Intervall einer beliebigen Normalverteilung kann so in ein „gleichwahrscheinliches" Intervall der Standardnormalverteilung überführt werden. Es gilt für jede $N(\mu, \sigma^2)$-verteilte Zufallsvariable X:

$$\begin{aligned} F_X(x_2) - F_X(x_1) &= \boldsymbol{P}(x_1 \leq X \leq x_2) \\ &= \boldsymbol{P}\left(\frac{x_1 - \mu}{\sigma} \leq Z \leq \frac{x_2 - \mu}{\sigma}\right) \\ &= \boldsymbol{P}(z_1 \leq Z \leq z_2) \\ &= F_Z(z_2) - F_Z(z_1). \end{aligned}$$

Die Wahrscheinlichkeit $\boldsymbol{P}(x_1 \leq X \leq x_2) = \boldsymbol{P}(z_1 \leq Z \leq z_2)$ entspricht somit der Differenz $F_Z(z_2) - F_Z(z_1)$, wobei die einzelnen Wahrscheinlichkeiten $F_Z(z_1)$ und $F_Z(z_2)$ mittels der Tabelle der Standardnormalverteilung (s. Gesamtglossar) bestimmt werden können.

Beispiel 3.3.1:

X ist $N(3, 16)$-verteilt. Mit welcher Wahrscheinlichkeit fällt X in das Intervall $3 \leq X \leq 7$?

Es gilt:

$$\begin{aligned} F_X(7) - F_X(3) &= \boldsymbol{P}(3 \leq X \leq 7) = \boldsymbol{P}\left(\frac{3-3}{4} \leq Z \leq \frac{7-3}{4}\right) \\ &= \boldsymbol{P}(0 \leq Z \leq 1) = F_Z(1) - F_Z(0) \\ &= 0.8413 - 0.5 = 0.3413. \end{aligned}$$

Es ist also die Wahrscheinlichkeit gesucht, mit der die standardnormalverteilte Zufallsvariable Z in das Intervall $[0, 1]$ fällt. Beide in Abbildung 3.3.4 schraffierten Flächen sind somit gleich.

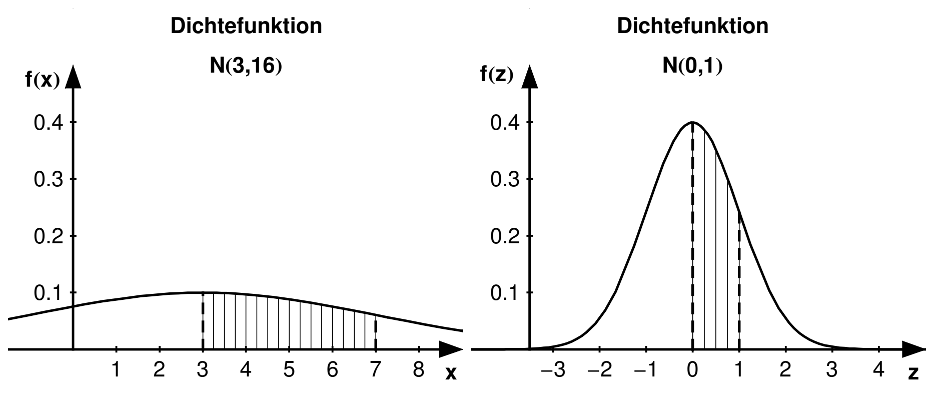


Abbildung 3.3.4: *Transformation in die Standardnormalverteilung*

Andererseits kann jedes Intervall einer standardnormalverteilten Zufallsvariablen auf ein „gleichwahrscheinliches“ Intervall einer $N(\mu, \sigma^2)$-verteilten Zufallsvariablen geführt werden. Dazu wird $Z = \frac{X-\mu}{\sigma}$ nach X aufgelöst, d.h. $X = \mu + Z\sigma$.

Die Wahrscheinlichkeit, dass Z zwischen z_1 und z_2 liegt, ist gleich der Wahrscheinlichkeit, dass X im Intervall mit den Grenzen $\mu + z_1\sigma$ mit $\mu + z_2\sigma$ liegt, also zwischen der z_1-fachen und z_2-fachen Standardabweichung vom Erwartungswert entfernt.

$$\boldsymbol{P}(z_1 \leq Z \leq z_2) = \boldsymbol{P}(\mu + z_1\sigma \leq X \leq \mu + z_2\sigma).$$

Beispiel 3.3.2:

Eine standardnormalverteilte Zufallsvariable Z fällt mit etwa 95.45% Wahrscheinlichkeit in das Intervall $-2 \leq Z \leq 2$.

Es ist also $\boldsymbol{P}(-2 \leq Z \leq 2) = 0.9545$. Für eine $N(3,16)$-verteilte Zufallsvariable X gilt:

$$\boldsymbol{P}(3 - 2 \cdot 4 \leq X \leq 3 + 2 \cdot 4) = \boldsymbol{P}(-5 \leq X \leq 11) = 0.9545.$$

Die normalverteilte Zufallsvariable X fällt daher mit etwa 95.45% Wahrscheinlichkeit in das Intervall $-5 \leq X \leq 11$.

Werden die Wahrscheinlichkeiten für Intervalle nicht direkt über numerische Integrationsverfahren bestimmt, werden die hier beschriebenen

Zusammenhänge ausgenutzt, um die gesuchten Wahrscheinlichkeiten leichter bestimmen zu können. Aufgrund der Tatsache, dass jede beliebige Normalverteilung sich in die Standardnormalverteilung überführen lässt, werden nur die Tabellen der Standardnormalverteilung benötigt, in denen die Wahrscheinlichkeiten für bestimmte Intervalle angegeben sind.

Bei der Benutzung solcher Tabellen ist zu beachten, dass die Normalverteilung eine symmetrische Verteilung ist. Die Wahrscheinlichkeit, dass die standardnormalverteilte Zufallsvariable Z zwischen $-z$ und 0 liegt, ist also genauso groß wie die Wahrscheinlichkeit, dass sie zwischen 0 und z liegt.

Um die Handhabung und Anwendung von Tabellen der Normalverteilung zu erleichtern, sind im Gesamtglossar mehrere Tabellen der Standardnormalverteilung angegeben, nämlich für $0 \leq z \leq 3$ die Werte F_Z, F_1 und F_2.

$$
\begin{aligned}
&(1) \quad F_Z &=& \quad \boldsymbol{P}(-\infty < Z \leq z) = \boldsymbol{P}(-\infty < X \leq \mu + z\sigma) \\
&(2) \quad F_1 &=& \quad \boldsymbol{P}(0 \leq Z \leq z) = \boldsymbol{P}(\mu \leq X \leq \mu + z\sigma) \\
&(3) \quad F_2 &=& \quad \boldsymbol{P}(-z \leq Z \leq z) = \boldsymbol{P}(\mu - z\sigma \leq X \leq \mu + z\sigma)
\end{aligned}
$$

Das z in (1) entspricht somit dem Quantil z_p mit $\boldsymbol{P}(Z \leq z) = p = F_z$. Grafisch kann die Wahrscheinlichkeit, dass die stetige Zufallsvariable Z in einem bestimmten Intervall liegt, als Fläche unterhalb der Dichte über dem betreffenden Intervall veranschaulicht werden. Die folgende Zeichnung verdeutlicht, welche Wahrscheinlichkeiten in der Tabelle des Glossars zusammengestellt sind.

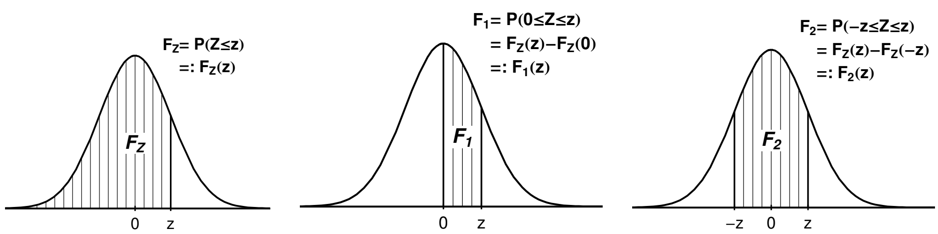
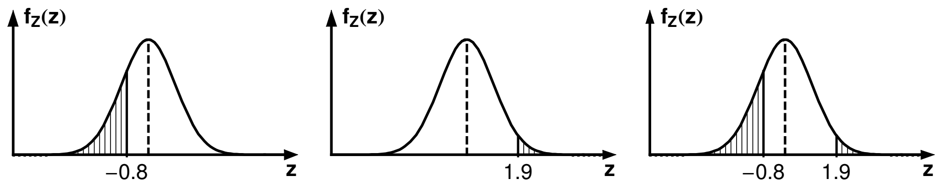


Abbildung 3.3.5: *Wahrscheinlichkeiten F_Z, F_1, F_2 einer standardnormalverteilten Zufallsvariablen*

Beispiel 3.3.3:

Im Folgenden werden für verschiedene Intervalle einer Standardnormalverteilung Wahrscheinlichkeiten bestimmt. Es ist zu beachten, dass nicht alle gesuchten Wahrscheinlichkeiten unmittelbar aus der Tabelle im Glossar abgelesen werden können. Zur Berechnung der Wahrscheinlichkeiten werden zusätzlich folgende Eigenschaften der Normalverteilung benötigt.

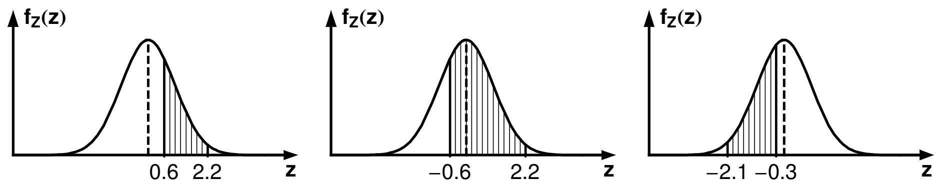
- *Die Dichtefunktion der Normalverteilung ist symmetrisch.*
- *Die Wahrscheinlichkeit, dass die Zufallsvariable größer als der Erwartungswert ist, ist gleich der Wahrscheinlichkeit, dass die Zufallsvariable einen Wert kleiner als der Erwartungswert annimmt und zwar 0.5. Diese Eigenschaft resultiert aus der Symmetrie der Verteilung.*
- *Wahrscheinlichkeiten für Intervalle lassen sich gegebenenfalls durch Differenzenbildung (Differenz zweier bekannter bzw. mit Hilfe der Tabelle zu ermittelnder Flächen) bestimmen.*



$$
\begin{aligned}
\boldsymbol{P}(Z < -0.8) &= 0.5 - F_1(0.8) = 0.5 - 0.2881 = 0.2119 \\
\boldsymbol{P}(Z < -0.8) &= \boldsymbol{P}(Z > 0.8) = 1 - \boldsymbol{P}(Z < 0.8) \\
&= 1 - F_Z(0.8) = 1 - 0.7881 = 0.2119
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{P}(Z > 1.9) &= 0.5 - F_1(1.9) = 0.5 - 0.4713 = 0.0287 \\
\boldsymbol{P}(Z > 1.9) &= 1 - F_Z(1.9) = 1 - 0.9713 = 0.0287
\end{aligned}
$$

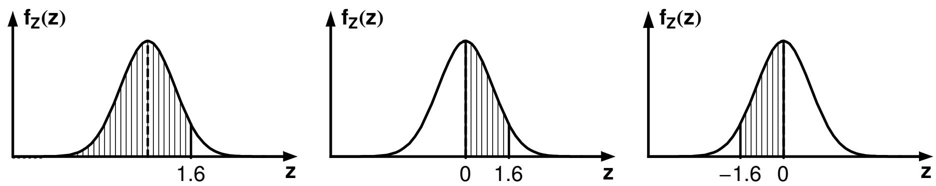
$$
\begin{aligned}
\boldsymbol{P}(Z < -0.8 \text{ oder } Z > 1.9)) &= \boldsymbol{P}(Z < -0.8) + \boldsymbol{P}(Z > 1.9) \\
&= 0.2119 + 0.0287 = 0.2406
\end{aligned}
$$



$$\begin{aligned}
\boldsymbol{P}(0.6 < Z < 2.2) &= F_1(2.2) - F_1(0.6) = 0.4861 - 0.2257 = 0.2604 \\
\boldsymbol{P}(0.6 < Z < 2.2) &= F_Z(2.2) - F_Z(0.6) = 0.9861 - 0.7257 = 0.2604
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{P}(-0.6 < Z < 2.2) &= F_1(0.6) + F_1(2.2) = 0.2257 + 0.4861 = 0.7118 \\
\boldsymbol{P}(-0.6 < Z < 2.2) &= F_Z(2.2) - F_Z(-0.6) = F_Z(2.2) - 1 + F_Z(0.6) \\
&= 0.9861 - 1 + 0.7257 = 0.7118
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{P}(-2.1 < Z < -0.3) &= F_1(2.1) - F_1(0.3) = 0.4821 - 0.1179 = 0.3642 \\
\boldsymbol{P}(-2.1 < Z < -0.3) &= F_Z(2.1) - F_Z(0.3) = 0.9821 - 0.6179 = 0.3642
\end{aligned}$$



$$\boldsymbol{P}(Z < 1.6) = F_Z(1.6) = 0.9452$$

$$\begin{aligned}
\boldsymbol{P}(0 < Z < 1.6) &= F_1(1.6) = 0.4452 \\
\boldsymbol{P}(0 < Z < 1.6) &= F_Z(1.6) - F_Z(0) = 0.9452 - 0.5 = 0.4452
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{P}(-1.6 < Z < 0) &= F_Z(0) - F_Z(-1.6) = 0.5 - 1 + F_Z(1.6) \\
&= -0.5 + 0.9452 = 0.4452 \\
\boldsymbol{P}(-1.6 < Z < 0) &= \boldsymbol{P}(0 < Z < 1.6) = F_1(1.6) = 0.4452
\end{aligned}$$

In späteren Ausführungen wird die Bezeichnung z_p verwendet, womit der Wert der standardnormalverteilten Zufallsvariablen Z gemeint ist, bei dem die Verteilungsfunktion den Wert p annimmt. Der Wert z_p entspricht somit dem p-Quantil.

Quantile

$$F_Z(z_p) = \boldsymbol{P}(Z \leq z_p) = p.$$

Wichtige z-Quantile sind:

$z_{0.95} = 1.65 \quad z_{0.975} = 1.96 \quad z_{0.99} = 2.33 \quad z_{0.995} = 2.58$

3.3.4 Der zentrale Grenzwertsatz

Der wohl wichtigste Grund für die Häufigkeit des Auftretens der Normalverteilung ergibt sich aus den sogenannten zentralen Grenzwertsätzen. Deren Sinn ist es zu zeigen, unter welchen Bedingungen die Verteilungsfunktionen der Summen von Zufallsvariablen gegen die Verteilungsfunktion der Normalverteilung konvergieren. Es existieren mehrere zentrale Grenzwertsätze, die alle die gleiche Konvergenzaussage unter verschiedenen Voraussetzungen beinhalten. Wichtig dabei ist, dass diese Voraussetzungen sehr allgemein sind, was oft eine Annahme der Normalverteilung rechtfertigt. In vielen experimentellen Fällen setzt sich die betrachtete Größe additiv aus einer großen Anzahl unabhängiger Zufallsvariablen zusammen, so dass nach den zentralen Grenzwertsätzen auf die Normalität der Wahrscheinlichkeitsverteilung dieser Größe geschlossen werden kann. Ein typisches Beispiel dafür sind Messfehler. Der resultierende Fehler setzt sich hier aus vielen verschiedenen kleinen Fehlern zusammen.

Liegen n unabhängige Zufallsvariablen $X_1, X_2, ..., X_n$ mit der gleichen Verteilung, dem gleichen Erwartungswert μ und der gleichen Varianz σ^2 vor, so besagt im Allgemeinen der zentrale Grenzwertsatz, dass die Summe der stochastisch unabhängigen Zufallsvariablen annähernd normalverteilt ist mit Erwarungswert $n\mu$ und Varianz $n\sigma^2$.

$$\begin{aligned} E\left(\sum_{i=1}^{n} X_i\right) &= \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} \mu = n\mu \\ \mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) &= \sum_{i=1}^{n} \mathrm{Var}(X_i) = \sum_{i=1}^{n} \sigma^2 = n\sigma^2 \end{aligned}$$

Die wohl einfachste Version des Grenzwertsatzes kann wie folgt formuliert werden:

Zentraler Grenzwertsatz

Gegeben seien die unabhängigen Zufallsvariablen $X_1, X_2, ..., X_n$, die alle die gleiche Verteilung und damit auch den gleichen Erwartungswert μ und die gleiche Varianz σ^2 besitzen. Dann ist die Zufallsvariable

$$Z_n = \frac{\sum\limits_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \quad \text{bzw.} \quad Z_n = \frac{\frac{1}{n}\sum\limits_{i=1}^{n} X_i - \mu}{\sigma/\sqrt{n}}$$

näherungsweise standardnormalverteilt, d.h. $N(0,1)$-verteilt. Für $n \to \infty$ konvergiert die Verteilung von Z_n gegen die Standardnormalverteilung.

3.3.5 Ergänzende Bemerkungen

Auch die **Normalverteilung** ist reproduktiv.

Reproduktivität der Normalverteilung

Sind X_1 und X_2 unabhängige Zufallsvariablen und $N(\mu_1, \sigma_1^2)$- bzw. $N(\mu_2, \sigma_2^2)$-verteilt, so gilt für die zusammengesetzte Zufallsvariable $\boldsymbol{X = a_1 X_1 + a_2 X_2 + b}$:

$$X \sim N(a_1\mu_1 + a_2\mu_2 + b, a_1^2\sigma_1^2 + a_2^2\sigma_2^2).$$

Beispiel 3.3.4:

Es sind $X_1 \sim N(4, 25)$ und $X_2 \sim N(5, 36)$. Dann gilt:

$$\begin{aligned} 2X_1 - 3X_2 - 5 &\sim N(2\mu_1 - 3\mu_2 - 5, 4\sigma_1^2 + 9^2\sigma_2^2) \\ &\sim N(-12, 424). \end{aligned}$$

Für die große Bedeutung der Normalverteilung gibt es verschiedene Gründe. Die wichtigsten sind:

1. Viele Zufallsvariablen, die bei Experimenten und Beobachtungen in der Praxis auftreten, sind annähernd normalverteilt. Wird etwa aufgrund einer Stichprobe vermutet, dass eine bestimmte Grundgesamtheit eine eingipflige Verteilung besitzt, ohne dass Näheres bekannt ist, so führt die Annahme einer Normalverteilung in vielen Fällen zu sinnvollen und praktisch brauchbaren Ergebnissen.

2. Bisweilen ist eine Transformation von nicht normalverteilten Zufallsvariablen in eine Normalverteilung möglich.

3. Gewisse komplizierte Verteilungen lassen sich in Grenzfällen durch die Normalverteilung brauchbar approximieren (zentrale Grenzwertsätze).

1,2,3-σ-Regel

Kurz erwähnt sei an dieser Stelle noch die **1,2,3-$\boldsymbol{\sigma}$-Regel** die besagt, dass die Wahrscheinlichkeit, dass die Zufallsvariable X um höchstens 1σ, 2σ, 3σ vom Erwartungswert μ abweicht 0.6827, 0.9545, 0.9973 beträgt. Formal:

$$\begin{aligned} \boldsymbol{P}(\mu - 1\sigma \leq X \leq \mu + 1\sigma) &= 0.6827 \\ \boldsymbol{P}(\mu - 2\sigma \leq X \leq \mu + 2\sigma) &= 0.9545 \\ \boldsymbol{P}(\mu - 3\sigma \leq X \leq \mu + 3\sigma) &= 0.9973 \end{aligned}$$

3.4 Approximation der Binomialverteilung durch die Normalverteilung

Die Anwendung der Binomialverteilung bereitet mitunter erheblichen Rechenaufwand bei der Bestimmung von Wahrscheinlichkeiten. Aufgrund der Gültigkeit des zentralen Grenzwertsatzes ist es möglich die Binomialverteilung durch die Normalverteilung zu approximieren.

Approximation

Für $n\pi \geq 5, n(1-\pi) \geq 5$ ist eine $B(n,\pi)$-verteilte Zufallsvariable X **approximativ** normalverteilt mit $X \overset{a}{\sim} N(n\pi, n\pi(1-\pi))$. Die Zufallsvariable Z mit

$$Z = \frac{X - n\pi}{\sqrt{n\pi(1-\pi)}}$$

ist dann approximativ standardnormalverteilt ($Z \overset{a}{\sim} N(0,1)$).

Um für die Approximation einer diskreten Verteilung durch eine stetige Verteilung eine Verbesserung zu erreichen, wird die sogenannte Stetigkeitskorrektur vorgenommen, welche auf folgenden Überlegungen basiert.

Die Wahrscheinlichkeit für einen bestimmten Wert der Zufallsvariablen ist, wenn sie im diskreten Fall positiv ist, bei der stetigen Verteilung Null ($\boldsymbol{P}(X = x_i) = 0$). Um mittels der Approximationsverteilung $\boldsymbol{P}(X = x_i)$ zu bestimmen, wird deshalb für ganzzahlige x_i

$$F_X(x_i + 0.5) - F_X(x_i - 0.5)$$

berechnet. Dies basiert auf der Überlegung, dass bei dem Übergang einer diskreten auf eine stetige Verteilung Säulen mit einer Breite von eins um die Werte x_i gelegt werden. Für die Berechnung der Wahrscheinlichkeiten von Intervallen ganzzahliger Zufallsvariablen werden entsprechend die Intervallgrenzen um 0.5 verlegt. Dieses Vorgehen wird als **Stetigkeitskorrektur** bezeichnet.

Stetigkeitskorrektur

Die Zufallsvariable X sei $B(n,\pi)$-verteilt. Unter Anwendung der **Stetigkeitskorrektur** wird für $n\pi \geq 5, n(1-\pi) \geq 5$

$$\begin{aligned}
\boldsymbol{P}(X \leq x) &:= \boldsymbol{P}\left(Z \leq \frac{x + 0.5 - n\pi}{\sqrt{n\pi(1-\pi)}}\right), \\
\boldsymbol{P}(x_1 \leq X \leq x_2) &:= \boldsymbol{P}\left(\frac{x_1 - 0.5 - n\pi}{\sqrt{n\pi(1-\pi)}} \leq Z \leq \frac{x_2 + 0.5 - n\pi}{\sqrt{n\pi(1-\pi)}}\right),
\end{aligned}$$

berechnet, wobei Z approximativ standardnormalverteilt ist.

Beispiel 3.4.1:

Aus einer Urne mit einem Anteil von $\pi = 0.2$ an schwarzen Kugeln werden 100 Kugeln entnommen. Wie groß ist die Wahrscheinlichkeit genau 15 bzw. mindestens 15 und höchstens 28 schwarze Kugeln zu finden?

Wegen $n\pi = 20 \geq 5, n(1-\pi) = 80 \geq 5$ ist die Zufallsvariable X (Anzahl der schwarzen Kugeln) näherungsweise normalverteilt mit den Parametern $\mu = n\pi = 20$ und $\sigma^2 = n\pi(1-\pi) = 16$.

$$\begin{aligned}
\boldsymbol{P}(X = 15) &\approx \boldsymbol{P}\left(\frac{15 + 0.5 - 20}{4} \leq Z\right) - \boldsymbol{P}\left(\frac{15 - 0.5 - 20}{4} \leq Z\right) \\
&= \boldsymbol{P}(-1.125 \leq Z) - \boldsymbol{P}(-1.375 \leq Z) \\
&= F_1(1.375) - F_1(1.125) = 0.4155 - 0.3697 = 0.0458 \\
\boldsymbol{P}(15 \leq X \leq 28) &\approx \boldsymbol{P}\left(\frac{15 - 0.5 - 20}{4} \leq Z \leq \frac{28 + 0.5 - 20}{4}\right) \\
&= F_Z(2.125) - F_Z(-1.375) \\
&= F_Z(2.125) - 1 + F_Z(1.375) \\
&= F_1(1.375) + F_1(2.125) = 0.4155 + 0.4832 = 0.8987
\end{aligned}$$

3.5 χ^2-Verteilung (Chi-Quadrat-Verteilung)

Die χ^2-Verteilung findet oft Anwendung im Bereich der Testverfahren. Sie beschreibt die Summe unabhängiger quadrierter standardnormalverteilter Zufallsvariablen. Von Interesse ist hauptsächlich die Bestimmung der Quantile und die Approximationsmöglichkeit durch die Normalverteilung. Daher wird hier nicht näher auf die Dichtefunktion bzw. die Verteilungsfunktion eingegangen.

χ^2-Verteilung

Sind $X_1, X_2, ..., X_n$ standardnormalverteilte stochastisch unabhängige Zufallsvariablen, so ist die Zufallsvariable

$$Y = \sum_{i=1}^{n} X_i^2$$

χ^2-verteilt mit n Freiheitsgraden ($Y \sim \chi^2(n)$). Es ist

Erwartungswert und Varianz

$$E(Y) = n \quad \text{und} \quad \mathrm{Var}(Y) = 2n.$$

Freiheitsgrad

Allgemein entspricht die Anzahl der **Freiheitsgrade** der Anzahl unabhängiger Einzelinformationen minus der Anzahl der in die Berechnung der jeweiligen Größe eingehenden zusätzlichen Informationen. Die Anzahl der Freiheitsgraden ergibt sich hier aus der Anzahl der unabhängigen Zufallsvariablen, welche die Summe $\sum_{i=1}^{n} X_i^2$ bilden. Aufgrund der Unabhängigkeit der Zufallsvariablen $X_1, X_2, ..., X_n$, d.h. es liegt keine zusätzliche Information über die Zufallsvariablen vor, können alle einen beliebigen „freien" Wert annehmen. Daher entspricht die Anzahl der Freiheitsgrade hier n.
Für negative y ist die Dichte der χ^2-Verteilung 0. Die χ^2-Verteilung ist asymmetrisch und die Gestalt hängt von dem Parameter $n \in \mathbb{N}$ ab.

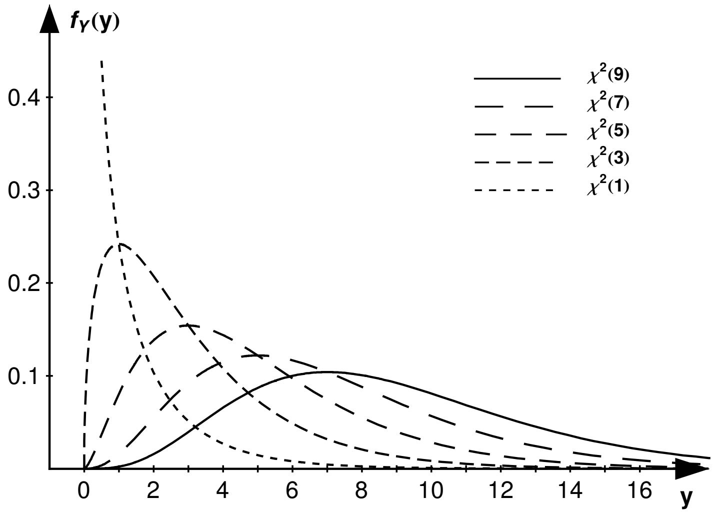


Abbildung 3.5.1: *χ^2-Verteilung*

Für wachsende Stichprobenumfänge wird die Verteilung immer symmetrischer.[6]

Reproduktivität der χ^2-Verteilung

> Sind $Y_1, ..., Y_m$ unabhängige χ^2-verteilte Zufallsvariablen mit $n_1, ..., n_m$ Freiheitsgraden, so ist $Y = \sum_{i=1}^{m} Y_i$ eine χ^2-verteilte Zufallsvariable mit $n = \sum_{i=1}^{m} n_i$ Freiheitsgraden.

Für $n > 100$ kann die χ^2-Verteilung direkt durch eine $N(n, 2n)$-Verteilung approximiert werden. Da hierbei die Symmetrieeigenschaft noch nicht optimal erfüllt ist, ist eine indirekte Approximation durch die Normalverteilung vorzuziehen. Die Zufallsvariable $Z = \sqrt{2Y} - \sqrt{2n-1}$ ist für $n > 30$ näherungsweise standardnormalverteilt.

Die Quantile der Chi-Quadrat-verteilten Zufallsvariablen werden im Folgenden mit $\chi^2_p(n)$ bezeichnet. Dies entspricht dem Wert der mit n Freiheitsgraden χ^2-verteilten Zufallsvariablen, an dessen Stelle die Verteilungsfunktion den Wert p annimmt. Die Zufallsvariable ist also mit einer Wahrscheinlichkeit von p kleiner als $\chi^2_p(n)$.

Beispiel 3.5.1:

Gegeben ist $X \sim \chi^2(3)$. Es sollen x_1, x_2 bestimmt werden, für die gilt:

$$\boldsymbol{P}(X \leq x_1) = 0.25, \qquad \boldsymbol{P}(X \geq x_2) = 0.25.$$

Gesucht sind somit das 0.25-Quantil (untere Quartil) und das 0.75-Quantil (obere Quartil) der χ^2-Verteilung, denn für x_2 gilt äquivalent $\boldsymbol{P}(X \leq x_2) = 0.75$. Die Quantile sind in der folgenden Graphik eingezeichnet.

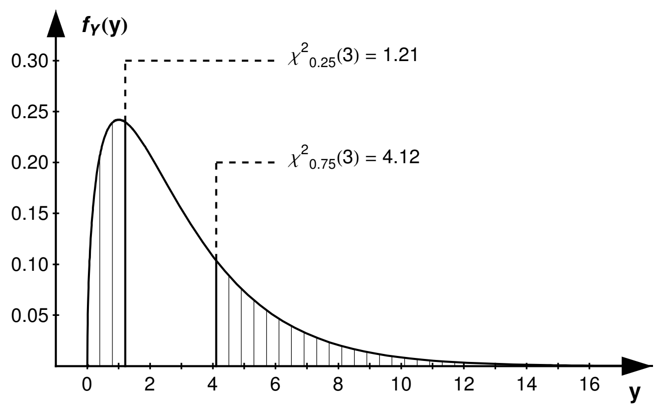


Abbildung 3.5.2: *Das 0.25-Quantil und das 0.75-Quantil einer χ^2-Verteilung mit 3 Freiheitsgraden*

[6] siehe interaktive Mathematica-Applets auf der Homepage des Lehrstuhls http://www.fernuni-hagen.de/ls_statistik/forschung/multimedia/

3.6 t-Verteilung (Student-Verteilung)

Die t-Verteilung wird oft für Parametertests und für Konfidenzintervalle benötigt. Auch an dieser Stelle ist lediglich die Bestimmung der Quantile und die Approximationsmöglichkeit der t-Verteilung von Interesse.

t-Verteilung

Es seien X und Y unabhängige Zufallsvariablen mit $X \sim N(0,1)$ und $Y \sim \chi^2(n)$. Dann ist

$$T = \frac{X}{\sqrt{Y/n}}$$

Student-verteilt oder **t-verteilt mit n Freiheitsgraden ($T \sim t(n)$)**. Die Parameter der t-Verteilung lauten:

Erwartungswert und Varianz

$$\begin{array}{llcll} \text{Erwartungswert:} & E(T) & = & 0 & \text{für } n \geq 2, \\ \text{Varianz:} & \operatorname{Var}(T) & = & \dfrac{n}{n-2} & \text{für } n \geq 3. \end{array}$$

Für $n > 30$ ist eine Student-verteilte Zufallsvariable näherungsweise $N(0,1)$-verteilt, d.h. für $n > 30$ kann die Student-Verteilung durch die Standardnormalverteilung approximiert werden.[7]

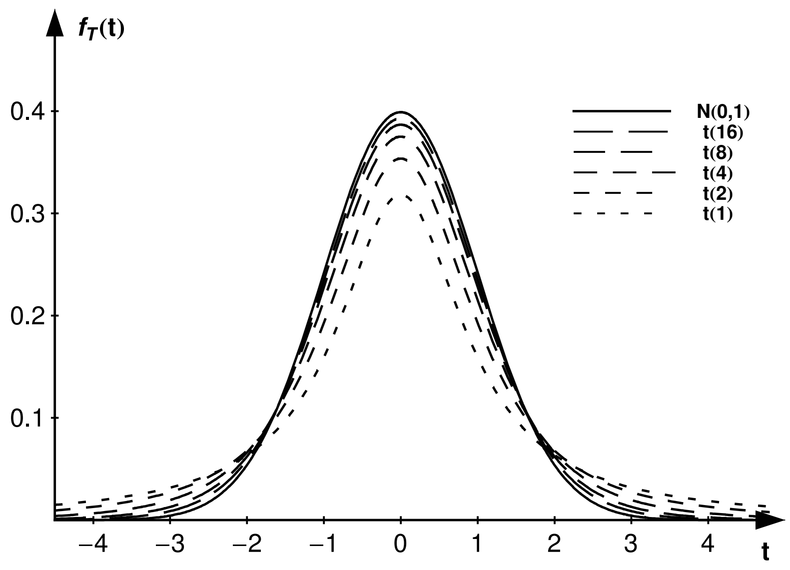


Abbildung 3.6.1: *Approximation der t-Verteilung durch die Normalverteilung*

[7]siehe interaktive Mathematica-Applets auf der Homepage des Lehrstuhls http://www.fernuni-hagen.de/ls_statistik/forschung/multimedia/

In der Tabelle der t-Verteilung (s. Gesamtglossar) sind die Wahrscheinlichkeiten $P(T < t)$ bei gegebener Anzahl von n Freiheitsgraden aufgeführt.

Beispiel 3.6.1:

Gegeben ist eine t-verteilte Zufallsvariable mit $n = 20$ Freiheitsgraden. Gesucht ist das 0.95-Quantil der t-Verteilung, welches hier mit $t_{0.95}(20)$ bezeichnet wird. Es gibt den Wert an, für den $\boldsymbol{P}(T \leq t) = 0.95$ gilt. Dieser Wert ist direkt aus der Tabelle im Gesamtglossar ablesbar. Es gilt $t_{0.95}(20) = 1.725$.

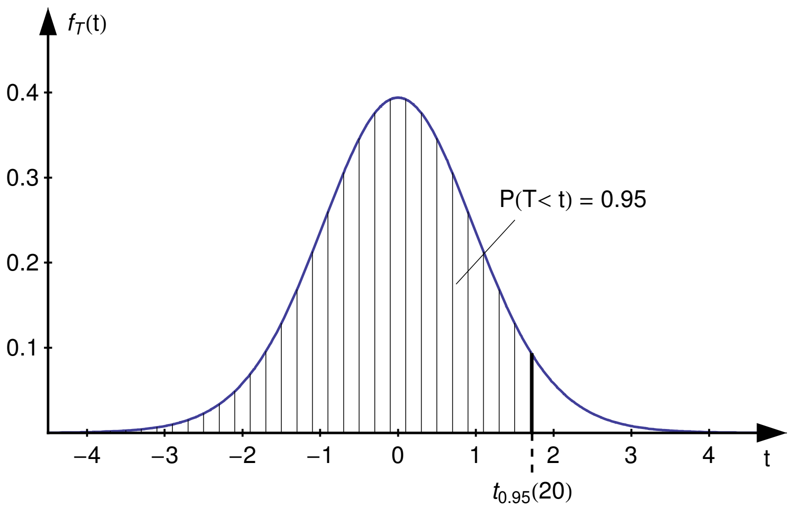


Abbildung 3.6.2: *Grafische Interpretation der tabellierten Werte einer t-Verteilung mit n Freiheitsgraden*

3.7 F-Verteilung (Fisher-Verteilung)

Die F-Verteilung ist eine stetige Wahrscheinlichkeitsverteilung, deren Dichtefunktion von zwei Parametern, den Freiheitsgraden n bzw. m abhängt. Sie entsteht aus standardnormalverteilten Zufallsvariablen. Die F-Verteilung wird bei Tests bezüglich des Verhältnisses verschiedener Varianzen benötigt.

F-**Verteilung**

Es sei $Y_1 \sim \chi^2(m)$ und $Y_2 \sim \chi^2(n)$, wobei Y_1 und Y_2 stochastisch unabhängig sind. Dann ist die Zufallsvariable

$$X = \frac{Y_1/m}{Y_2/n}$$

***F*-verteilt mit *m* und *n* Freiheitsgraden** (m: Freiheitsgrad des Zählers, n: Freiheitsgrad des Nenners). Die Schreibweise lautet $\boldsymbol{X \sim F(m,n)}$. Weiter gilt

Erwartungswert und Varianz

$$\begin{aligned} E(X) &= \frac{n}{n-2} \text{ für } n \geq 3, \\ \mathrm{Var}(X) &= \frac{2n^2(m+n-2)}{m(n-4)(n-2)^2} \text{ für } n \geq 5. \end{aligned}$$

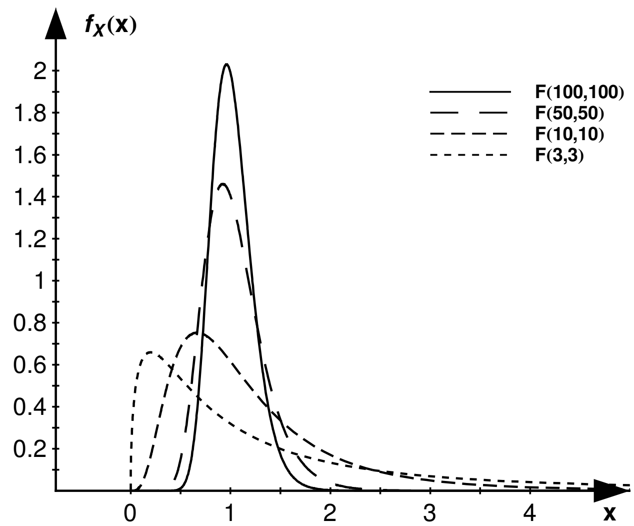


Abbildung 3.7.1: *F-Verteilung für verschiedene Freiheitsgrade*

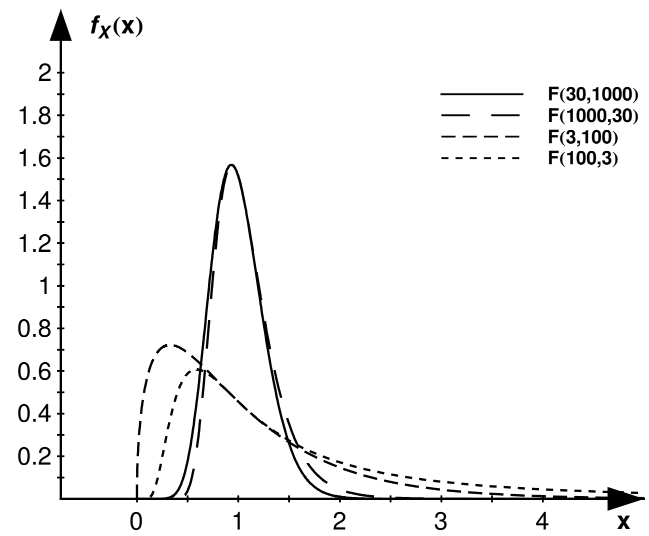


Abbildung 3.7.2: *F-Verteilung für verschiedene Freiheitsgrade*

Aus den Abbildungen ist zu erkennen, dass für große Stichprobenumfänge die F-Verteilung sich einer symmetrischen Verteilung nähert.[8]

Der Wert x einer F-verteilten Zufallsvariablen mit m und n Freiheitsgraden, bei dem die Verteilungsfunktion den Wert p annimmt wird mit $F_p(m,n)$ bezeichnet (p-Quantil).

Allgemein gilt bei der Bestimmung der Quantile:

$$
\begin{aligned}
&F_p(m,n) = \frac{1}{F_{1-p}(n,m)} \\
&m = 1,\, n \text{ beliebig:} \quad F_p(1,n) = [t_{(1+p)/2}(n)]^2 \\
&m = 1,\, n \geq 30\text{:} \qquad F_p(1,n) = (z_{(1+p)/2})^2 \\
&m > 1,\, n \geq 200\text{:} \quad\ F_p(m,n) = \tfrac{1}{m}\chi^2_p(m)
\end{aligned}
$$

In Abhängigkeit von m und n sind im Gesamtglossar die Werte x der F-verteilten Zufallsvariablen tabelliert, bei denen die Verteilungsfunktion den Wert $p = 0.95$ bzw. $p = 0.99$ annimmt.

[8]siehe interaktive Mathematica-Applets auf der Homepage des Lehrstuhls http://www.fernuni-hagen.de/ls_statistik/forschung/multimedia/

Beispiel 3.7.1:

Für eine mit $m = 50$ *und* $n = 15$ *Freiheitsgraden* F*-verteilte Zufallsvariable* X *gilt:*

$$\boldsymbol{P}(X \geq 2.18) = 0.05 \qquad \boldsymbol{P}(X \geq 3.08) = 0.01$$

Der Wert 2.18 entspricht somit dem 0.95*-Quantil der* F*-Verteilung mit* 50 *und* 15 *Freiheitsgraden* ($F_{0.95}(50, 15)$)*. Der Wert* 3.08 *entspricht dem Quantil* $F_{0.99}(50, 15)$*.*

Wird das Quantil $F_{0.05}(50, 15)$ *gesucht, also der Wert* x*, für den gilt* $\boldsymbol{P}(X \leq x) = 0.05$*, wird ausgenutzt, dass*

$$F_{0.05}(50, 15) = \frac{1}{F_{0.95}(15, 50)}$$

entspricht. Mit $F_{0.95}(15, 50) = 1.87$ *ergibt sich somit* $F_{0.05}(50, 15) = 0.535$*.*

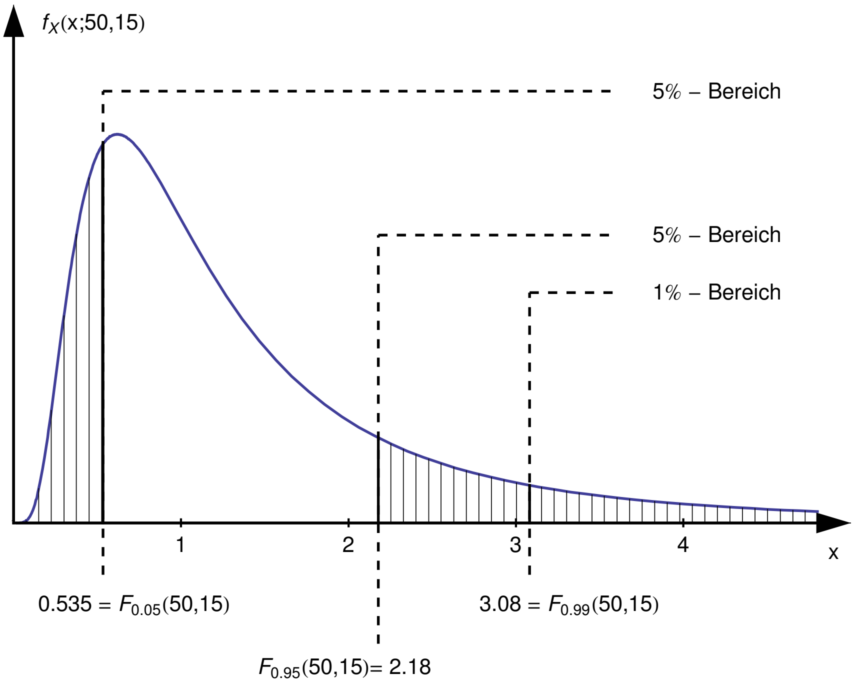


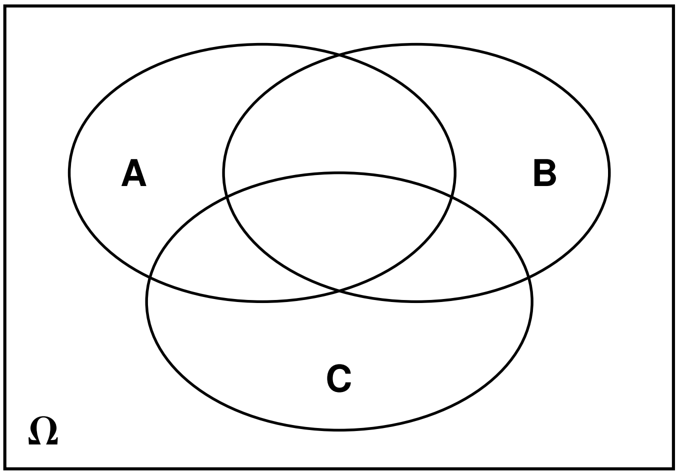
Abbildung 3.7.3: *Grafische Interpretation der tabellierten Werte der* F*-Verteilung (nicht maßstabsgetreu!)*

Übungsaufgaben

Aufgaben zu Kapitel 1

Übungsaufgabe 1.1:

Sei Ω ein Ergebnisraum und A, B, C drei Ereignisse, die eine gemeinsame Schnittmenge besitzen.



Markieren Sie folgende Ereignisse in Venn-Diagramme der obigen Form.

- $(A \cap B) \cap \overline{C}$
- $(A \cap B) \cup C$
- $(\overline{A} \cap \overline{B}) \cap C$
- $(A \cup B) \cap \overline{C}$

Übungsaufgabe 1.2:

Geben Sie zu folgenden Zufallsexperimenten die Ergebnisräume an:

a) *Werfen mit zwei Münzen, die nicht unterschieden werden können.*

$\Omega =$

b) *Gleichzeitiges Ziehen von zwei Kugeln aus einer Urne mit roten, grünen und blauen Kugeln; wobei von jeder Farbe mindestens zwei Kugeln in der Urne liegen.*

$\Omega =$

c) *Zweimaliges Ziehen einer Kugel aus einer Urne mit roten, grünen und blauen Kugeln. (Es wird erst die erste Kugel gezogen und die Farbe registriert. Danach wird die zweite Kugel gezogen und deren Farbe registriert.)*

$\Omega =$

Übungsaufgabe 1.3:

Es werden 5 *Münzen geworfen. Geben Sie alle Anordnungen an, bei denen genau zweimal Zahl* (Z) *oben liegt.*

Übungsaufgabe 1.4:

Aus einem Kartenspiel mit 32 Karten werden zufällig Karten gezogen. Welche Ereignisse schließen sich gegenseitig aus?

$A = \{\text{Kreuz}\}$ $B = \{8\}$ $C = \{\text{Karo 7, Karo 8, Karo 9}\}$ $D = \{\text{Herz 7}\}$

Übungsaufgabe 1.5:

a) *In einer Urne befinden sich* 3 *grüne und* 6 *rote Kugeln. Wie groß ist die Wahrscheinlichkeit, bei zufälliger Entnahme einer Kugel eine grüne zu greifen?*

$\boldsymbol{P}(\mathit{grün}) =$

b) *Eine ideale Münze wird* 3*mal nacheinander geworfen. Wie lauten die gleichmöglichen Fälle für das Ergebnis dieses Zufallsexperiments* $(K = Kopf, Z = Zahl)$*?*

Wie groß ist die Wahrscheinlichkeit für das Ereignis, dass genau zweimal Zahl oben liegt?

$\boldsymbol{P}(\textit{zweimal } Z) =$

Übungsaufgabe 1.6:

Für welche Ereignisse sind die Wahrscheinlichkeiten bei Verwendung des Wahrscheinlichkeitsbegriffs nach Laplace richtig angegeben?

a) *Beim zweimaligen Würfeln ist die Wahrscheinlichkeit für die Augensumme* 12 *gleich* $\frac{1}{36}$. *(Hinweis: Bild 1.5.2 veranschaulicht den Ergebnisraum.)*

b) *In einer Urne befinden sich* 7 *grüne und* 21 *blaue Kugeln. Die Wahrscheinlichkeit, eine grüne Kugel zu ziehen, beträgt* $\frac{1}{3}$.

Übungsaufgabe 1.7:

a) *Eine Urne enthält* 5 *verschieden farbige Kugeln. Es werden nacheinander* 3 *Kugeln gezogen, wobei jede Kugel nach dem Ziehen zurückgelegt wird.*
Wie viele verschiedene Möglichkeiten gibt es?

b) *Eine Urne enthält* 5 *verschieden farbige Kugeln. Es werden gleichzeitig* 3 *Kugeln gezogen.*
Wie viele verschiedene Möglichkeiten gibt es?

Übungsaufgabe 1.8:

In einer Schachtel sind 3 *rote,* 2 *blaue und* 4 *gelbe Perlen. Alle neun Perlen sollen aufgefädelt werden.*
Wie viele verschiedene Anordnungen gibt es?

Übungsaufgabe 1.9:

Ein Dominospiel besteht aus Steinen, die mit den Zahlen $0, 1, 2, 3, 4, 5, 6$ *beschriftet sind. Auf jedem Stein steht eine Zahlenkombination* $\{i, j\}$. *Aus wie vielen Steinen besteht ein Dominospiel, wenn jede Zahlenkombination* $\{i, j\}$ *mit* $i, j \in \{0, 1, ..., 6\}$ *einmal vorkommt?*

Übungsaufgabe 1.10:

In einer Lostrommel mit 1000 *gut gemischten Losen befinden sich* 10 *erste Gewinne* (I) *und* 80 *zweite Gewinne* (II). *Wird ein Los gezogen, so beträgt die Wahrscheinlichkeit für einen ersten Gewinn*

$$\boldsymbol{P}(I) =$$

und für einen zweiten Gewinn

$$\boldsymbol{P}(II) = \qquad\qquad .$$

Die Wahrscheinlichkeit für einen ersten oder zweiten Gewinn beträgt dann

$$\boldsymbol{P}(I \cup II) = \qquad\qquad .$$

Übungsaufgabe 1.11:

a) *In einer Urne befinden sich* 200 *Kugeln, von denen* 70 *blau sind und die übrigen gelb. Auf* 20 *blaue Kugeln und* 30 *gelbe Kugeln ist ein Stern gemalt. Wie groß ist die Wahrscheinlichkeit, dass eine zufällig gezogene Kugel blau ist oder mit einem Stern bemalt ist?*

b) *Aus einem Spiel mit* 32 *Karten wird zufällig eine Karte gezogen. Es ist* $\boldsymbol{P}(\text{„}Kreuz\text{“}) = 0.25$ *und* $\boldsymbol{P}(\text{„}As\text{“}) = 0.125$. *Wie groß ist* $\boldsymbol{P}(\text{„}Kreuz\text{“}oder\text{„}As\text{“})$*?*

Übungsaufgabe 1.12:

In welchen Fällen wird durch die Zuordnung von reellen Zahlen $\boldsymbol{P}(A), \boldsymbol{P}(B)$ und $\boldsymbol{P}(C)$ zu den Ereignissen A, B und C eine Wahrscheinlichkeit auf $\mathcal{P}(\Omega)$ definiert $(A \cup B \cup C = \Omega)$?

(1) $\boldsymbol{P}(A) = \frac{1}{2}$ $\boldsymbol{P}(B) = \frac{3}{4}$ $\boldsymbol{P}(C) = \frac{3}{2}$

(2) $\boldsymbol{P}(A) = \frac{1}{2}$ $\boldsymbol{P}(B) = \frac{1}{4}$ $\boldsymbol{P}(C) = \frac{1}{8}$

(3) $\boldsymbol{P}(A) = 1$ $\boldsymbol{P}(B) = \frac{1}{2}$ $\boldsymbol{P}(C) = \frac{3}{4}$

(4) $\boldsymbol{P}(A) = \frac{1}{4}$ $\boldsymbol{P}(B) = \frac{1}{4}$ $\boldsymbol{P}(C) = \frac{1}{2}$

(5) $\boldsymbol{P}(A) = \frac{1}{4}$ $\boldsymbol{P}(B) = -\frac{1}{4}$ $\boldsymbol{P}(C) = 1$

(6) $\boldsymbol{P}(A) = \frac{1}{2}$ $\boldsymbol{P}(B) = \frac{1}{2}$ $\boldsymbol{P}(A \cup B) = \frac{1}{4}$

Übungsaufgabe 1.13:

Es wird mit 10 Würfeln gewürfelt. Wie groß ist die Wahrscheinlichkeit dafür, dass die Augensumme größer als 11 ist?

Übungsaufgabe 1.14:

a) *In einer Urne befinden sich 7 blaue und 6 gelbe Kugeln. Es werden nacheinander (ohne Zurücklegen) zwei Kugeln gezogen. Wie groß ist die Wahrscheinlichkeit, beim zweiten Zug eine gelbe Kugel zu ziehen unter der Bedingung, dass (1) beim ersten Zug eine blaue Kugel bzw. (2) beim ersten Zug eine gelbe Kugel gezogen wurde?*

b) *Es wird ein roter und ein grüner Würfel geworfen. Wie groß ist die Wahrscheinlichkeit dafür, dass die Augensumme größer als 8 (A) ist unter der Bedingung, dass der grüne Würfel eine 4 (B) zeigt?*

Übungsaufgabe 1.15:

Betrachtet wird eine Familie mit zwei Kindern. Die Geburtswahrscheinlichkeit für Junge und Mädchen sei je 0.5*. Wie groß ist die Wahrscheinlichkeit dafür, dass beide Kinder Jungen sind, wenn*

a) keine sonstigen Angaben vorliegen?

b) bekannt ist, dass ein Kind ein Junge ist?

c) bekannt ist, dass das älteste Kind ein Junge ist?

Übungsaufgabe 1.16:

Ein Student besteht die Klausur im Fach Statistik mit der Wahrscheinlichkeit 0.7 *und im Fach Finanzmathematik mit der Wahrscheinlichkeit* 0.8*. Weiter sei die Wahrscheinlichkeit für das Bestehen beider Klausuren mit* 0.6 *angegeben.*

a) Sind die Ereignisse „Bestehen der Klausur im Fach Statistik“ und „Bestehen der Klausur im Fach Finanzmathematik“ unabhängig voneinander?

b) Wie groß ist die Wahrscheinlichkeit dafür, dass wenigstens eine Klausur bestanden wird?

Übungsaufgabe 1.17:

In einer Stadt erscheinen zwei Zeitungen. Zeitung 1 *wird von* 50% *der Erwachsenen gelesen.* 15% *der Erwachsenen lesen Zeitung* 1*, aber nicht Zeitung* 2*.* 20% *der Erwachsenen lesen Zeitung* 2*, aber nicht Zeitung* 1*. Zufällig wird ein Erwachsener ausgewählt.*

Mit welcher Wahrscheinlichkeit liest er

a) *wenigstens eine Zeitung?*

b) *beide Zeitungen?*

c) *höchstens eine Zeitung?*

d) *keine Zeitung?*

Übungsaufgabe 1.18:

Es wird mit einem Würfel zweimal geworfen.

a) *Bestimmen Sie die Wahrscheinlichkeit für das Auftreten der Augenzahl* 12*.*

b) *Bestimmen Sie die Wahrscheinlichkeit für das Auftreten der Augenzahl* 10*.*

Übungsaufgabe 1.19:

Ein Fußballspieler soll mit der Wahrscheinlichkeit von 0.5 *bei jedem Schuss auf das gegnerische Tor Erfolg haben, d.h. für seine Mannschaft ein Tor erzielen. Weiterhin soll angenommen werden, dass die Ergebnisse der einzelnen Schüsse voneinander unabhängig sind. Wie oft muss der betreffende Spieler mindestens auf das gegnerische Tor schießen, um mit der Wahrscheinlichkeit von* 0.99 *wenigstens ein Tor zu erzielen?*

Übungsaufgabe 1.20:

Eine Maschine besteht aus 3 Einzelaggregaten, die - unabhängig voneinander - mit den Wahrscheinlichkeiten 0.3, 0.2, 0.1 *ausfallen. Die Maschine kann nur genutzt werden, wenn keines der 3 Einzelaggregate ausfällt. Wie hoch ist die Wahrscheinlichkeit für den Ausfall der Maschine?*

Übungsaufgabe 1.21:

Drei Maschinen A, B und C produzieren 50%, 30% *bzw.* 20% *der gesamten Produktion eines Betriebes. Die Ausschussanteile der Maschinen betragen* 3%, 4% *bzw.* 5%.

a) *Wie groß ist die Wahrscheinlichkeit für das Ereignis X „Ein zufällig ausgewähltes Stück ist defekt“?*

b) *Wie groß ist die Wahrscheinlichkeit für das Ereignis „Ein zufällig ausgewähltes defektes Stück stammt von Maschine A“?*

Übungsaufgabe 1.22:

Für eine Infektionskrankheit wurde ein neuer medizinischer Test für die Diagnostik entwickelt. Von Interesse ist zu wissen, wie hoch die Sicherheit ist, dass aufgrund eines positiven Testergebnisses der Patient als infiziert eingestuft werden kann. Gegeben sind die Ereignisse I=„Patient ist infiziert“ *und* T=„Testergebnis ist positiv“ *mit den Wahrscheinlichkeiten*

$$\begin{aligned} \boldsymbol{P}(I) &= 0.001, \\ \boldsymbol{P}(T/I) &= 0.99, \\ \boldsymbol{P}(\overline{T}/\overline{I}) &= 0.98. \end{aligned}$$

a) *Wie groß ist die Wahrscheinlichkeit, dass bei Vorliegen eines positiven Testergebnisses der Patient tatsächlich infiziert ist?*

b) *Wie groß ist die Wahrscheinlichkeit, dass ein infizierter Patient nicht erkannt wird?*

Aufgaben zu Kapitel 2

Übungsaufgabe 2.1:

Gegeben ist die folgende diskrete Wahrscheinlichkeitsverteilung:

x_i	1	2	4	5
$f_X(x_i)$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{4}$

a) *Bestimmen Sie die Verteilungsfunktion.*

$$F_X(x) =$$

b) *Bestimmen Sie folgenden Wahrscheinlichkeiten.*

$$\boldsymbol{P}(0 < X < 4) =$$

$$\boldsymbol{P}(1 < X < 4) =$$

$$\boldsymbol{P}(1 \leq X \leq 4) =$$

$$\boldsymbol{P}(2 < X \leq 5) =$$

Übungsaufgabe 2.2:

Gegeben ist die Verteilungsfunktion einer diskreten Zufallsvariablen.

$$F_X(x) = \begin{cases} 0 & \text{für } x < 2 \\ 0.1 & \text{für } 2 \leq x < 3 \\ 0.3 & \text{für } 3 \leq x < 5 \\ 0.7 & \text{für } 5 \leq x < 8 \\ 0.9 & \text{für } 8 \leq x < 9 \\ 1 & \text{für } 9 \leq x \end{cases}$$

Wie lautet die Wahrscheinlichkeitsfunktion?

Übungsaufgabe 2.3:

Welche der folgenden Funktionen sind Dichtefunktionen?

a) $f(x) = \begin{cases} \sin x & \text{für } 0 \leq x \leq \frac{3}{2}\pi \\ 0 & \text{sonst} \end{cases}$

b) $f(x) = \begin{cases} \frac{1}{3} & \text{für } 1 \leq x \leq 4 \\ 0 & \text{sonst} \end{cases}$

c) $f(x) = \begin{cases} 2x & \text{für } 0 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$

d) $f(x) = \begin{cases} 2e^{-2x} & \text{für } x \geq 0 \\ 0 & \text{sonst} \end{cases}$

Übungsaufgabe 2.4:

Welche der folgenden Funktionen erfüllen die Eigenschaften einer Verteilungsfunktion?

a)
$$F_X(x) = \begin{cases} 0 & \text{für } x < 0 \\ \frac{1}{2}x & \text{für } 0 \leq x \leq 2 \\ 1 & \text{für } 2 < x \end{cases}$$

b)
$$F_Y(y) = \begin{cases} 0 & \text{für } y \leq \frac{1}{2} \\ y^2 & \text{für } \frac{1}{2} < y < 1 \\ 1 & \text{für } 1 \leq y \end{cases}$$

c)
$$F_Z(z) = \begin{cases} 0 & \text{für } z < 0 \\ \frac{1}{4}z & \text{für } 0 \leq z < 1 \\ \frac{1}{4}z^2 & \text{für } 1 \leq z < 2 \\ 1 & \text{für } 2 \leq z \end{cases}$$

Übungsaufgabe 2.5:

Gegeben sind zwei Verteilungsfunktionen.

a)
$$F_X(x) = \begin{cases} 0 & \text{für} \quad x < 0 \\ x^2 & \text{für } 0 \leq x < 1 \\ 1 & \text{für } 1 \leq x \end{cases}$$

b)
$$F_X(x) = \begin{cases} 0 & \text{für} \quad x < 0 \\ x^4 & \text{für } 0 \leq x < 1 \\ 1 & \text{für } 1 \leq x \end{cases}$$

Bestimmen Sie:

$\boldsymbol{P}(-1 \leq X \leq \frac{1}{2})=$

$\boldsymbol{P}(\frac{1}{4} \leq X \leq \frac{3}{4}) =$

Übungsaufgabe 2.6:

Gegeben sind zwei Dichtefunktionen.

a)
$$f_X(x) = \begin{cases} \frac{1}{3} & \text{für } 1 \leq x \leq 4 \\ 0 & \text{sonst} \end{cases}$$

b)
$$f_X(x) = \begin{cases} 3\mathrm{e}^{-3x} & \text{für } x \geq 0 \\ 0 & \text{sonst} \end{cases}$$

Bestimmen Sie die zugehörigen Verteilungsfunktionen und berechnen Sie $\boldsymbol{P}(0 \leq X \leq 2)$ *und* $\boldsymbol{P}(1 \leq X \leq 2)$.

Übungsaufgabe 2.7:

Eine diskrete Zufallsvariable bestitzt folgende Wahrscheinlichkeitsverteilung:

x_i	2	3	5	8	9
$f_X(x_i)$	0.1	0.4	0.2	0.1	0.2

Berechnen Sie den Erwartungswert der Zufallsvariablen X.

$$E(X) =$$

Übungsaufgabe 2.8:

Gegeben sind zwei Dichtefunktionen.

a) $$f_X(x) = \begin{cases} 2x & \text{für } 0 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$$

a) $$f_X(x) = \begin{cases} 4x^3 & \text{für } 0 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$$

Bestimmen Sie $E(X)$.

Übungsaufgabe 2.9:

Zwei Spieler A und B würfeln abwechselnd mit je zwei Würfeln. Dabei ist folgende Regel vereinbart:

A gewinnt 3 €, wenn B zehn und mehr Augen würfelt, B gewinnt 1 €, wenn A weniger als zehn Augen würfelt.

Untersuchen Sie, ob diese Spielregeln beiden Spielern gleiche Gewinnchancen gewährt.

Übungsaufgabe 2.10:

Berechnen Sie die Varianz und Standardabweichung für die Zufallsvariable aus Aufgabe 2.7.

$$\mathrm{Var}(X) =$$

$$\sigma_X =$$

Übungsaufgabe 2.11:

Berechnen Sie die Varianz für die in Aufgabe 2.8 gegebenen stetigen Zufallsvariablen.

Übungsaufgabe 2.12:

Betrachtet wird die stochastische (zufällig schwankende) Nachfrage eines Gutes. Die Höhe der Nachfrage X sei folgendermaßen verteilt:

$$f_X(x) = \begin{cases} \frac{1}{12} & \text{für } 0 \leq x \leq 12 \\ 0 & \text{sonst.} \end{cases}$$

Die Produktion des Gutes wird unmittelbar abgesetzt, d.h., es existieren keine Absatzlager. Die Kostenfunktion des Gutes lautet $K = 2X+10$.

Geben Sie den Erwartungswert der Kosten $E(K)$ an und ermitteln Sie $\mathrm{Var}(K)$.

Übungsaufgabe 2.13:

An einer Kreuzung münden 4 Straßen. Die Anzahl der innerhalb einer halben Stunde aus den Straßen auf die Kreuzung fahrenden Kraftfahrzeuge sind voneinander unabhängige Zufallsvariablen X_1, X_2, X_3 und X_4. Es gilt $E(X_1) = 18, E(X_2) = 12, E(X_3) = 5, E(X_4) = 28, \mathrm{Var}(X_1) = 6, \mathrm{Var}(X_2) = 4, \mathrm{Var}(X_3) = 2, \mathrm{Var}(X_4) = 15$. *Bestimmen Sie Erwartungswert und Varianz für die Gesamtzahl der die Kreuzung innerhalb einer halben Stunde passierenden Kraftfahrzeuge.*

Übungsaufgabe 2.14:

Ein Röhrenwerk produziert Stahlröhren, deren Durchmesser produktionsbedingten Schwankungen unterliegen. Für den Innendurchmesser X_1 gilt $E(X_1) = 800$ *mm und* $\mathrm{Var}(X_1) = 0.01$ *und für den Außendurchmesser X_2 gilt* $E(X_2) = 810$ *mm und* $\mathrm{Var}(X_2) = 0.02$. *Bestimmen Sie Erwartungswert und Varianz für die Wandstärke der Röhren, wenn angenommen werden kann, dass Innen- und Außendurchmesser voneinander unabhängig schwanken.*

Übungsaufgabe 2.15:

Gegeben ist folgende Häufigkeitstabelle der Zufallsvariablen X *und* Y:

	y_j		
x_i	1	2	3
1	0.2	0.2	0
2	0	0.2	0.4

a) *Bestimmen Sie den Erwartungswert und die Varianz der Zufallsvariablen* X *und* Y.

b) *Bestimmen Sie die bedingten Verteilungen* $f_Y(y_j|x_i)$ *und* $f_X(x_i|y_j)$.

c) *Bestimmen Sie die Kovarianz von* X *und* Y.

d) *Bestimmen Sie die Varianz von* $2X + 3Y$ *bzw.* $2X - 3Y + 5$.

Aufgaben zu Kapitel 3

Übungsaufgabe 3.1:

Eine diskrete Zufallsvariable X nimmt die Werte 2, 4 und 6 mit jeweils gleicher Wahrscheinlichkeit an. Geben Sie die zugehörige Verteilungsfunktion an und berechnen Sie Erwartungswert und Varianz von X.

Übungsaufgabe 3.2:

Eine Zufallsvariable X ist gleichverteilt über $[0; 2]$.

a) *Bestimmen Sie den Erwartungswert und die Varianz von X.*

b) *Bestimmen Sie den Erwartungswert und die Varianz von X^2.*

Übungsaufgabe 3.3:

Eine Münze wird viermal geworfen. Wie groß ist die Wahrscheinlichkeit dafür, dass dabei

a) *zweimal Zahl (Z) oben liegt?*

b) *dreimal Zahl (Z) oben liegt?*

Übungsaufgabe 3.4:

Aus einer Produktionsserie, von der ein Anteil π leichte Fehler enthält, werden zufällig 4 Artikel entnommen. Wie groß ist die Wahrscheinlichkeit darunter 0, 1, 2, 3 oder 4 fehlerhafte Stücke zu finden, wenn a) $\pi = 0.2$, b) $\pi = 0.5$, c) $\pi = 0.6$ beträgt?

Wie groß sind für Fall a) die Wahrscheinlichkeiten, dass unter den 4 zufällig entnommenen Stücken 0, 1, 2, 3, 4 einwandfreie Stücke zu finden sind?

Übungsaufgabe 3.5:

Die Zufallsvariable Z ist $N(0,1)$-verteilt, d.h. $Z \sim N(0,1)$. Bestimmen Sie folgende Wahrscheinlichkeiten.

a) $\boldsymbol{P}(0 \leq Z \leq 2.4)$

b) $\boldsymbol{P}(-1.3 \leq Z \leq 0)$

c) $\boldsymbol{P}(-0.8 \leq Z \leq 0.8)$

d) $\boldsymbol{P}(Z \leq 2.1)$

e) $\boldsymbol{P}(Z \leq -0.4)$

f) $\boldsymbol{P}(Z \leq -0.1)$

g) $\boldsymbol{P}(0.2 \leq Z \leq 1.6)$

h) $\boldsymbol{P}(-1.4 \leq Z \leq 1.2)$

i) $\boldsymbol{P}(-2 \leq Z \leq -1)$

Übungsaufgabe 3.6:

Die Zufallsvariable Z ist $N(0,1)$-verteilt. Bestimmen Sie A, B, C, D.

a) $\boldsymbol{P}(Z \leq A) = 0.6$ A=

b) $\boldsymbol{P}(Z \leq B) = 0.8$ B=

c) $\boldsymbol{P}(|Z| \leq C) = 0.6$ C=

d) $\boldsymbol{P}(|Z| \geq D) = 0.3$ D=

Übungsaufgabe 3.7:

Die Brenndauer von Glühbirnen sei normalverteilt mit einem Erwartungswert von 900 *Stunden und einer Varianz von* 10000 *Stunden.*
Bestimmen Sie die Wahrscheinlichkeit für eine Brenndauer

a) zwischen 750 *und* 1050 *Stunden,*

b) zwischen 800 *und* 1050 *Stunden,*

c) kleiner als 650 *Stunden,*

d) größer als 1200 *Stunden,*

e) kleiner als 800 *oder größer als* 1200 *Stunden.*

Übungsaufgabe 3.8:

Es ist $X \sim N(100, 100)$. *Bestimmen Sie* A, B *und* C.

a) $\boldsymbol{P}(X \leq A) = 0.7$ A=

b) $\boldsymbol{P}(X \geq B) = 0.65$ B=

c) $\boldsymbol{P}(|X - 100| \leq C) = 0.5$ C=

Übungsaufgabe 3.9:

Die Kapazität einer Lieferung von Kondensatoren sei normalverteilt mit Erwartungswert $\mu = 100\,(pF)$ *und einer Standardabweichung von* $\sigma = 0.2$. *Wie viel Prozent Ausschuss sind zu erwarten, wenn die Kapazität der Kondensatoren*

a) *mindestens* 99.8 *pF,*

b) *höchstens* 100.6 *pF betragen soll,*

c) *um maximal* 0.3 *pF vom Sollwert* 100 *pF abweichen darf?*

d) *Wie müssen die Toleranzgrenzen* $100 + C$ *und* $100 - C$ *gewählt werden, um genau* 5% *Ausschuss zu erhalten?*

e) *Wie ändert sich der Ausschussprozentsatz für die in Frage d) bestimmten Toleranzgrenzen, wenn sich* μ *nach* 100.1 *verschiebt?*

Übungsaufgabe 3.10:

Von einem Betrieb werden Metallfolien hergestellt, von denen nur Folien, deren Durchmesser zwischen 0.82 *mm und* 1.18 *mm liegen, zur Weiterverarbeitung verwendet werden können. (Der Rest wird als Ausschuss betrachtet.) Zur Herstellung werden dem Betrieb zwei Maschinen* A *und* B *angeboten. Die Foliendurchmesser der mit diesen Maschinen hergestellten Folien sind um den auf den Maschinen einstellbaren Sollwert (Erwartungswert) normalverteilt, und zwar bei Maschine* A *mit einer Standardabweichung von* 0.1 *mm und bei* B *von* 0.18 *mm.*

a) Wie sollte der Sollwert eingestellt werden, um den Ausschussanteil zu minimieren?

b) Die Produktionskosten pro 1000 *Folien betragen für Maschine A* 20 € *und für B* 16 €*. Für welche der beiden Maschinen sollte sich der Betrieb entscheiden, wenn einwandfreie Folien zu minimalen Stückkosten hergestellt werden sollen?*

Übungsaufgabe 3.11:

Es sind $X_1 \sim N(12, 16)$ *und* $X_2 \sim N(18, 9)$. *Bestimmen Sie die Wahrscheinlichkeiten in a) und b).*

$$a) \qquad \boldsymbol{P}(X_1 + X_2 \leq 21)$$

$$b) \qquad \boldsymbol{P}(24 \leq X_1 + X_2 \leq 42)$$

Übungsaufgabe 3.12:

X *sei* χ^2*-verteilt mit* $n = 25$ *Freiheitsgraden. Bestimmen Sie die Quantile* x_1, x_2, x_3 *und* x_4.

a) $F_X(x_1) = 0.95$

b) $F_X(x_2) = 0.1$

c) $1 - F_X(x_3) = 0.1$

d) $1 - F_X(x_4) = 0.01$

Übungsaufgabe 3.13:

Y sei χ^2-verteilt mit $n = 18$ Freiheitsgraden. Bestimmen Sie

a) $P(8.23 \leq Y \leq 37.16)$

b) $P(6.26 \leq Y \leq 31.53)$

Übungsaufgabe 3.14:

T ist Student-verteilt mit $n = 12$ Freiheitsgraden. Bestimmen Sie t_1, t_2, t_3 und t_4.

a) $\boldsymbol{P}(-t_1 \leq T \leq t_1) = 0.9$

b) $\boldsymbol{P}(T \leq t_2) = 0.99$

c) $\boldsymbol{P}(T \geq t_3) = 0.9$

d) $\boldsymbol{P}(T \geq t_4) = 0.1$

Übungsaufgabe 3.15:

Gegeben sind die stochastisch unabhängig standardnormalverteilten Zufallsvariablen $X_1, ..., X_6, Y_1, ..., Y_4$. Bestimmen Sie das 0.95, 0.99, 0.05 und 0.01 Quantil der Verteilung der Zufallsvariablen W.

$$W = \frac{2 \sum\limits_{i=1}^{6} X_i^2}{3 \sum\limits_{j=1}^{4} Y_j^2}$$

Übungsaufgabe 3.16:

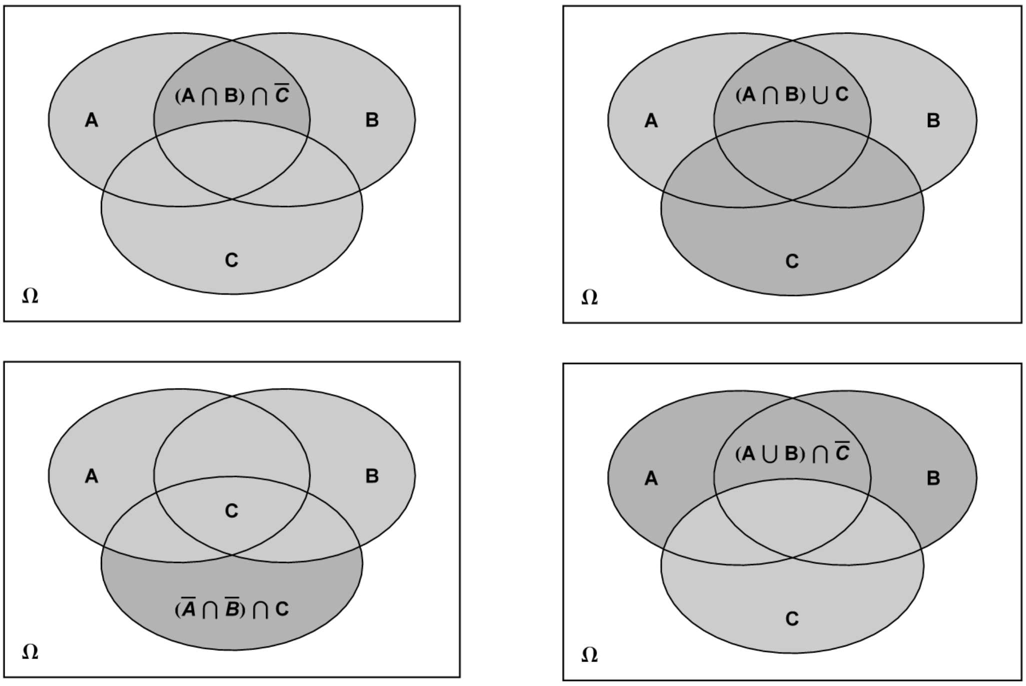
Gegeben sind die stochastisch unabhängig standardnormalverteilten Zufallsvariablen $X_1, ..., X_3, Y_1, ..., Y_5$. Bestimmen Sie den Erwartungswert und die Varianz der Zufallsvariablen W.

$$W = \frac{20 \sum\limits_{i=1}^{3} X_i^2}{6 \sum\limits_{j=1}^{5} Y_j^2}$$

Lösung der Übungsaufgaben

Lösung der Aufgaben zu Kapitel 1

Lösung 1.1:



Lösung 1.2:

a) $\Omega = \{(Kopf, Kopf), (Kopf, Zahl), (Zahl, Zahl)\}$

b) $\Omega = \{(rot, rot), (rot, grün), (rot, blau), (grün, grün),$
$(grün, blau), (blau, blau)\}$

c) $\Omega = \{(rot, rot), (rot, grün), (rot, blau), (grün, rot), (grün, grün),$
$(grün, blau), (blau, rot), (blau, grün), (blau, blau)\}$

Lösung 1.3:

(Z, Z, K, K, K) (Z, K, Z, K, K) (Z, K, K, Z, K) (Z, K, K, K, Z)
(K, Z, Z, K, K) (K, Z, K, Z, K) (K, Z, K, K, Z) (K, K, Z, Z, K)
(K, K, Z, K, Z) (K, K, K, Z, Z)

Lösung 1.4:

$A \cap C = \emptyset \quad A \cap D = \emptyset \quad B \cap D = \emptyset \quad C \cap D = \emptyset$

Lösung 1.5:

a) $\boldsymbol{P}(grün) = \frac{3}{9} = \frac{1}{3}$

b) $(K,K,K)\ \ (K,K,Z)\ \ (K,Z,K)\ \ (Z,K,K)\ \ (K,Z,Z)$
$(Z,K,Z)\ \ (Z,Z,K)\ \ (Z,Z,Z)$

$\boldsymbol{P}(zweimal\ \ Z) = \frac{3}{8}$

Lösung 1.6:

Lösung a) ist richtig.

Lösung b) ist falsch, da von insgesamt 28 *Kugeln* 7 *grün sind.*
Die richtige Wahrscheinlichkeit im Fall b) ist $\frac{1}{4}$.

Lösung 1.7:

a) *Ziehen mit Zurücklegen, mit Berücksichtigung der Reihenfolge:*

$$5^3 = 125 \ \textit{Möglichkeiten}$$

b) *Ziehen ohne Zurücklegen, ohne Berücksichtigung der Reihenfolge:*

$$\binom{5}{3} = 10 \ \textit{Möglichkeiten}$$

Lösung 1.8:

$$\frac{9!}{3!2!4!} = 1260$$

Lösung 1.9:

Da die Steine $\{1,3\}$ und $\{3,1\}$ gleich sind und Wiederholungen zugelassen sind, z.B. $\{1,1\}$, ist nach den Anordnungen einer 2-er Kombination aus einer 7-elementigen Reihe mit Wiederholung, ohne Berücksichtigung der Reihenfolge gefragt.

$$\binom{7+2-1}{2} = \binom{8}{2} = \frac{8!}{2!6!} = \frac{8 \cdot 7}{1 \cdot 2} = 28$$

Lösung 1.10:

$\boldsymbol{P}(I) = \frac{1}{1000} = \frac{1}{100} = 0.01$

$\boldsymbol{P}(II) = \frac{80}{1000} = \frac{8}{100} = 0.08$

$\boldsymbol{P}(I \cup II) = \boldsymbol{P}(I) + \boldsymbol{P}(II) = 0.01 + 0.08 = 0.09$

Lösung 1.11:

a) $\boldsymbol{P}(B \cup *) = \boldsymbol{P}(B) + \boldsymbol{P}(*) - \boldsymbol{P}(B \cap *) = \frac{70}{200} + \frac{50}{200} - \frac{20}{200} = \frac{1}{2}$

b) $\boldsymbol{P}(\text{„Kreuz“ oder „AS“}) = \boldsymbol{P}(\text{„Kreuz“}) + \boldsymbol{P}(\text{„As“}) - \boldsymbol{P}(\text{„Kreuz As“})$

$$= 0.25 + 0.125 + 0.03125 = 0.34375$$

Lösung 1.12:

(1) *Da* $\boldsymbol{P}(C) = \frac{3}{2} > 1$ *dem Axiom* (3) *aus Abschnitt* **??** *widerspricht, liegt keine Wahrscheinlichkeit vor.*

(2) *Es wird keine Wahrscheinlichkeit auf* $\mathcal{P}(\Omega)$ *definiert, denn es liegt ein Widerspruch vor. Allgemein gilt nämlich:*

$$\begin{aligned}
\boldsymbol{P}(A \cup B \cup C) &= \boldsymbol{P}((A \cup B) \cup C) \\
&= \boldsymbol{P}(A \cup B) + \boldsymbol{P}(C) - \boldsymbol{P}((A \cup B) \cap C) \\
&\leq \boldsymbol{P}(A \cup B) + \boldsymbol{P}(C) \\
&= \boldsymbol{P}(A) + \boldsymbol{P}(B) - \boldsymbol{P}(A \cap B) + \boldsymbol{P}(C) \\
&\leq \boldsymbol{P}(A) + \boldsymbol{P}(B) + \boldsymbol{P}(C).
\end{aligned}$$

Hieraus leitet sich folgender Widerspruch ab:
$1 = \boldsymbol{P}(\Omega) = \boldsymbol{P}(A \cup B \cup C) \geq \boldsymbol{P}(A) + \boldsymbol{P}(B) + \boldsymbol{P}(C) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8}$.

(3) *Es wird eine Wahrscheinlichkeit auf* $\mathcal{P}(\Omega)$ *definiert.*

(4) *Es wird eine Wahrscheinlichkeit auf* $\mathcal{P}(\Omega)$ *definiert.*

(5) *Es liegt keine Wahrscheinlichkeit vor, denn* $\boldsymbol{P}(B) = -\frac{1}{4} < 0$ *widerspricht Axiom* (1) *aus Abschnitt* **??**.

(6) *Es wird keine Wahrscheinlichkeit auf* $\mathcal{P}(\Omega)$ *definiert, denn es liegt ein Widerspruch vor. Allgemein gilt nämlich:*

$$\begin{aligned} \boldsymbol{P}(A \cup B) &= \boldsymbol{P}(A \cup (B \cap \overline{A})) = \boldsymbol{P}(A) + \boldsymbol{P}(B \cap \overline{A}) \\ &\geq \boldsymbol{P}(A), \text{ denn } A \cap (B \cap \overline{A}) = \emptyset. \end{aligned}$$

Hieraus leitet sich folgender Widerspruch ab:
$\frac{1}{4} = \boldsymbol{P}(A \cup B) \leq \boldsymbol{P}(A) = \frac{1}{2}$.

Lösung 1.13:

A = {*Augensumme ist größer als* 11 *bei* 10 *Würfen*}
$\overline{A}$ = {*Augensumme ist kleiner oder gleich* 11 *bei* 10 *Würfen*}

$\boldsymbol{P}(A) + \boldsymbol{P}(\overline{A}) = 1$, *da* $A \cup \overline{A} = \Omega$

$\overline{A} = \{(1111111111), (2111111111), (1211111111), (...),$
$(...), (1111111112)\}$ *(insg. 11 Fälle)*

$\boldsymbol{P}(\overline{A}) = \frac{11}{6^{10}}$ $\quad \boldsymbol{P}(A) = 1 - \boldsymbol{P}(\overline{A}) = 1 - \frac{11}{6^{10}} = 0.999999818$

Lösung 1.14:

a) $\boldsymbol{P}(G_2|B_1) = \frac{1}{2}$ $\quad \boldsymbol{P}(G_2|G_1) = \frac{5}{12}$

b) $\boldsymbol{P}(A|B) = \dfrac{\boldsymbol{P}(A \cap B)}{\boldsymbol{P}(B)}$

$A \cap B = \{(5,4), (6,4)\}$ $\quad \boldsymbol{P}(A \cap B) = \frac{2}{36}$ $\quad \boldsymbol{P}(B) = \frac{1}{6}$

$\boldsymbol{P}(A|B) = \dfrac{\frac{2}{36}}{\frac{1}{6}} = \frac{1}{3}$

Lösung 1.15:

a) *Hier können alle möglichen Fälle der Form* $(1.Kind, 2.Kind)$ *vorliegen, d.h.* (J, J),(J, M),(M, J),(M, M). *Somit ergibt sich für die Wahrscheinlichkeit nach Laplace* $\frac{1}{4}$.

b) *Aufgrund der Vorinformation kommen hier nur die Fälle* (J, J),(J, M),(M, J) *in Betracht, so dass die gesuchte Wahrscheinlichkeit* $\frac{1}{3}$ *beträgt.*

c) *Möglich sind hier nur die Fälle* (J, J) *und* (J, M). *Für die gesuchte Wahrscheinlichkeit gilt* $\frac{1}{2}$.

Lösung 1.16:

$$\begin{aligned} S &= \{\text{Bestehen der Klausur Statistik}\} \\ M &= \{\text{Bestehen der Klausur Mathematik}\} \end{aligned}$$

Die Ereignisse S und M sind unabhängig, wenn gilt:

$$\boldsymbol{P}(S|M) = \boldsymbol{P}(S) \quad \text{bzw.} \quad \boldsymbol{P}(M|S) = \boldsymbol{P}(M)$$

a) $\boldsymbol{P}(S) = 0.7 \quad \boldsymbol{P}(M) = 0.8 \quad \boldsymbol{P}(S \cap M) = 0.6$ *Es gilt:*

$$\boldsymbol{P}(M|S) = \frac{\boldsymbol{P}(S \cap M)}{\boldsymbol{P}(S)} = \frac{0.6}{0.7} = 0.857 \neq \boldsymbol{P}(M)$$

und entsprechend:

$$\boldsymbol{P}(S|M) = \frac{0.6}{0.8} = 0.75 \neq \boldsymbol{P}(S).$$

Die Ereignisse S und M sind abhängig.

b) $C = \{\text{wenigstens eine Klausur bestanden}\}$

$C = \{S \cup M\}$

$\boldsymbol{P}(C) = \boldsymbol{P}(S) + \boldsymbol{P}(M) - \boldsymbol{P}(S \cap M) = 0.7 + 0.8 - 0.6 = 0.9$

Lösung 1.17:

$A = \{Zeitung\ 1\ wird\ gelesen\}, B = \{Zeitung\ 2\ wird\ gelesen\}$

Gegeben: $\boldsymbol{P}(A) = 0.5 \quad \boldsymbol{P}(A \cap \overline{B}) = 0.15 \quad \boldsymbol{P}(B \cap \overline{A}) = 0.2$

Vierfeldertafel:

	A	$\overline{A}$	$\sum$
B	0.35	0.2*	0.55
$\overline{B}$	0.15*	0.3	0.45
$\sum$	0.5*	0.5	1.0*

* *gegebene Wahrscheinlichkeiten*

a)

$$\begin{aligned} C &= \{Erwachsener\ liest\ wenigstens\ eine\ Zeitung\} \\ &= \{A \cup B\} \\ \boldsymbol{P}(C) &= \boldsymbol{P}(A \cup B) = \boldsymbol{P}(A) + \boldsymbol{P}(B) - \boldsymbol{P}(A \cap B) \\ &= 0.5 + 0.55 - 0.35 = 0.7 \end{aligned}$$

b)

$$\begin{aligned} D &= \{Erwachsener\ liest\ beide\ Zeitungen\} = \{A \cap B\} \\ \boldsymbol{P}(D) &= \{A \cap B\} = 0.35 \end{aligned}$$

c)

$$\begin{aligned} E &= \{Erwachsener\ liest\ höchstens\ eine\ Zeitung\} \\ \overline{E} &= \{Erwachsener\ liest\ zwei\ Zeitungen\} = \{A \cap B\} \\ \boldsymbol{P}(\overline{E}) &= \boldsymbol{P}(D) = 0.35 \quad \boldsymbol{P}(E) = 1 - \boldsymbol{P}(\overline{E}) = 0.65 \end{aligned}$$

d)

$$\begin{aligned} F &= \{Erwachsener\ liest\ keine\ Zeitung\} = \{\overline{A} \cap \overline{B}\} \\ \boldsymbol{P}(F) &= \boldsymbol{P}(\overline{A} \cap \overline{B}) = 0.3 \end{aligned}$$

Lösung 1.18:

a) $\boldsymbol{P}(12) = \boldsymbol{P}(6) \cdot \boldsymbol{P}(6) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$

b) $\boldsymbol{P}(10) = \boldsymbol{P}((6,4) \cup (5,5) \cup (4,6)) = \boldsymbol{P}(6,4) + \boldsymbol{P}(5,5) + \boldsymbol{P}(4,6)$

Da sich die Ereignisse gegenseitig ausschließen, gilt dann:

$$\boldsymbol{P}(10) = \boldsymbol{P}(6) \cdot \boldsymbol{P}(4) + \boldsymbol{P}(5) \cdot \boldsymbol{P}(5) + \boldsymbol{P}(4) \cdot \boldsymbol{P}(6)$$

$$= \tfrac{1}{6} \cdot \tfrac{1}{6} + \tfrac{1}{6} \cdot \tfrac{1}{6} + \tfrac{1}{6} \cdot \tfrac{1}{6} = \tfrac{1}{12}$$

Lösung 1.19:

$A = \{\textit{Spieler schießt ein Tor}\}$ $\boldsymbol{P}(A) = 0.5$
$\overline{A} = \{\textit{Spieler schießt kein Tor}\}$
$A \cup \overline{A} = \Omega_1 \Rightarrow \boldsymbol{P}(\overline{A}) = 1 - \boldsymbol{P}(A) = 0.5$
$A_n = \{\textit{Spieler schießt nach n Schüssen mindestens ein Tor}\}$
$\overline{A}_n = \{\textit{Spieler schießt nach n Schüssen kein Tor}\}$
$A \cup \overline{A} = \Omega_n \Rightarrow \boldsymbol{P}(A_n) = 1 - \boldsymbol{P}(\overline{A}_n) = 1$

$n = 1 \Rightarrow \overline{A}_1 = \{\overline{A}\} \Rightarrow \boldsymbol{P}(\overline{A}_1) = 0.5$
$n = 2 \Rightarrow \overline{A}_2 = \{\overline{A}\,\overline{A}\} \Rightarrow \boldsymbol{P}(\overline{A}_2) = 0.5 \cdot 0.5$*(Schüsse unabhängig)*
.
.
.
$n = n^* \Rightarrow \overline{A}_n = \{\underbrace{\overline{A}, \overline{A}, ..., \overline{A}}_{n^*\ mal}\} \Rightarrow \boldsymbol{P}(\overline{A}_{n^*}) = 0.5^{n^*}$

gegeben:

$\boldsymbol{P}(A_n) = 0.99$
$\boldsymbol{P}(\overline{A}_n) = 0.01$
$\boldsymbol{P}(\overline{A}_n) = 0.01 = 0.5^n$

$$n \cdot \log 0.5 = \log 0.01$$

$$n = \frac{\log 0.01}{\log 0.5} = \frac{0 - 2}{0.7 - 1} = \frac{-2}{-0.3} = 6.67$$

Der Spieler muss mindestens siebenmal aufs Tor schießen.

Lösung 1.20:

$A = \{Maschine\ fällt\ aus\}$
$\overline{A} = \{Maschine\ fällt\ nicht\ aus\}$

$\boldsymbol{P}(A) + \boldsymbol{P}(\overline{A}) = 1$

$B = \{Aggregat\ I\ fällt\ aus\ \}$	$\boldsymbol{P}(B) = 0.3$	$\boldsymbol{P}(\overline{B}) = 0.7$
$C = \{Aggregat\ II\ fällt\ aus\ \}$	$\boldsymbol{P}(C) = 0.2$	$\boldsymbol{P}(\overline{C}) = 0.8$
$D = \{Aggregat\ III\ fällt\ aus\ \}$	$\boldsymbol{P}(D) = 0.1$	$\boldsymbol{P}(\overline{D}) = 0.9$

$\overline{A} = \overline{B} \cap \overline{C} \cap \overline{D} \Rightarrow \boldsymbol{P}(\overline{A}) = \boldsymbol{P}(\overline{B}) \cdot \boldsymbol{P}(\overline{C}) \cdot \boldsymbol{P}(\overline{D}) = 0.7 \cdot 0.8 \cdot 0.9 = 0.504$

$\boldsymbol{P}(A) = 1 - \boldsymbol{P}(\overline{A}) = 1 - 0.504 = 0.496$

Lösung 1.21:

$A = (das\ ausgewählte\ Stück\ stammt\ von\ Maschine\ A)$
$B = (das\ ausgewählte\ Stück\ stammt\ von\ Maschine\ B)$
$C = (das\ ausgewählte\ Stück\ stammt\ von\ Maschine\ C)$
$X = (das\ ausgewählte\ Stück\ ist\ defekt)$

A, B und C schließen sich paarweise aus

Gegeben:

$\boldsymbol{P}(A) = 0.5$	$\boldsymbol{P}(B) = 0.3$	$\boldsymbol{P}(C) = 0.2$
$\boldsymbol{P}(X\|A) = 0.03$	$\boldsymbol{P}(X\|B) = 0.04$	$\boldsymbol{P}(X\|C) = 0.05$

a)

$$\begin{aligned} \boldsymbol{P}(X) &= \boldsymbol{P}(X|A) \cdot \boldsymbol{P}(A) + \boldsymbol{P}(X|B) \cdot \boldsymbol{P}(B) + \boldsymbol{P}(X|C) \cdot \boldsymbol{P}(C) \\ &= 0.03 \cdot 0.5 + 0.04 \cdot 0.3 + 0.05 \cdot 0.2 = 0.037 \end{aligned}$$

b)

$$\begin{aligned} \boldsymbol{P}(A|X) &= \frac{\boldsymbol{P}(X|A) \cdot \boldsymbol{P}(A)}{\boldsymbol{P}(X|A) \cdot \boldsymbol{P}(A) + \boldsymbol{P}(X|B) \cdot \boldsymbol{P}(B) + \boldsymbol{P}(X|C) \cdot \boldsymbol{P}(C)} \\ &= \frac{0.03 \cdot 0.5}{0.037} = 0.405 \end{aligned}$$

Lösung 1.22:

$$\boldsymbol{P}(I) = 0.001 \qquad \boldsymbol{P}(T|I) = 0.99 \qquad \boldsymbol{P}(T|\overline{I}) = 1 - \boldsymbol{P}(\overline{T}|\overline{I}) = 0.02$$

a)

$$\begin{aligned}\boldsymbol{P}(I|T) &= \frac{\boldsymbol{P}(T|I)\cdot\boldsymbol{P}(I)}{\boldsymbol{P}(T|I)\cdot\boldsymbol{P}(I) + \boldsymbol{P}(T|\overline{I})\cdot\boldsymbol{P}(\overline{I})} \\ &= \frac{0.99\cdot 0.001}{0.99\cdot 0.001 + 0.02\cdot 0.999} = 0.05\end{aligned}$$

b)

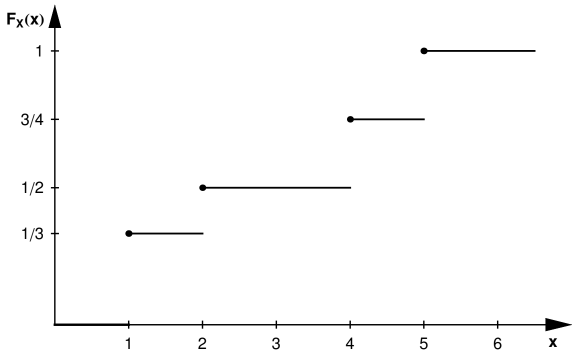
$$\begin{aligned}\boldsymbol{P}(I|\overline{T}) &= \frac{\boldsymbol{P}(\overline{T}|I)\cdot\boldsymbol{P}(I)}{\boldsymbol{P}(\overline{T}|I)\cdot\boldsymbol{P}(I) + \boldsymbol{P}(\overline{T}|\overline{I})\cdot\boldsymbol{P}(\overline{I})} \\ &= \frac{0.01\cdot 0.001}{0.01\cdot 0.001 + 0.98\cdot 0.999} = 0.00001\end{aligned}$$

Tatsächlich sind nur 5% der Patienten infiziert, wenn ein positiver Test vorliegt und die Wahrscheinlichkeit einen infizierten Patienten nicht zu erkennen ist gering. Dieser Zusammenhang lässt sich ähnlich wie in dem Beispiel **??***.17 anhand einer Vierfeldertafel gut erklären. Da in der Bevölkerung viel mehr gesunde Personen sind, die fälschlicherweise als infiziert eingestuft werden, so ist dementsprechend auch die Wahrscheinlichkeit gering, dass bei einem positiven Testergebnis die Person tatsächlich infiziert ist.*

Lösung der Aufgaben zu Kapitel 2

Lösung 2.1:

$$F_X(x) = \begin{cases} 0 & \text{für} \quad x < 1 \\ \frac{1}{3} & \text{für } 1 \leq x < 2 \\ \frac{1}{2} & \text{für } 2 \leq x < 4 \\ \frac{3}{4} & \text{für } 4 \leq x < 5 \\ 1 & \text{für } 5 \leq x \end{cases}$$



Für eine diskrete Wahrscheinlichkeitsverteilung können Wahrscheinlichkeiten für Intervalle der Zufallsvariablen durch Addition der Wahrscheinlichkeiten für die bestehenden diskreten Werte der Zufallsvariablen oder über die Verteilungsfunktion bestimmt werden.

Unter Verwendung der Verteilungsfunktion ist im diskreten Fall auf die Intervallgrenzen (<oder $\leq$) zu achten, denn es gilt $\boldsymbol{P}(a < x \leq b) = F_X(b) - F_X(a)$. Bei Abweichungen von diesen Intervallgrenzen sind gegebenenfalls die Wahrscheinlichkeiten für $x = a$ bzw. $x = b$ zu addieren bzw. zu subtrahieren. Es werden beide Lösungswege behandelt.

$$\begin{aligned} \boldsymbol{P}(0 < x < 4) &= \boldsymbol{P}(x = 1) + \boldsymbol{P}(x = 2) = \tfrac{1}{3} + \tfrac{1}{6} = \tfrac{1}{2} \\ \boldsymbol{P}(0 < x < 4) &= F_X(4) - F_X(0) - \boldsymbol{P}(x = 4) = \tfrac{3}{4} - 0 - \tfrac{1}{4} = \tfrac{1}{2} \\ \boldsymbol{P}(1 \leq x \leq 4) &= \boldsymbol{P}(x = 1) + \boldsymbol{P}(x = 2) + \boldsymbol{P}(x = 4) = \tfrac{1}{3} + \tfrac{1}{6} + \tfrac{1}{4} = \tfrac{3}{4} \\ \boldsymbol{P}(1 \leq x \leq 4) &= F_X(4) - F_X(1) + \boldsymbol{P}(x = 1) = \tfrac{3}{4} - \tfrac{1}{3} + \tfrac{1}{3} = \tfrac{3}{4} \end{aligned}$$

$$\begin{aligned}
\boldsymbol{P}(1 < x < 4) &= \boldsymbol{P}(x = 2) = \tfrac{1}{6} \\
\boldsymbol{P}(1 < x < 4) &= F_X(4) - F_X(1) - \boldsymbol{P}(x = 4) = \tfrac{3}{4} - \tfrac{1}{3} - \tfrac{1}{4} = \tfrac{1}{6} \\
\boldsymbol{P}(2 < x \leq 5) &= \boldsymbol{P}(x = 4) + \boldsymbol{P}(x = 5) = \tfrac{1}{4} + \tfrac{1}{4} = \tfrac{1}{2} \\
\boldsymbol{P}(2 < x \leq 5) &= F_X(5) - F_X(2) = 1 - \tfrac{1}{2} = \tfrac{1}{2}
\end{aligned}$$

Lösung 2.2:

x_i	2	3	5	8	9
$f_X(x_i)$	0.1	0.2	0.4	0.2	0.1

Lösung 2.3:

Eine Dichtefunktion muss die Bedingungen

$$\begin{aligned}
&(1) \quad \int_{-\infty}^{\infty} f_X(x)dx = 1 \quad \textit{und} \\
&(2) \quad f_X(x) \geq 0 \quad \textit{für alle } x \in \mathbb{R}
\end{aligned}$$

erfüllen. Diese beiden Bedingungen sind für die angegebenen Funktionen zu überprüfen:

a) *Keine Dichtefunktion, da* $f(x) = \sin x < 0$ *für* $\pi < x \leq \frac{3}{2}\pi$.

b) (1) $f(x) \geq 0$

(2) $\int_{-\infty}^{\infty} f(x)dx = \int_1^4 \frac{1}{3}dx = \left[\frac{1}{3}x\right]_1^4 = \frac{4}{3} - \frac{1}{3} = 1$

Es handelt sich um eine Dichtefunktion.

c) (1) $f(x) \geq 0$

(2) $\int_{-\infty}^{\infty} f(x)dx = \int_0^1 2xdx = \left[x^2\right]_0^1 = 1 - 0 = 1$

Es handelt sich um eine Dichtefunktion.

d) (1) $f(x) \geq 0$

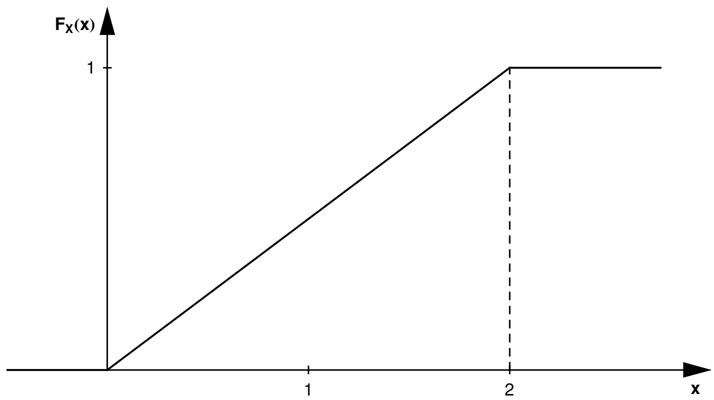
(2) $\int_{-\infty}^{\infty} f(x)dx = \int_{0}^{\infty} 2 \cdot e^{-2x}dx = \left[-e^{-2x}\right]_0^{\infty} = 0 - (-1) = 1$

Es handelt sich um eine Dichtefunktion.

Lösung 2.4:

Für die gegebenen Funktionen ist zu prüfen, ob sie die Eigenschaften einer Verteilungsfunktion besitzen (vgl. Abschnitt 2.2).

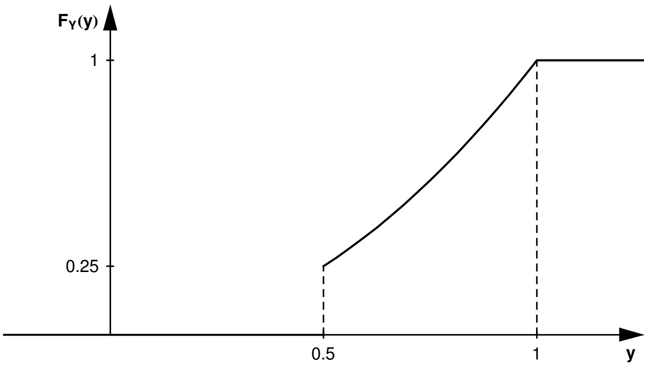
$$a) \qquad F_X(x) = \begin{cases} 0 & \text{für} \quad x < 0 \\ \frac{1}{2}x & \text{für } 0 \leq x \leq 2 \\ 1 & \text{für } 2 < x \end{cases}$$



1. $\lim\limits_{x\to-\infty} F_X = 0$
2. $\lim\limits_{x\to\infty} F_X = 1$
3. F_X *monoton steigend*
4. *rechtsseitig stetig*

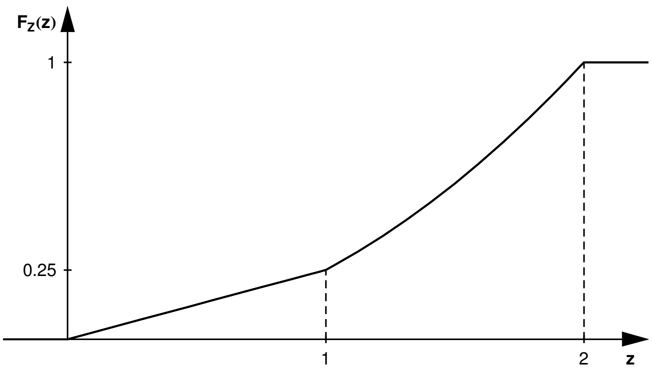
$F_X(x)$ *ist eine Verteilungsfunktion.*

b) $$F_Y(y) = \begin{cases} 0 & \text{für} \quad y \leq \frac{1}{2} \\ y^2 & \text{für } \frac{1}{2} < y < 1 \\ 1 & \text{für } 1 \leq y \end{cases}$$



$F_Y(y)$ ist keine Verteilungsfunktion. Bedingung 4 ist nicht erfüllt.
$F_Y(y)$ ist für $y = \frac{1}{2}$ linksseitig stetig.

c) $$F_Z(z) = \begin{cases} 0 & \text{für} \quad z < 0 \\ \frac{1}{4}z & \text{für } 0 \leq z < 1 \\ \frac{1}{4}z^2 & \text{für } 1 \leq z < 2 \\ 1 & \text{für } 2 \leq z \end{cases}$$



Alle Eigenschaften einer Verteilungsfunktion werden erfüllt.

Lösung 2.5:

a) $\boldsymbol{P}(-1 \leq x \leq \frac{1}{2}) = F_X(\frac{1}{2}) - F_X(-1) = \frac{1}{4} - 0 = \frac{1}{4}$

$\boldsymbol{P}(\frac{1}{4} \leq x \leq \frac{3}{4}) = F_X(\frac{3}{4}) - F_X(\frac{1}{4}) = \frac{9}{16} - \frac{1}{16} = \frac{1}{2}$

b) $\boldsymbol{P}(-1 \leq x \leq \frac{1}{2}) = F_X(\frac{1}{2}) - F_X(-1) = \frac{1}{16} - 0 = \frac{1}{16}$

$\boldsymbol{P}(\frac{1}{4} \leq x \leq \frac{3}{4}) = F_X(\frac{3}{4}) - F_X(\frac{1}{4}) = \frac{81}{256} - \frac{1}{256} = \frac{5}{16}$

Lösung 2.6:

$$\begin{aligned} a) \quad F_X(x) &= \int_{-\infty}^{x} f_X(\xi) d\xi = \int_1^x \frac{1}{3} d\xi = \left[\frac{1}{3}\xi\right]_1^x \\ &= \frac{1}{3}x - \frac{1}{3} \quad \textit{für } 1 \leq x < 4 \end{aligned}$$

Es ist also:

$$F_X(x) = \begin{cases} 0 & \text{für} \quad x < 1 \\ \frac{1}{3}x - \frac{1}{3} & \text{für } 1 \leq x < 4 \\ 1 & \text{für } 4 \leq x. \end{cases}$$

$$\begin{aligned} \boldsymbol{P}(0 \leq x \leq 2) &= F_X(2) - F_X(0) = \frac{1}{3} \cdot 2 - \frac{1}{3} - 0 = \frac{1}{3} \\ \boldsymbol{P}(1 \leq x \leq 2) &= F_X(2) - F_X(1) = \frac{1}{3} \cdot (2 - 1) = \frac{1}{3} \end{aligned}$$

$$\begin{aligned} b) \quad f_X(x) &= \begin{cases} 3 \cdot e^{-3x} & \text{für } x \geq 0 \\ 0 & \text{sonst} \end{cases} \\ F_X(x) &= \int_0^x 3 \cdot e^{-3\xi} d\xi = 3 \cdot \left[-\frac{1}{3} e^{-3\xi}\right]_0^x = \left[-e^{-3\xi}\right]_0^x \\ &= -e^{-3x} + 1 = 1 - e^{-3x} \quad \textit{für } x \geq 0 \end{aligned}$$

daraus folgt:

$$F_X(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 - e^{-3x} & \text{für } 0 \leq x \end{cases}$$

$$\begin{aligned} \boldsymbol{P}(0 \leq x \leq 2) &= F_X(2) - F_X(0) = (1 - e^{-6}) - (1 - e^{0}) \\ &= (1 - 0.00248) - (1 - 1) = 0.99752 \\ \boldsymbol{P}(1 \leq x \leq 2) &= F_X(2) - F_X(1) = (1 - e^{-6}) - (1 - e^{-3}) \\ &= (1 - 0.00248) - (1 - 0.04979) = 0.04731 \end{aligned}$$

Lösung 2.7:

$$\begin{aligned} E(X) &= \sum x_i f_X(x_i) = 2 \cdot 0.1 + 3 \cdot 0.4 + 5 \cdot 0.2 + 8 \cdot 0.1 + 9 \cdot 0.2 \\ &= 0.2 + 1.2 + 1.0 + 0.8 + 1.8 = 5 \end{aligned}$$

Lösung 2.8:

$$\begin{aligned} a)\ E(X) &= \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x2x\, dx = \int_0^1 2x^2\, dx = \left[\frac{2}{3}x^3\right]_0^1 = \frac{2}{3} \\ b)\ E(X) &= \int_0^1 x4x^3\, dx = \int_0^1 4x^4\, dx = \left[\frac{4}{5}x^5\right]_0^1 = \frac{4}{5} \end{aligned}$$

Lösung 2.9:

$$\begin{aligned} &\boldsymbol{P} \quad (\textit{Augenzahl bei zwei Würfeln}\ ) \geq 10) = \tfrac{6}{36} = \tfrac{1}{6} \\ &\boldsymbol{P} \quad (\textit{Augenzahl bei zwei Würfeln}\ ) < 10) = \tfrac{5}{6} \\ &E \quad (\textit{Gewinn des Spielers A}\ ) = 3 \cdot \tfrac{1}{6} - 1 \cdot \tfrac{5}{6} = -\tfrac{1}{3} \\ &E \quad (\textit{Gewinn des Spielers B}\ ) = 1 \cdot \tfrac{5}{6} - 3 \cdot \tfrac{1}{6} = +\tfrac{1}{3} \\ &E \quad (\textit{Gewinn des Spielers B}\ ) > E(\textit{Gewinn des Spielers A}\ ) \end{aligned}$$

Es ist kein faires Spiel.

Lösung 2.10:

$$\begin{aligned}
\operatorname{Var}(X) &= \sum (x_i - E(X))^2 f_X(x_i) \\
&= (2-5)^2 \cdot 0.1 + (3-5)^2 \cdot 0.4 + (5-5)^2 \cdot 0.2 \\
&\quad +(8-5)^2 \cdot 0.1 + (9-5)^2 \cdot 0.2 \\
&= 9 \cdot 0.1 + 4 \cdot 0.4 + 0 + 9 \cdot 0.1 + 16 \cdot 0.2 \\
&= 0.9 + 1.6 + 0.9 + 3.2 = 6.6 \\
\sigma_X &= \sqrt{6.6} = 2.57
\end{aligned}$$

Lösung 2.11:

$$\begin{aligned}
a)\ \operatorname{Var}(X) &= \int_{-\infty}^{\infty} x^2 f_X(x)dx - (E(X))^2 \\
&= \int_0^1 x^2 \cdot 2x\,dx - (\frac{2}{3})^2 = \int_0^1 2x^3\,dx - \frac{4}{9} \\
&= \left[\frac{x^4}{2}\right]_0^1 - \frac{4}{9} = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}
\end{aligned}$$

$$\begin{aligned}
b)\ \operatorname{Var}(X) &= \int_0^1 x^2 \cdot 4x^3\,dx - (\frac{4}{5})^2 = \int_0^1 4 \cdot x^5\,dx - \frac{16}{25} \\
&= \left[\frac{2x^6}{3}\right]_0^1 - \frac{16}{25} = \frac{2}{3} - \frac{16}{25} = \frac{50-48}{75} = \frac{2}{75}
\end{aligned}$$

Lösung 2.12:

$$\begin{aligned}
E(K) &= E(2X+10) = 2 \cdot E(X) + 10 \\
E(X) &= \int_{-\infty}^{\infty} f_X(x)x\,dx = \int_0^{12} \frac{1}{12}x\,dx \\
&= \frac{1}{12} \cdot \left[\frac{x^2}{2}\right]_0^{12} = \frac{1}{12} \cdot \frac{144}{2} = 6 \\
E(K) &= 2 \cdot 6 + 10 = 22
\end{aligned}$$

$$\begin{aligned}
\operatorname{Var}(K) &= \operatorname{Var}(2X+10) = 4\operatorname{Var}(X) \\
\operatorname{Var}(X) &= \int_{-\infty}^{\infty} f_X(x) \cdot (x - E(X))^2 dx \\
&= \int_0^{12} \frac{1}{12}(x-6)^2 dx = \frac{1}{12}\int_0^{12} (x^2 - 12x + 36)dx \\
&= \frac{1}{12}\left[\frac{x^3}{3} - 6x^2 + 36x\right]_0^{12} = 12 \\
\operatorname{Var}(K) &= 4 \cdot 12 = 48
\end{aligned}$$

Lösung 2.13:

Es sei $Y = X_1 + X_2 + X_3 + X_4$*. Dann gilt:*

$E(Y) = E(X_1) + E(X_2) + E(X_3) + E(X_4) = 18 + 12 + 5 + 28 = 63$

und

$$\begin{aligned} \mathrm{Var}(Y) &= \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \mathrm{Var}(X_3) + \mathrm{Var}(X_4) \\ &= 6 + 4 + 2 + 15 = 27. \end{aligned}$$

Lösung 2.14:

Für die Wandstärke Y *gilt* $Y = \frac{1}{2}(X_2 - X_1)$.

Es ist dann

$$E(Y) = \frac{E(X_2)}{2} - \frac{E(X_1)}{2} = \frac{810}{2} - \frac{800}{2} = \frac{10}{2} = 5$$

und

$$\mathrm{Var}(Y) = \tfrac{1}{4}\mathrm{Var}(X_2) + \tfrac{1}{4}\mathrm{Var}(X_1) = 0.005 + 0.0025 = 0.0075.$$

Lösung 2.15:

x_i \ y_j	1	2	3	$\sum$
1	0.2	0.2	0	0.4
2	0	0.2	0.4	0.6
$\sum$	0.2	0.4	0.4	1

a)

$$\begin{aligned} E(X) &= 1 \cdot 0.4 + 2 \cdot 0.6 = 1.6 \\ E(Y) &= 1 \cdot 0.2 + 2 \cdot 0.4 + 3 \cdot 0.4 = 2.2 \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(X) &= E(X^2) - (E(X))^2 = 1 \cdot 0.4 + 4 \cdot 0.6 - 2.56 = 0.24 \\ \mathrm{Var}(Y) &= E(Y^2) - (E(Y))^2 = 1 \cdot 0.2 + 4 \cdot 0.4 + 9 \cdot 0.4 - 4.84 = 0.56 \end{aligned}$$

b)

$f_y(y_j|x_i)$

x_i \ y_j	1	2	3	$\sum$
1	$\frac{1}{2}$	$\frac{1}{2}$	0	1
2	0	$\frac{1}{3}$	$\frac{2}{3}$	1

$f_x(x_i|y_j)$

x_i \ y_j	1	2	3
1	1	$\frac{1}{2}$	0
2	0	$\frac{1}{2}$	1
$\sum$	1	1	1

c)

$$\begin{aligned} \mathrm{Cov}(X,Y) &= E(X \cdot Y) - E(X) \cdot E(Y) \\ &= 1 \cdot 0.2 + 2 \cdot 0.2 + 4 \cdot 0.2 + 6 \cdot 0.4 - 1.6 \cdot 2.2 \\ &= 3.8 - 3.52 = 0.28 \end{aligned}$$

d)

$$\begin{aligned} \mathrm{Var}(2X + 3Y) &= 4\,\mathrm{Var}(X) + 9\,\mathrm{Var}(Y) + 2 \cdot 2 \cdot 3\,\mathrm{Cov}(X,Y) \\ &= 0.96 + 5.04 + 3.36 \\ &= 9.36 \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(2X - 3Y + 5) &= \mathrm{Var}(2X - 3Y) \\ &= 4\,\mathrm{Var}(x) + 9\,\mathrm{Var}(Y) + 2 \cdot 2 \cdot (-3)\,\mathrm{Cov}(X,Y) \\ &= 0.96 + 5.04 - 3.36 \\ &= 2.64 \end{aligned}$$

Lösung der Aufgaben zu Kapitel 3

Lösung 3.1:

Es liegt eine diskrete Gleichverteilung vor.

$$F_X(x) = \begin{cases} 0 & \text{für} \quad x < 2 \\ \frac{1}{3} & \text{für } 2 \leq x < 4 \\ \frac{2}{3} & \text{für } 4 \leq x < 6 \\ 1 & \text{für } 6 \leq x \end{cases}$$

$$\begin{aligned} E(X) &= \sum_{i=1}^{3} x_i f_X(x_i) = \frac{1}{3}\sum_{i=1}^{3} x_i \\ &= \frac{1}{3} \cdot 12 = 4 \\ \mathrm{Var}(X) &= E(X^2) - (E(X))^2 = \frac{1}{3}\sum_{i=1}^{3} x_i^2 - 4^2 \\ &= \frac{1}{3} \cdot (4 + 16 + 36) - 16 = \frac{56}{3} - \frac{48}{3} = \frac{8}{3} \end{aligned}$$

Lösung 3.2:

Eine Zufallsvariable X ist gleichverteilt über $[0;2]$.

a) *Für eine über* $[a;b]$ *stetige gleichverteilte Zufallsvariable gilt* $E(X) = \frac{b+a}{2}$ *und* $\mathrm{Var}(X) = \frac{(b-a)^2}{12}$. *Somit gilt hier* $E(X) = 1$ *und* $\mathrm{Var}(X) = \frac{1}{3}$.

b) *Der Erwartungswert von X^2 berechnet sich zu*

$$E(X^2) = \int_0^2 \frac{1}{2}x^2 \, dx = \frac{1}{2}\left[\frac{1}{3}x^3\right]_0^2 = \frac{1}{2} \cdot \frac{8}{3} = \frac{4}{3}.$$

Die Varianz von X^2 berechnet sich zu

$$\begin{aligned}
\mathrm{Var}(X^2) &= E(X^4) - (E(X^2))^2 \\
&= \int_0^2 \frac{1}{2}x^4\,dx - \frac{16}{9} = \frac{1}{2}\left[\frac{1}{5}x^5\right]_0^2 - \frac{16}{9} \\
&= \frac{16}{5} - \frac{16}{9} \\
&= \frac{144-80}{45} = \frac{64}{45}.
\end{aligned}$$

Lösung 3.3:

Es liegt eine Binomialverteilung vor mit $n = 4$ und $\pi = \frac{1}{2}$.

a)

$$x = 2$$

$$\binom{n}{k}\pi^x(1-\pi)^{n-x} = \binom{4}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2 = \frac{4!}{2!\;2!}\cdot\frac{1}{2^4} = \frac{6}{16} = \frac{3}{8}$$

b)

$$x = 3$$

$$\binom{n}{k}\pi^x(1-\pi)^{n-x} = \binom{4}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^1 = \frac{4!}{1!\;3!}\cdot\frac{1}{2^4} = \frac{4}{16} = \frac{2}{8}$$

Lösung 3.4:

	a) $\pi = 0.20 = \frac{1}{5}$ $1-\pi = \frac{4}{5}$	b) $\pi = 0.50 = \frac{1}{2}$ $1-\pi = \frac{1}{2}$	c) $\pi = 0.60 = \frac{3}{5}$ $1-\pi = \frac{2}{5}$
$x = 0$	$\boldsymbol{P}(0) = \frac{256}{625} = 0.4096$	$\boldsymbol{P}(0) = \frac{1}{16} = 0.0625$	$\boldsymbol{P}(0) = \frac{16}{625} = 0.0256$
$x = 1$	$\boldsymbol{P}(1) = \frac{256}{625} = 0.4096$	$\boldsymbol{P}(1) = \frac{4}{16} = 0.25$	$\boldsymbol{P}(1) = \frac{96}{625} = 0.1536$
$x = 2$	$\boldsymbol{P}(2) = \frac{96}{625} = 0.1536$	$\boldsymbol{P}(2) = \frac{6}{16} = 0.375$	$\boldsymbol{P}(2) = \frac{216}{625} = 0.3456$
$x = 3$	$\boldsymbol{P}(3) = \frac{16}{625} = 0.0256$	$\boldsymbol{P}(3) = \frac{4}{16} = 0.25$	$\boldsymbol{P}(3) = \frac{216}{625} = 0.3456$
$x = 4$	$\boldsymbol{P}(4) = \frac{1}{625} = 0.0016$	$\boldsymbol{P}(4) = \frac{1}{16} = 0.0625$	$\boldsymbol{P}(4) = \frac{81}{625} = 0.1296$

*Y ist die Zahl der fehlerfreien Stücke. Sind unter n Stücken y **fehlerfreie** Stücke, so sind auch gleichzeitig $x = n - y$ **fehlerhafte** Stücke enthalten. Es gilt also: $\boldsymbol{P}$(y-fehlerfreie Stücke) = $\boldsymbol{P}$(x-fehlerhafte Stücke). Die gesuchten Wahrscheinlichkeiten können also unmittelbar in a) abgelesen werden ($n = 4$, $\pi = 0.2$).*

$$\begin{aligned}
\boldsymbol{P}(Y = 0) &= \boldsymbol{P}(X = 4) = 0.0016 \\
\boldsymbol{P}(Y = 1) &= \boldsymbol{P}(X = 3) = 0.0256 \\
\boldsymbol{P}(Y = 2) &= \boldsymbol{P}(X = 2) = 0.1536 \\
\boldsymbol{P}(Y = 3) &= \boldsymbol{P}(X = 1) = 0.4096 \\
\boldsymbol{P}(Y = 4) &= \boldsymbol{P}(X = 0) = 0.4096
\end{aligned}$$

Lösung 3.5:

a) 0.4918 *b)* 0.4032 *c)* 0.5763 *d)* 0.9821

e) $0.5 - 0.1554 = 0.3446$ *f)* $0.5 - 0.0398 = 0.4602$

g) $0.9452 - 0.5793 = 0.3659$ *h)* $0.3849 + 0.4192 = 0.8041$

i) $0.9772 - 0.8413 = 0.1359$

Lösung 3.6:

a) $A = 0.25$ *b)* $B = 0.84$

c) $C = 0.84$ *d)* $D = 1.04$

Lösung 3.7:

a) 0.8664 *b)* 0.7745

c) 0.0062 *d)* 0.0013

e) 0.16

Lösung 3.8:

a) $$\boldsymbol{P}(X \leq A) = \boldsymbol{P}\left(Z \leq \frac{A-100}{10}\right) = \boldsymbol{P}(Z \leq z_1) = F_Z(z_1) = 0.7$$

$$z_1 = 0.52, \ \textit{daraus folgt} \ \frac{A-100}{10} = 0.52 \ \textit{oder} \ A = 105.2.$$

b) $$\boldsymbol{P}(X \geq B) = \boldsymbol{P}\left(Z \geq \frac{B-100}{10}\right) = \boldsymbol{P}(Z \geq z_2) = 0.5 + F_1(z_2) = 0.65$$

$$F_1(z_2) = 0.15, \ z_2 = -0.38, \ \frac{B-100}{10} = -0.38 \ \textit{oder} \ B = 96.2.$$

c) $$\begin{aligned} \boldsymbol{P}(|X-100| \leq C) &= \boldsymbol{P}(100 - C \leq X \leq 100 + C) \\ &= \boldsymbol{P}\left(-\frac{C}{10} \leq Z \leq \frac{C}{10}\right) = F_2\left(\frac{C}{10}\right) = 0.5 \end{aligned}$$

$$\frac{C}{10} = 0.67 \text{ und somit } C = 6.7.$$

Lösung 3.9:

a) $$\begin{aligned} \boldsymbol{P}(X \leq 99.8) &= \boldsymbol{P}\left(\frac{X-\mu}{\sigma} \leq \frac{99.8-100}{0.2}\right) \\ &= \boldsymbol{P}(Z \leq -1) = 1 - F_Z(1) = 0.1587 \\ &\mathrel{\widehat{=}} 15.87\% \ \textit{Ausschuss} \end{aligned}$$

b) $$\begin{aligned} \boldsymbol{P}(X \geq 100.6) &= \boldsymbol{P}\left(Z \geq \frac{100.6-100}{0.2}\right) = \boldsymbol{P}(Z \geq 3) \\ &= 1 - F_Z(3) = 0.0013 \mathrel{\widehat{=}} 0.13\% \ \textit{Ausschuss} \end{aligned}$$

c) $$\begin{aligned} \boldsymbol{P}(|X-100| \geq 0.3) &= \boldsymbol{P}\left(|Z| \geq \frac{0.3}{0.2}\right) = \boldsymbol{P}(|Z| \geq 1.5) \\ &= 1 - F_2(1.5) = 0.1336 \mathrel{\widehat{=}} 13.36\% \ \textit{Ausschuss} \end{aligned}$$

d) $\boldsymbol{P}(|X-100| \geq C) = 0.05 \Leftrightarrow \boldsymbol{P}\left(|Z| \geq \frac{C}{0.2}\right) = 0.05$

$F_2(1.96) = 0.95$ *bzw.* $1 - F_2(1.96) = 0.05$

$\Rightarrow \frac{C}{0.2} = 1.96$ *bzw.* $C = 0.392$

Die Toleranzgrenzen lauten daher:

$99.608 \leq X \leq 100.392.$

e) *Die Zufallsvariable* X_1 *hat den Erwartungswert* $\mu_1 = 100.1$ *und die Standardabweichung* $\sigma_1 = \sigma = 0.2$. *Es gilt dann für den Ausschussprozentsatz:*

$$1 - \boldsymbol{P}(99.608 \leq X_1 \leq 100.392)$$

$$= 1 - \boldsymbol{P}\left(\frac{99.608 - 100.1}{0.2} \leq \frac{X_1 - \mu_1}{\sigma_1} \leq \frac{100.392 - 100.1}{0.2}\right)$$

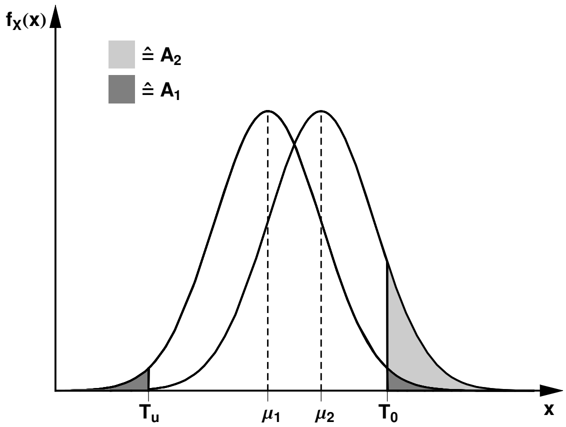
$$= 1 - \boldsymbol{P}(-2.46 \leq Z \leq 1.46) = 1 - [F_1(2.46) + F_1(1.46)]$$

$$= 1 - (0.4931 + 0.4279) = 1 - 0.9210 = 0.079$$

$\widehat{=}$ 7.9% *Ausschuss.*

<u>*Lösung 3.10:*</u>

a) *Es ist am günstigsten, den Sollwert so einzustellen, dass ein symmetrisches Intervall für die zulässigen Foliendurchmesser um den Sollwert vorliegt. Die Wahrscheinlichkeit dafür, dass ein Foliendurchmesser innerhalb des Toleranzintervalls liegt, ist dann am größten. Dies gilt für alle symmetrischen Verteilungen, deren Dichte mit wachsender Entfernung zum Erwartungswert monoton abnimmt.*



$\mu_1 = \dfrac{T_u + T_o}{2}$ *(symmetrisches Intervall)*

$2 \cdot A_1 =$ *Ausschuss für Sollwert* μ_1

$\mu_2 > \mu_1$ *(asymmetrisches Intervall)*

$A_2 =$ *Ausschuss für Sollwert* μ_2

Also verringert sich die Wahrscheinlichkeit dafür, dass eine Folie innerhalb der Toleranzgrenzen liegt, wenn der Sollwert von der Intervallmitte verschoben wird.

Als Sollwert ergibt sich $\mu = 1.0$ *mm.*

b) *Zunächst wird berechnet, wie viel gute Folien bei der Herstellung von* 1000 *Stück produziert werden, wenn Maschine A bzw. Maschine B eingesetzt wird.*

Berechnet wird die absolute Häufigkeit der Folien, die bei 1000 *produzierten Stücken kein Ausschuss sind. Dazu wird zunächst die Wahrscheinlichkeit berechnet, dass ein Foliendurchmesser innerhalb der Toleranzgrenzen liegt.*

$$\begin{aligned}
\boldsymbol{P}(T_u \le X \le T_o) &= \boldsymbol{P}\left(\frac{T_u - \mu}{\sigma} \le Z \le \frac{T_o - \mu}{\sigma}\right) \\
&= \boldsymbol{P}\left(-\frac{T_o - T_u}{2\sigma} \le Z \le \frac{T_o - T_u}{2\sigma}\right) \\
&= \boldsymbol{P}\left(|Z| \le \frac{T_o - T_u}{2\sigma}\right)
\end{aligned}$$

Zu $z = \dfrac{T_o - T_u}{2\sigma}$ *ist der* F_2*-Wert gesucht.*

$$A : \frac{1.18 - 0.82}{2 \cdot 0.1} = 1.8 \Rightarrow F_2 = 0.9282$$

$$B : \frac{1.18 - 0.82}{2 \cdot 0.18} = 1 \Rightarrow F_2 = 0.6826$$

D.h., bei Maschine A sind bei einer Produktion von 1000 *Stück* 928 *gute Folien und bei B* 683 *gute Folien.*

Daraus ergeben sich die Stückkosten für A und B.

$A : \quad 20 : 928 = 0.02155$

$B : \quad 16 : 683 = 0.02343$

Da die Stückkosten bei der Maschine A für einwandfreie Stücke niedriger sind als bei Maschine B, wird die Entscheidung zugunsten Maschine A gefällt.

Lösung 3.11:

a) $\boldsymbol{P}(X_1 + X_2 \leq 21) = \boldsymbol{P}\big(Z \leq \frac{21-30}{5} = -1.8\big)$

$= 0.5 - F_1(1.8) = 0.5 - 0.4641 = 0.0359$

b) $\boldsymbol{P}(24 \leq X_1 + X_2 \leq 42) = \boldsymbol{P}\big(\frac{24-30}{5} \leq Z \leq \frac{42-30}{5}\big)$

$= \boldsymbol{P}(-1.2 \leq Z \leq 2.4)$

$= F_1(1.2) + F_1(2.4) = 0.3849 + 0.4918 = 0.8767$

Lösung 3.12:

a) $x_1 = \chi^2_{0.95}(25) = 37.65$ *b)* $x_2 = \chi^2_{0.1}(25) = 16.47$

c) $x_3 = \chi^2_{0.9}(25) = 34.38$ *d)* $x_4 = \chi^2_{0.99}(25) = 44.31$

Lösung 3.13:

a) 0.97 *b)* 0.97

Lösung 3.14:

a) $t_1 = 1.782$ *b)* $t_2 = 2.681$

c) $t_3 = -1.356$ *d)* $t_4 = 1.356$

Lösung 3.15:

Die Zufallsvariable $X = \sum_{i=1}^{6} X_i^2$ *ist* χ^2*-verteilt mit* 6 *Freiheitsgraden und die Zufallsvariable* $Y = \sum_{j=1}^{4} Y_j^2$ *ist* χ^2*-verteilt mit* 4 *Freiheitsgraden.*

$$W = \frac{2X}{3Y} = \frac{X/6}{Y/4} \sim F(6,4)$$

$$F_{0.95}(6,4) = 6.26 \qquad F_{0.05}(6,4) = \frac{1}{F_{0.95}(4,6)} = \frac{1}{4.53} = 0.221$$

$$F_{0.99}(6,4) = 15.2 \qquad F_{0.01}(6,4) = \frac{1}{F_{0.99}(4,6)} = \frac{1}{9.15} = 0.109$$

Lösung 3.16:

Die Zufallsvariable $X = \sum_{i=1}^{3} X_i^2$ *ist* χ^2*-verteilt mit* 3 *Freiheitsgraden und die Zufallsvariable* $Y = \sum_{j=1}^{5} Y_j^2$ *ist* χ^2*-verteilt mit* 5 *Freiheitsgraden.*

$$W = \frac{20X}{6Y} = 2 \cdot \frac{X/3}{Y/5} = 2F \text{ mit } F \sim F(3,5).$$

$$\begin{aligned} E(2F) &= 2E(F) = 2 \cdot \frac{5}{3} = \frac{10}{3} \\ \mathrm{Var}(2F) &= 4\mathrm{Var}(F) = 4 \cdot \frac{2 \cdot 5^2 \cdot 6}{3 \cdot 1 \cdot 3^2} = 4 \cdot \frac{100}{9} = \frac{400}{9} \end{aligned}$$