Statistik für VWL

Inhaltsverzeichnis

Universität St.Gallen

Teil II: Statistik

Teil I:

Wahrscheinlichkeitsrechnung

Universität St.Gallen

Kapitel 8.1

1. Grundlagen der Wahrscheinlichkeitstheorie

1.1. Ereignisse, Ereignisraum und Ereignismenge

Ein Experiment heisst ***Zufallsexperiment***, wenn es:

- ➔ nach einer ganz bestimmten Vorschrift aus-geführt wird;
- ➔ unter gleichen Bedingungen beliebig oft wiederholbar ist; und
- ➔ das Ergebnis ungewiss ist und nicht vorausgesagt werden kann.

Definition

Die einzelnen, nicht weiter zerlegbaren und sich gegenseitig ausschliessenden Ausgänge oder Ergebnisse eines Zufallsexperimentes heissen

Elementarereignisse.

Bsp.:

Elementar-ereignisse: Z, K

Universität St.Gallen

Definition

Die Menge S aller Elementarereignisse eines Zufallsexperiments heisst

Ereignisraum S

dieses Zufallsexperiments.

Bsp: (Münzwurf) $S = \{Z, K\}$

Beispiel 1.1.1:

•) Beim Werfen eines Würfels:

$$S = \{1, 2, \dots, 6\}$$

•) Wirft man eine Münze zweimal:

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

•) Man werfe eine Münze so lange, bis Kopf erscheint:

•) Für die Lebensdauer einer Glühbirne gilt:

$$S = \{t : t \geq 0\} = \mathbb{R}_+$$

•) Für die Entwicklung eines Aktienkurses ist eine mögliche Wahl des Funktionenraumes

$$S = \{\text{alle funktionen } \mathbb{R}_+ \rightarrow \mathbb{R}_+\}$$

Beispiel 1.1.1:

•) Beim Werfen eines Würfels:

$$S = \{1, 2, 3, 4, 5, 6\}$$

•) Wirft man eine Münze zweimal:

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

•) Man werfe eine Münze so lange, bis Kopf erscheint:

$$S = \{K, ZK, ZZK, ZZZK, ...\}$$

(unendlich viele Elemente, aber abzählbar)

•) Für die Lebensdauer einer Glühbirne gilt:

$$S = \{t : t \geq 0\} = [0,\infty) = \mathbf{R}^+$$

(mehr als endlich viele oder abzählbar unendlich viele Elemente) ➔ stetiges Kontinuum von Elementarereignissen.

•) Für die Entwicklung eines Aktienkurses ist eine mögliche Wahl des Funktionenraumes

$$S = \{\text{alle Funktionen: } R_+ \to R_+\}.$$

Merke:

Ein Zufallsexperiment kann durch mehrere Ereignisräume beschrieben werden.

Beispiel 1.1.2:

2 Münzen werden geworfen:

→ falls die Ausgänge K und Z beider Münzen interessieren:

$S_1 = \{KK, KZ, ZK, ZZ\}$

→ falls die Anzahl K und die Anzahl Z interessieren:

$S_2 = \{(2,0), (1,1), (0,2)\}$

→ falls interessiert, ob die Bilder gleich (g) oder verschieden (u) sind:

$S_3 = \{\text{"g"}, \text{"u"}\}$

Beispiel 1.1.2:

2 Münzen werden geworfen:

→ falls die Ausgänge K und Z beider Münzen interessieren:

$$S_1 = \{(K, K), (K, Z), (Z, K), (Z, Z)\}$$

→ falls die Anzahl K und die Anzahl Z interessieren:

$$S_2 = \{(2, 0), (1, 1), (0, 2)\}$$

→ falls interessiert, ob die Bilder gleich (g) oder verschieden (u) sind:

$$S_3 = \{\{g\}, \{u\}\}.$$

$S = \{s_1, s_2, \ldots\}$

$A \subseteq S$

Definition

Ein ***zufälliges Ereignis*** A ist eine Teilmenge des Ereignisraumes S.

Das Ereignis A ist eingetreten, wenn das Ergebnis des Zufallsexperimentes ein Element dieser Teilmenge A ist.

$\rightarrow s_i$, wenn $s_i \in A \Rightarrow A$ ist eingetreten

Universität St.Gallen

• Ereignisraum S

Definition

Alle Ereignisse eines Zufallsexperimentes mit dem Ereignisraum S bilden die dazugehörige

Ereignismenge $\underbrace{E(S)}_{E}$. ← Menge der Teilmengen von S

Bsp: Münzwurf: $S = \{K, Z\}$

$E = \{K, Z, \{K, Z\}, \emptyset\}$

Zur Ereignismenge E müssen zwei besondere Teilmengen von S gehören:

1. Der Ereignisraum selbst als das sogenannte ***sichere Ereignis:*** $S \subset E$;

2. Die leere Menge als das sogenannte ***unmögliche Ereignis***: $\emptyset \subset E$.

Kapitel 8.2

$\bar{A} = S \setminus A$

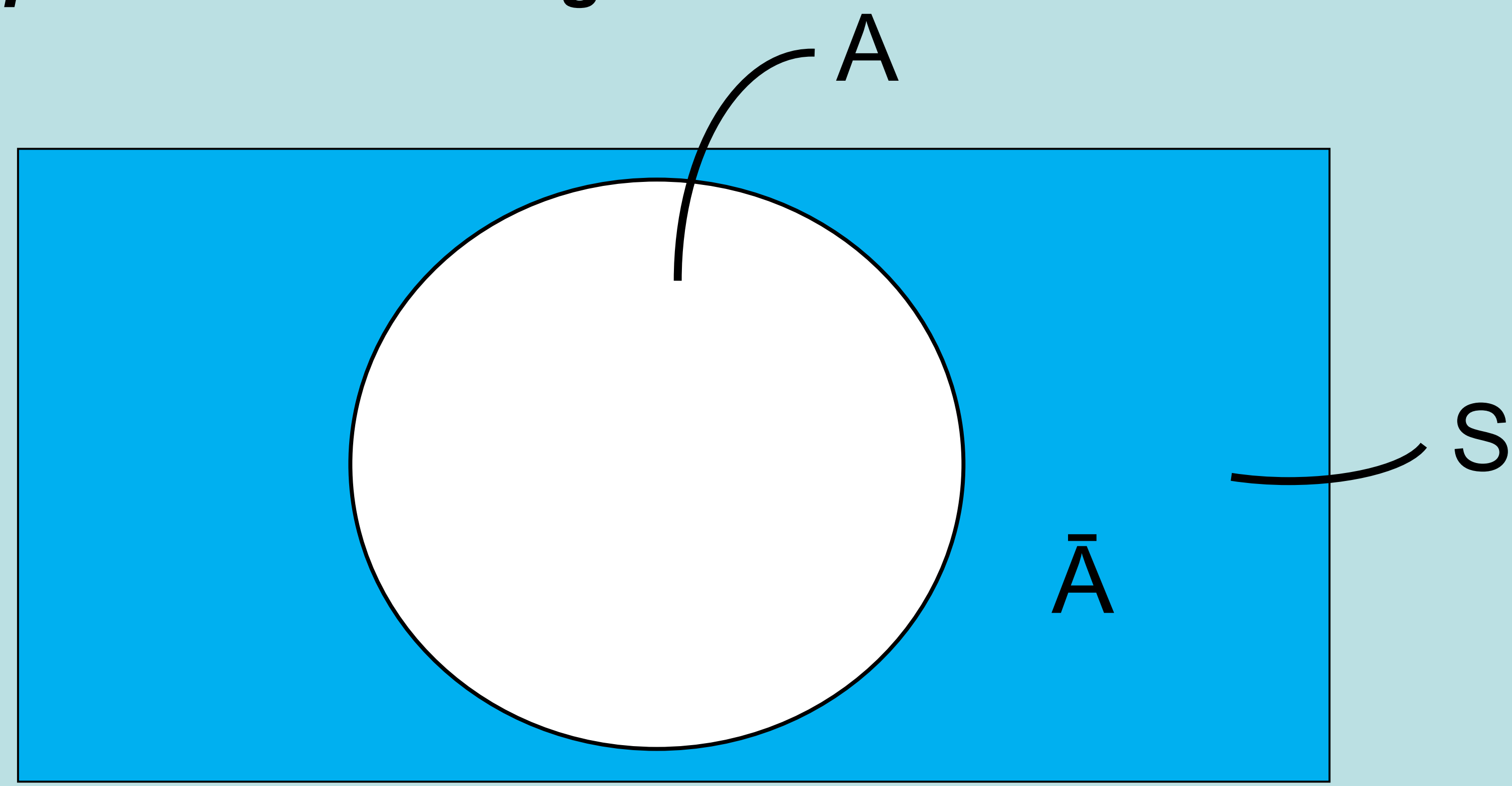
1.2. Das Rechnen mit Ereignissen

Negation: $\bar{A}$ ("nicht A")

tritt genau dann ein, wenn A nicht eintritt.

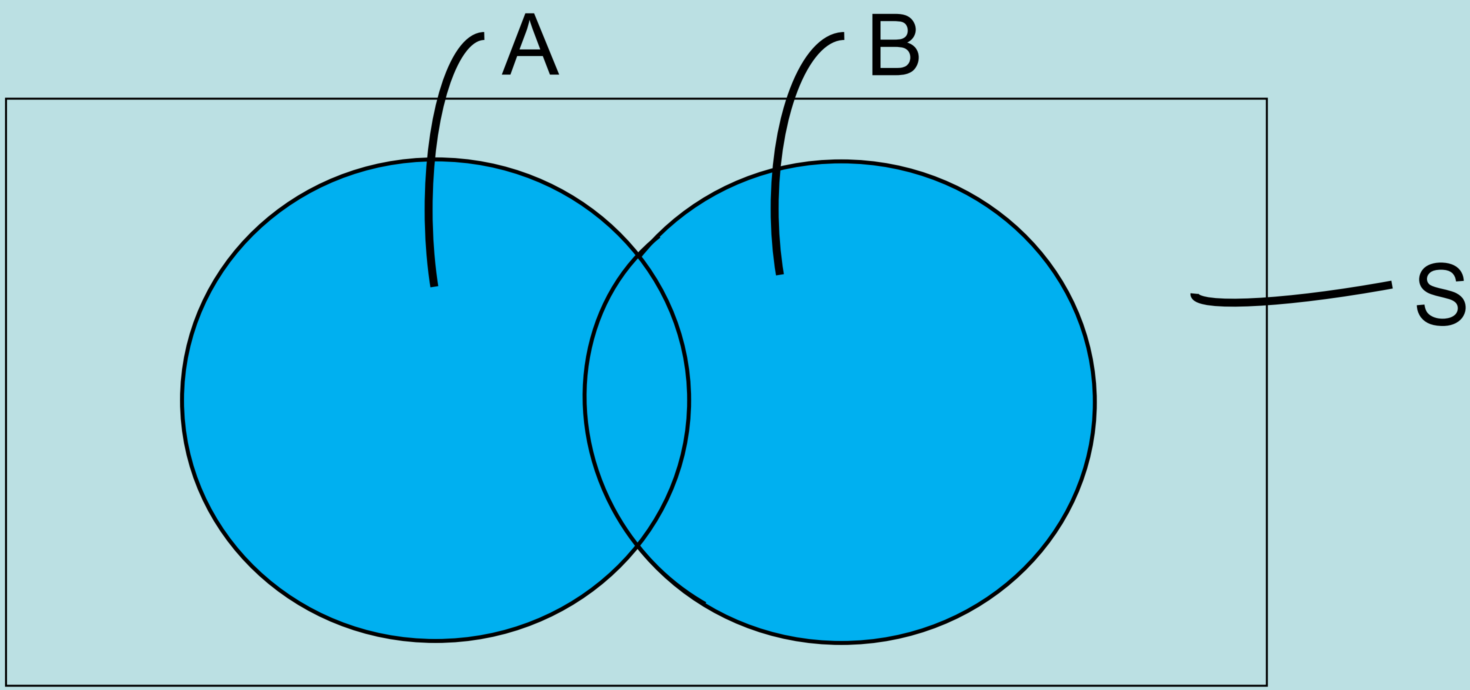
Definition

Das Ereignis $\bar{A}$ = S\A heisst das zu A ***komplementäre Ereignis***.



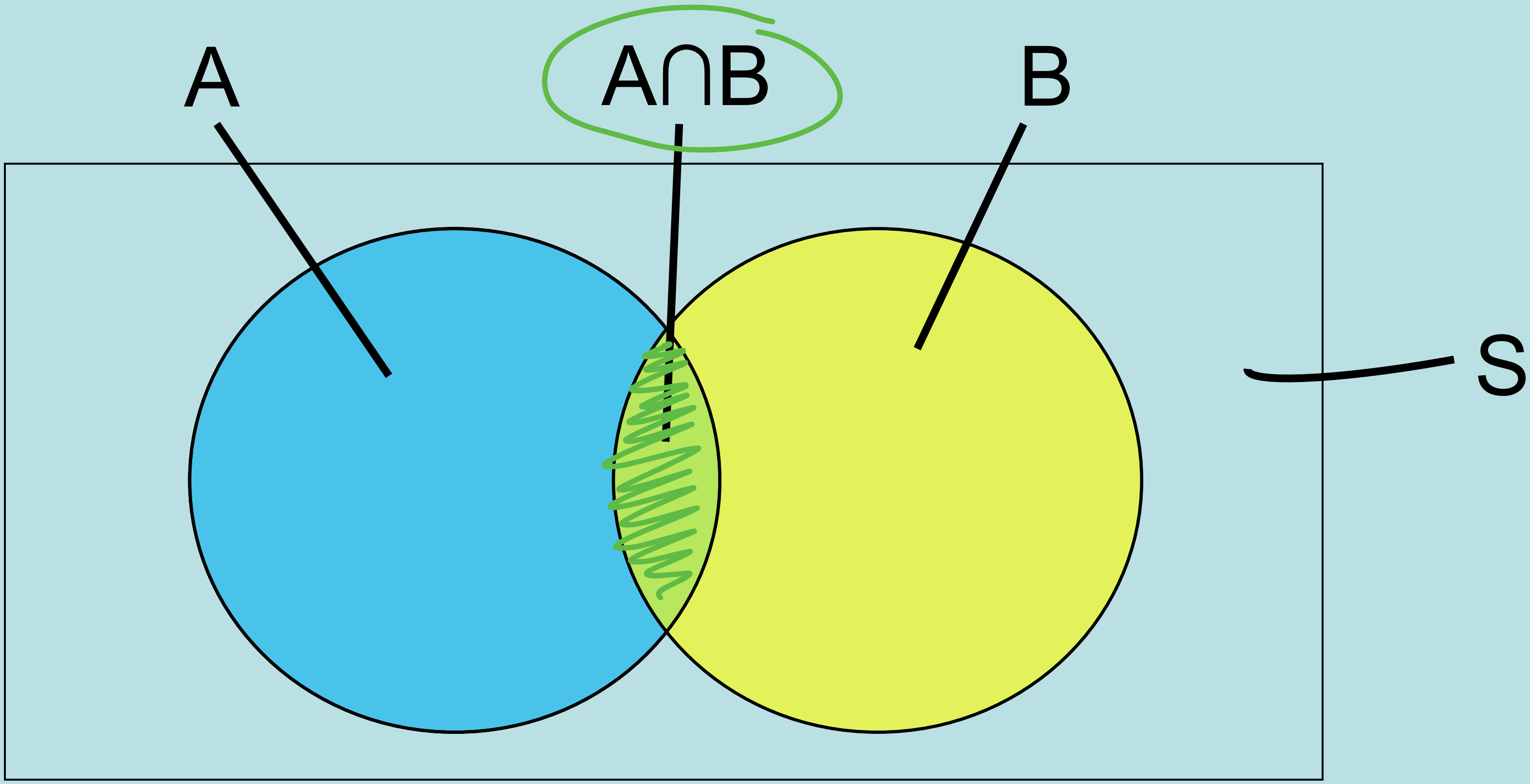
Definition

Vereinigung: $A \cup B$ („A oder B„) tritt genau dann ein, wenn Ereignis A oder Ereignis B oder beide zugleich eintreten.



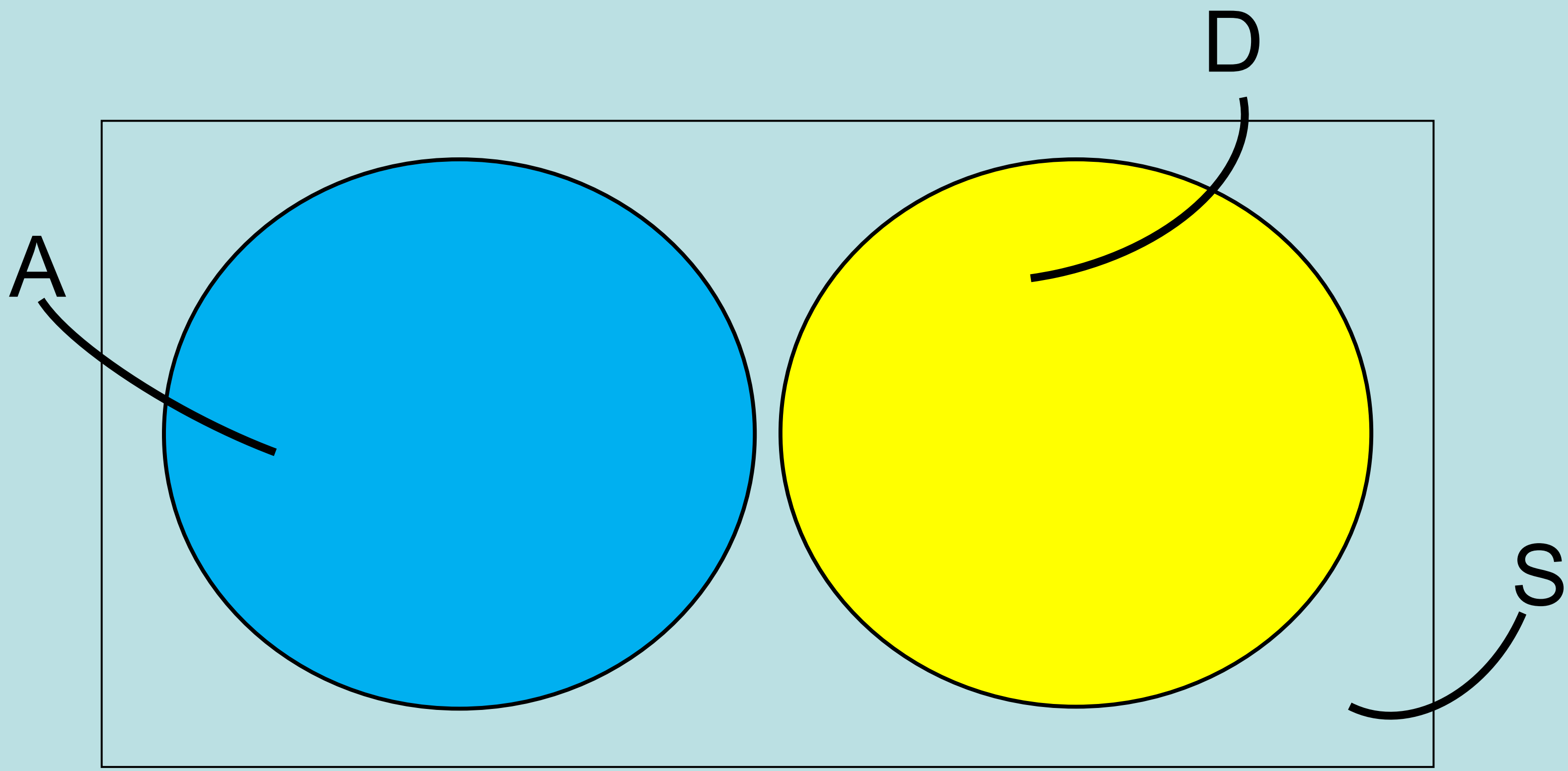
Definition

Durchschnitt: A ∩ B („A und B") tritt genau dann ein, wenn Ereignis A und Ereignis B gleichzeitig eintreten.



Definition

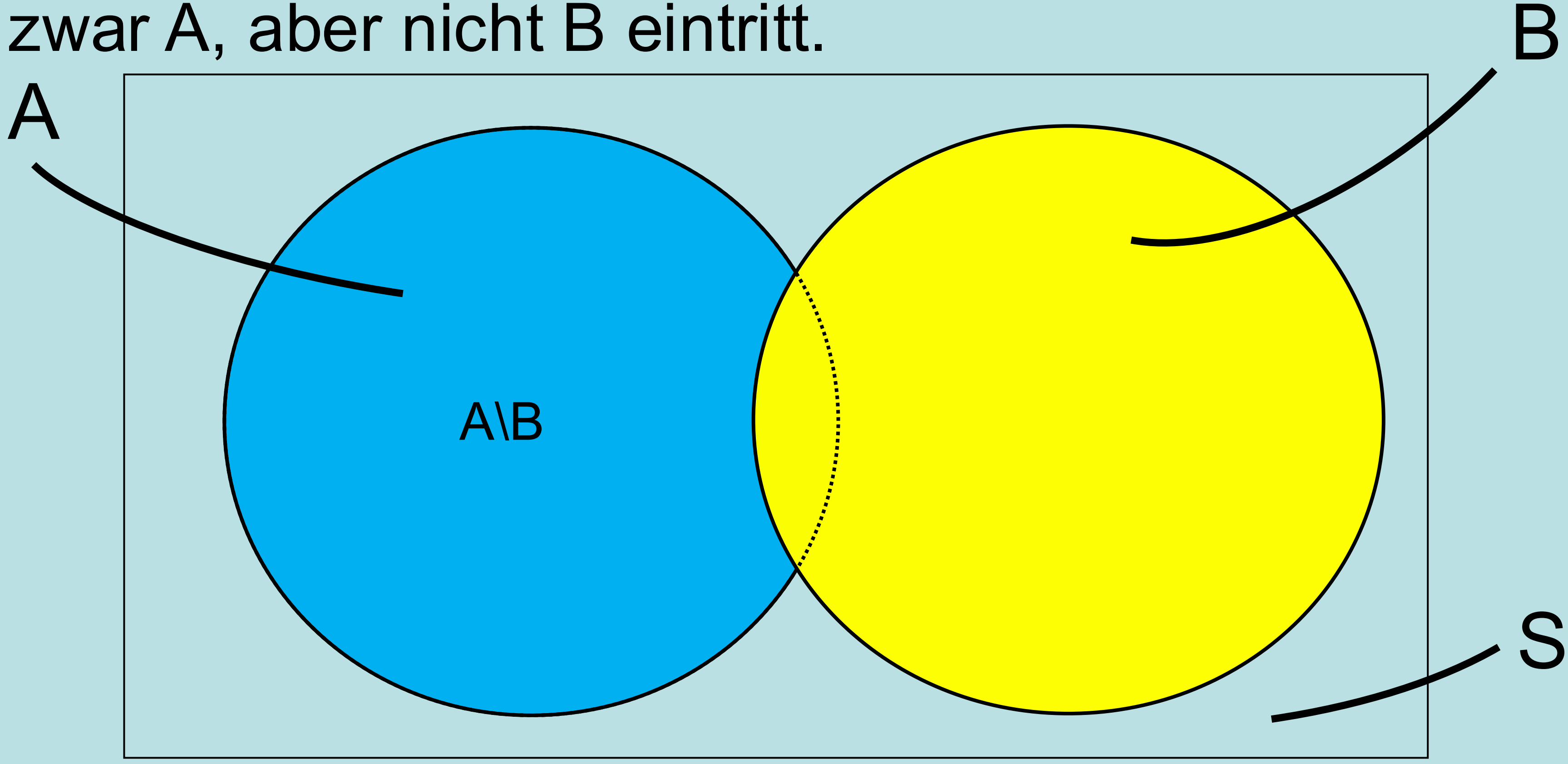
Zwei Ereignisse A und D heissen ***disjunkt*** oder ***unvereinbar***, wenn $A \cap D = \emptyset$ ist.



Disjunkte Ereignisse

Definition

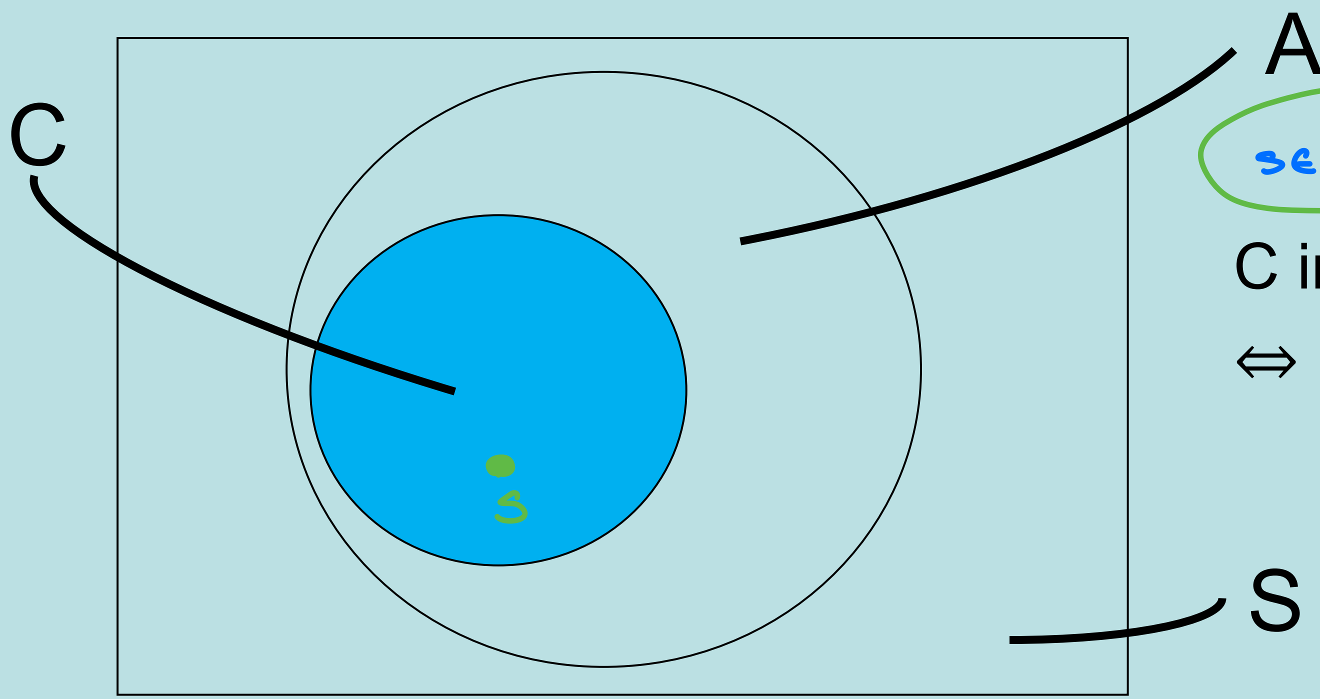
Differenz: A\B („A ohne B“) tritt genau dann ein, wenn zwar A, aber nicht B eintritt.



Universität St.Gallen

Definition

Tritt ein Ereignis A stets ein, wenn ein Ereignis C eintritt, so sagt man,
Ereignis C impliziert Ereignis A.



$s \in C \Rightarrow s \in A$

C impliziert A
$\Leftrightarrow C \subset A$

C impliziert A

Universität St.Gallen

Kapitel 8.3-5

1.3. Der Wahrscheinlichkeitsbegriff

Wahrscheinlichkeitsmass

$P(A)$

$\mathbb{P}(A)$

$A \subseteq S$

$$P: \quad E \longrightarrow \mathbb{R} = [0,1]$$

$A \in E$

$$A \longmapsto P(A)$$

ist eine reellwertige Funktion:

Jedem Element von E wird genau ein Element von $\mathbb{R}$ zugeordnet.

2) Statistische Wahrscheinlichkeit

→ Experiment als einzig zulässiges Verfahren zur Bestimmung von Eintrittswahrscheinlichkeiten

Definition

Der Grenzwert $P(A) = \lim_{n \to \infty} h_n(A)$

heisst die ***statistische Wahrscheinlichkeit*** des Ereignisses A, mit

$h_n(A)$ = "relative Häufigkeit des Eintretens von A".

in n Versuchen

Bsp: Münzwurf $A = \{K\}$

$h_n(A)$ = relative Häufigkeit von Kopf in n Würfen

$\mathbb{P}$("eine 6 zu würfeln")
$= 1/6$

Beispiel 1.3.1:

Ein Würfel werde 3‘000-mal hintereinander geworfen. Nach jedem Wurf notieren wir, wie viele Sechser bisher gefallen sind.

(z.B. Tabelle 8.1 im Buch)

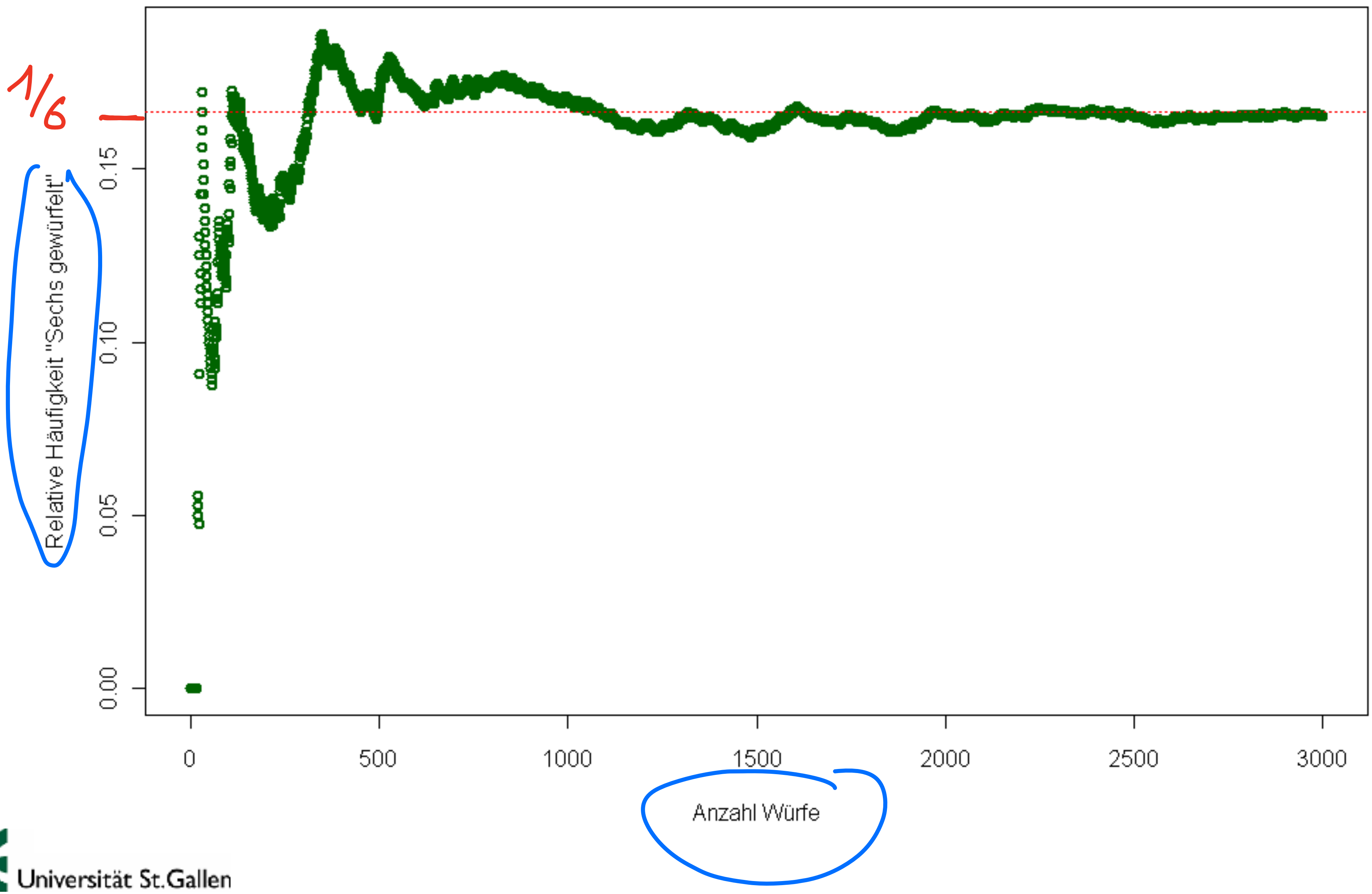
Was hätten wir für $P(\text{"}6\text{"})$ erwartet?

$$P(\text{"}6\text{"}) = \tfrac{1}{6} \rightarrow \text{drittes Konzept} \cong 0.1\overline{666}$$

"Klassische Wahrscheinlichkeit"

Universität St.Gallen

Illustration Beispiel 1.3.1:



Universität St.Gallen

3) Klassische Wahrscheinlichkeit

$\longrightarrow$ entwickelt von Laplace, 1812

Allerdings hat Bernoulli schon über 100 Jahre früher denselben Vorschlag gemacht.

“Wenn ein Experiment eine Anzahl verschiedener und **gleich möglicher** Ausgänge hervorbringen kann und einige davon als günstig anzusehen sind, dann ist die Wahrscheinlichkeit eines günstigen Ausgangs (Ereignis A) gleich dem Verhältnis der Anzahl der günstigen zur Anzahl der möglichen Ausgänge.“

Laplace Methode

$$P(A) = \frac{\#\text{ der günstigen Ausgänge}}{\#\text{ der möglichen Ausgänge}} = \frac{g}{m}$$

Definition

Ein Zuffallsexperiment mit endlich vielen ***gleichwahrscheinlichen*** Elementarereignissen heisst

Laplace-Experiment.

Universität St.Gallen

Beispiel 1.3.2:

Eine Münze und ein Würfel werden gemeinsam geworfen. Wir stellen uns folgende Frage:

Wahrscheinlichkeit, dass

A = “Kopf und Augenzahl strikt grösser als 4 erscheint“

Elementarereignisse

$$S = \{(K, \leq 4), (K, > 4), (Z, \leq 4), (Z, > 4)\}$$

Sind diese Elementarereignisse gleichwahrscheinlich?

$$S_2 = \{(K,1), (K,2), (K,3), \ldots, (K,6), (Z,1), \ldots, (Z,6)\}$$

gleichwahrscheinlich? → Ja

$$A = \{(K,5), (K,6)\} \subseteq S_2$$

$$\mathbb{P}(A) = \frac{\#\text{ günstige Ausgänge}}{\#\text{ mögliche Ausgänge}} = \frac{2}{12} = \underline{\underline{\frac{1}{6}}}$$

Beispiel 1.3.2:

Eine Münze und ein Würfel werden gemeinsam geworfen. Wir stellen uns folgende Frage:

Wahrscheinlichkeit, dass

A = “Kopf und Augenzahl grösser als 4 erscheint“

Konstruktion des Ereignisraumes:

Elementarereignisse: $(\mathrm{K}, \leq 4), (\mathrm{K}, > 4), (\mathrm{Z}, \leq 4), (\mathrm{Z}, > 4)$

sind gleichwahrscheinlich? Nein!

$\rightarrow$ Elementarereignisse:

$$(K,1),(Z,1),(K,2),(Z,2),\ldots,(K,6),(Z,6)$$

gleichwahrscheinlich!

Benützen wir das Laplace-Modell

$$\Rightarrow A = \{(K,5),(K,6)\} \text{ und } P(A) = \frac{2}{12} = 1/6$$

Universität St.Gallen

Beispiel 1.3.3:

Beim zweimaligen Münzwurf ist

$S = \{KK, KZ, ZK, ZZ\}$.

Frage: $\mathbb{P}$("mindestens einmal Kopf") = ?

$$= \frac{3}{4} = 3/4$$

Beispiel 1.3.3:

Beim zweimaligen Münzwurf ist

$$S = \{KK, KZ, ZK, ZZ\}.$$

$$P\left[\text{"mindestens einmal Kopf"}\right] = \frac{3}{4}$$

Kapitel 8.6

Kolmogorov 1933

1.4. Axiome der Wahrscheinlichkeitstheorie

Definition: Eine Funktion $\mathbb{P}: E \to \mathbb{R}$, $A \mapsto \mathbb{P}(A)$, heisst Wahrscheinlichkeitsfunktion (Wahrscheinlichkeitsmass), wenn

$\boxed{A1}$ $\mathbb{P}(A) \geq 0$ für alle $A \in E$

$\boxed{A2}$ $\mathbb{P}(S) = 1$ (sicheres Ereignis)

$\boxed{A3}$ $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ falls $A \cap B = \emptyset$

Thm 1: $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$ für jedes $A \in E$

Bew: $\mathbb{P}(S) \overset{A2}{=} 1 \overset{A3}{=} \mathbb{P}(A) + \mathbb{P}(\bar{A})$ ✓

$S = A \cup \bar{A}$

disjunkt

Thm 2: $\mathbb{P}(\emptyset) = 0$

Bew:

$\mathbb{P}(\underbrace{S \cup \emptyset}_{S}) \overset{A3}{=} \mathbb{P}(S) + \mathbb{P}(\emptyset)$

$\Longrightarrow \mathbb{P}(\emptyset) = 0$

1.4. Axiome der Wahrscheinlichkeitstheorie

Definition

Eine jede Funktion P

$$P: \quad E \longrightarrow \mathbb{R}$$

$$A \longrightarrow P(A)$$

die jedem Ereignis A aus der Ereignismenge E eine reelle Zahl zuordnet, heisst Wahrscheinlichkeitsfunktion und P(A) heisst Wahrscheinlichkeit von A wenn sie die folgenden Axiome (von Kolmogorov, 1933) erfüllt:

Definition (Fortsetzung)

Axiom 1: $P(A) \geq 0, \quad \forall A \in E$

Axiom 2: $P(S) = 1$

Axiom 3: $P(A \cup B) = P(A) + P(B),$

falls $A \cap B = \emptyset$

(Additionsregel für disjunkte Ereignisse)

Universität St.Gallen

Kapitel 8.7

1.5. Wichtige Regeln der Wahrscheinlichkeitsrechnung

Sei $n \in \mathbb{N}$ beliebig

Thm 3: Seien Ereignisse $A_1, A_2, \dots, A_n \in E$ paarweise disjunkt. Dann gilt

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

Beweis: per induktion → Übungsaufgabe

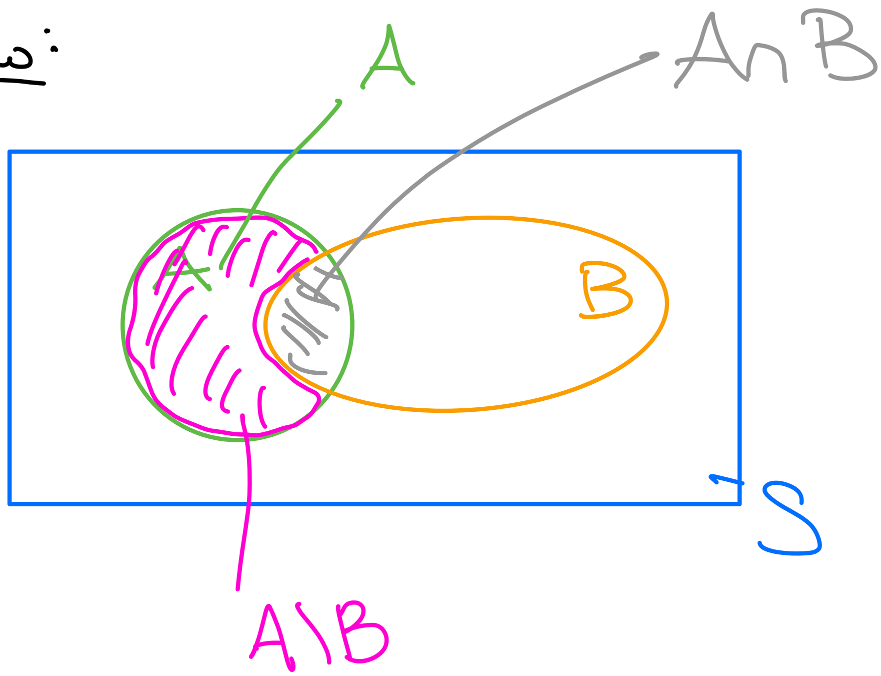
$n=2$ $\quad A_1, A_2 \quad , \quad A_1 \cap A_2 = \emptyset$

$$\mathbb{P}(A_1 \cup A_2) \overset{A3}{=} \mathbb{P}(A_1) + \mathbb{P}(A_2)$$

Theorem 4: Seien $A, B \in E$ zwei beliebige Ereignisse. Dann gilt

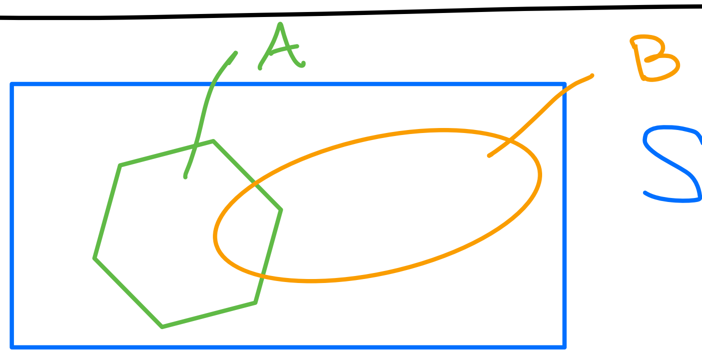
$$\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$$

Bew:



Thm 5: Seien A, B beliebige Ereignisse

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



Thm: (Monotonie)

Seien $A, B \in \mathcal{E}$ Ereignisse so dass

$A \subseteq B$. Dann gilt

$$\mathbb{P}(A) \leq \mathbb{P}(B)$$

→ Bew:

1.5. Wichtige Regeln der Wahrscheinlichkeitsrechnung

Theorem 1

Die Wahrscheinlichkeit des zu A komplementären Ereignisses ist stets

$$P(\bar{A}) = 1 - P(A), \text{ für jedes } A \in E$$

Theorem 2

Das unmögliche Ereignis hat die Wahrscheinlichkeit Null: $P(\emptyset) = 0$

Theorem 3

Sind die Ereignisse $A_1, A_2, \ldots, A_n \in E$ paarweise disjunkt, so ist die Wahrscheinlichkeit für das Ereignis, das aus der Vereinigung all dieser Ereignisse entsteht, gleich der Summe der Einzelwahrscheinlichkeiten. Das heisst

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

Theorem 4

Für eine Differenzmenge A\B gilt stets

$$P(A \setminus B) = P(A) - P(A \cap B).$$

Theorem 5 (Additionssatz)

Für zwei ***beliebige*** Ereignisse A und B aus E gilt

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

Theorem 6 (Monotonie Eigenschaft)

Impliziert das Ereignis A das Ereignis B, dann ist die Wahrscheinlichkeit von B niemals kleiner als die von A, das heisst

$$A \subset B \quad \Longrightarrow \quad P(A) \leq P(B)$$

Beispiel 1.5.1:

Wie gross ist die Wahrscheinlichkeit, dass eine zufällig gewählte dreistellige Zahl wiederholte Ziffern enthält? (Benützen Sie das Laplace-Modell)

$A = \{$3stellige Zahl enthält wiederholte Ziffern$\}$

$\bar{A} = \{$Zahl enthält keine wiederholte Ziffern$\}$

Idee: $\mathbb{P}(A) = 1 - \mathbb{P}(\bar{A}) = 1 - \frac{|\bar{A}|}{|S|}$

$|\bar{A}| = 10 \cdot 9 \cdot 8 = 720$

$|S| = 1000$

$\Rightarrow \mathbb{P}(A) = 1 - \frac{720}{1000} = \underline{\underline{0.28}}$

Beispiel 1.5.1:

Wie gross ist die Wahrscheinlichkeit, dass eine zufällig gewählte dreistellige Zahl wiederholte Ziffern enthält? (Benützen Sie das Laplace-Modell)

Hier ist offenbar

$$S = \{000, \dots, 999\}, \quad |S| = \#\ \text{Elemente von } S = 10^3$$

Universität St.Gallen

Betrachten wir das Ereignis

$$A = \{\text{Zahl enthält wiederholte Ziffer}\},$$

so ist

$$P(A) \overset{\text{Th.1}}{=} 1 - P(\overline{A}) = 1 - \frac{|\overline{A}|}{|S|}, \quad \text{und}$$

$$|\overline{A}| = \#\ \text{dreistellige Zahlen mit lauter}$$

$$\text{verschiedenen Ziffern} = 10 \cdot 9 \cdot 8 = 720$$

$$\Rightarrow P(A) = 1 - \frac{720}{1000} = 0.28$$

Kapitel 8.8

2. Vorlesung

S: Ereignisraum
E: Ereignismenge
$A \in E$ Ereignis

Bsp: Würfel $S = \{1,2,3,4,5,6\}$

$E = \{\emptyset, 1,2,3,4,5,6, \{1,2\}, \{1,3\}, \dots, \underbrace{\{1,2,3,4,5,6\}}_{S}\}$

Repetition:

Kolmogorov 1933

Definition

Eine jede Funktion P

$$P: \quad E \longrightarrow \mathbb{R}$$
$$A \longrightarrow P(A)$$

Axiom 1:	$P(A)$	$\geq$	$0, \quad \forall A \in E$
Axiom 2:	$P(S)$	$=$	1
Axiom 3:	$P(A \cup B)$	$=$	$P(A) + P(B)$, falls $A \cap B = \emptyset$

(Additionsregel für disjunkte Ereignisse)

$A \in E$

$A = \{1,2\}$

$\mathbb{P}(A)$

$\bar{A} = S \setminus A$

Theorem 1

Die Wahrscheinlichkeit des zu A komplementären Ereignisses ist stets

$$P(\bar{A}) = 1 - P(A), \text{ für jedes } A \in E$$

Theorem 2

Das unmögliche Ereignis hat die Wahrscheinlichkeit Null: $P(\emptyset) = 0$

$$\mathbb{P}(A) = 1 - \mathbb{P}(\bar{A})$$

Theorem 3

Sind die Ereignisse $A_1, A_2, \ldots, A_n \in E$ paarweise disjunkt, so ist die Wahrscheinlichkeit für das Ereignis, das aus der Vereinigung all dieser Ereignisse entsteht, gleich der Summe der Einzelwahrscheinlichkeiten. Das heisst

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

<u>Beweis:</u> IDEE Induktion über n

$\boxed{n=2}$ ✓ (A3) Induktionsverankerung

$\boxed{n \to n+1}$ Annahme: Aussage gilt für n

Zu zeigen: Aussage gilt auch für n+1

$$\mathbb{P}\left(\bigcup_{i=1}^{n+1} A_i\right) = \mathbb{P}\left(\underbrace{\left(\bigcup_{i=1}^{n} A_i\right)}_{=:B} \cup A_{n+1}\right)$$

$$\overset{(A3)}{=} \mathbb{P}(B) + \mathbb{P}(A_{n+1})$$

$$= \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) + \mathbb{P}(A_{n+1})$$

Induktionsannahme

$$= \sum_{i=1}^{n} \mathbb{P}(A_i) + \mathbb{P}(A_{n+1})$$

$$= \sum_{i=1}^{n+1} \mathbb{P}(A_i)$$

☒

Theorem 4

Für eine Differenzmenge A\B gilt stets

$$P(A \setminus B) = P(A) - P(A \cap B).$$

Theorem 5 (Additionssatz)

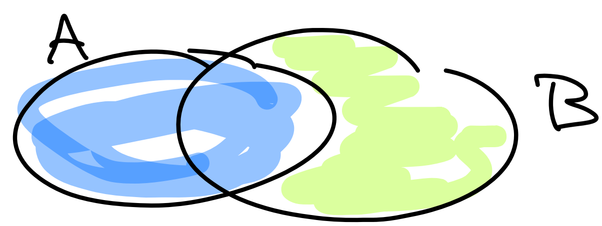
Für zwei ***beliebige*** Ereignisse A und B aus E gilt

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

Theorem 6 (Monotonie Eigenschaft)

Impliziert das Ereignis A das Ereignis B, dann ist die Wahrscheinlichkeit von B niemals kleiner als die von A, das heisst

$$A \subset B \longrightarrow P(A) \leq P(B)$$



Beweis: Theorem 5:

- $B = (B \setminus A) \cup (A \cap B)$

$$\Rightarrow \mathbb{P}(B) \overset{(A3)}{=} \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B) \qquad (1)$$

- $A \cup B = (B \setminus A) \cup A$

$$\Rightarrow \mathbb{P}(A \cup B) = \mathbb{P}(B \setminus A) + \mathbb{P}(A) \qquad (2)$$

$$\xRightarrow{(2) \text{ in } (1)} \mathbb{P}(B) = \mathbb{P}(A \cup B) - \mathbb{P}(A) + \mathbb{P}(A \cap B)$$

☒

1.6. Wahrscheinlichkeitsräume

S

Diskrete Wahrscheinlichkeitsräume

Definition (Diskreter Wahrscheinlichkeitsraum)

Ist ein Ereignisraum S endlich oder abzählbar unendlich, heisst er ***diskret***.

Sei nun P(•) ein Wahrscheinlichkeitsmass, das

funktion

den Axiomen genügt, dann wird jedem Elementar-ereignis e_i eine Wahrscheinlichkeit $p_i = P(e_i)$ zugeordnet. Also:

e_1	e_2	e_3	.....	e_i	.....
p_1	p_2	p_3	.....	p_i	.....

$$S = \{e_1, e_2, e_3, \ldots\ldots\}$$

$$P(e_i) =: p_i$$

Axiome

(1) $p_i \geq 0$ für jedes $i = 1,2,...$

(2) $\sum_{\text{alle } i} p_i = 1$, weil $e_i \cap e_j = \emptyset \quad \forall\, i \neq j$

(paarweise disjunkt) und $\bigcup_{\text{alle } i} e_i = S$

$$\sum_{\text{alle } i} p_i \overset{\text{Th. 3}}{=} P\left(\bigcup_{\text{alle } i} e_i\right) = P(S) = 1$$

(3) $\boxed{P(A) = \sum_{e_i \in A} p_i}$ $\quad i : e_i \in A$

d.h. die Wahrscheinlichkeit für jedes Ereignis $A \in E$ berechnet sich dann aus der Summe der Wahrscheinlichkeiten derjenigen Elementarereignisse, die in A enthalten sind.

Beispiel 1.6.1:

1. Gleichwahrscheinlichkeitsraum mit m Elementarereignissen (Laplace-Experiment):

$$S = \{e_1, e_2, \ldots, e_m\}$$

$$p_i = \mathbb{P}(e_i) = \frac{1}{m}$$

$$\sum_{i=1}^{m} p_i = \sum_{i=1}^{m} \frac{1}{m} = \frac{1}{m} \sum_{i=1}^{m} 1 = \frac{m}{m} = 1$$

2. (abzählbar unendlicher Ereignisraum)

Zufallsexperiment: “Man werfe eine Münze solange, bis 'Kopf' erscheint“.

$$S = \{ \underset{e_1}{K}, \underset{e_2}{ZK}, \underset{e_3}{ZZK}, \underset{e_4}{ZZZK}, \ldots$$

$$e_i = \underbrace{Z \cdot \ldots \cdot Z}_{(i-1)\,\text{mal}} \cdot K \quad \text{für } i = 1, 2, \ldots$$

$$p_1 = \underbrace{P(e_1)}_{P(K)} = 1/2 \quad , \quad p_2 = P(ZK) = \left(\tfrac{1}{2}\right)^2$$

$$p_3 = P(ZZK) = \left(\tfrac{1}{2}\right)^3$$

$$\Longrightarrow p_i = P(e_i) = \left(\tfrac{1}{2}\right)^i = \frac{1}{2^i} \qquad 0 \leq q < 1$$

$$\sum_{i=1}^{\infty} p_i = \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^i = 1 \qquad \sum_{i=0}^{\infty} q^i = \frac{1}{1-q} = 2$$

$$\underbrace{q^0}_{1} + \sum_{i=1}^{\infty} q^i \qquad \boxed{q = \tfrac{1}{2}}$$

$$\Longrightarrow \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^i = 1$$

Beispiel 1.6.1:

1. Gleichwahrscheinlichkeitsraum mit m Elementarereignissen (Laplace-Experiment):

$$e_1, \ldots, e_m \quad ; \quad p_i = {}^1\!/_m \quad \forall\ i = 1, \ldots, m$$

$$\Rightarrow \sum_{i=1}^{m} p_i = m \cdot {}^1\!/_m = 1$$

2. (abzählbar unendlicher Ereignisraum)

Zufallsexperiment: “Man werfe eine Münze solange, bis 'Kopf' erscheint“.

$$S = \{\underset{e_1}{K}, \underset{e_2}{ZK}, \underset{e_3}{ZZK}, \underset{e_4}{ZZZK}, \underset{e_5}{ZZZZK}, \underset{\ldots\ldots}{\ldots}\}$$

$$p_1 = P(K) = 1/2,\ p_2 = P(ZK) = 1/4,\ p_3 = 1/8, \ldots$$

$$p_i = P(\underbrace{ZZ\ldots ZK}_{i}) = \frac{1}{2^i}$$

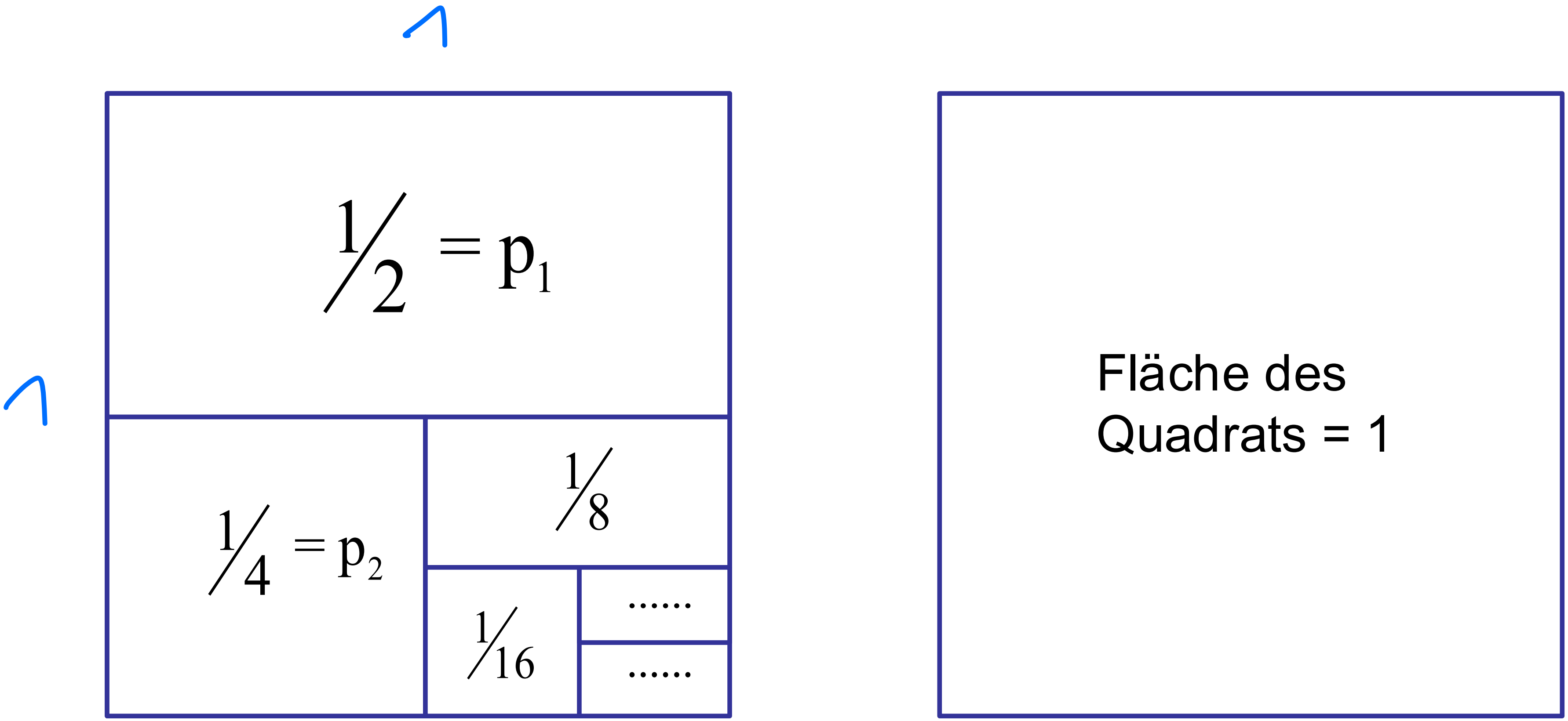
$$\sum_{i=0}^{\infty} q^i = \frac{1}{1-q}$$

$$= 2$$

$$\Rightarrow \sum_{\text{alle } i} p_i = \sum_{i=1}^{\infty} p_i = \sum_{i=1}^{\infty} \frac{1}{2^i} = \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^i = 1,$$ unendliche Reihe,

die aber endlich bleibt und konvergiert!

Beweis durch Eigenschaften der Reihen
oder Visualisierung der Konvergenz:



Stetige Wahrscheinlichkeitsräume

Definition (Stetiger Wahrscheinlichkeitsraum)

Ist ein Ereignisraum S nicht abzählbar, heisst er ***stetig***.

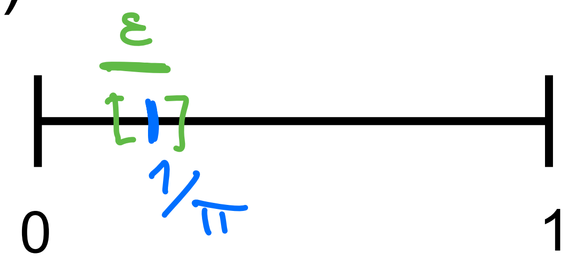
Bsp. $S = \mathbb{R}$

oder $S = [a, b]$, $a < b$

Universität St.Gallen

Beispiel 1.6.2:

Man denke an ein Geradenstück zwischen Null und Eins (abgeschlossen).



Auf diesem Geradenstück befinden sich Punkte, die einzeln betrachtet keinerlei Ausdehnung und damit die Länge Null haben. Allerdings sind es unendlich viele Punkte.

Nun denken wir uns ein Zufallsexperiment, welches darin besteht, dass aus dem abgeschlossenen Intervall [0,1] der reellen Zahlen eine Zahl $0 \leq a \leq 1$ zufällig gewählt wird.

In jedem Fall ist dann $P(a)=0 \quad \forall a \in S=[0,1]$.

Definieren wir drei Teilintervalle als Ereignisse:

$$A = \{a \mid a < 0.4\}$$

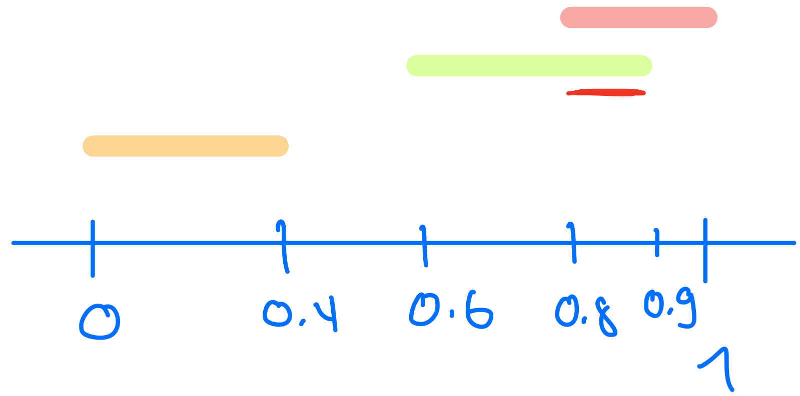
$$B = \{a \mid 0.6 < a < 0.9\}$$

$$C = \{a \mid a > 0.8\}$$

$A = \{a \mid a < 0.4\}$

$B = \{a \mid 0.6 < a \le 0.9\}$

$C = \{a \mid a > 0.8\}$



"Intuitiv:"

$$\left.\begin{aligned} \mathbb{P}(A) &= 0.4 \\ \mathbb{P}(B) &= 0.3 \\ \mathbb{P}(C) &= 0.2 \end{aligned}\right\}$$

Laplac'sche stetige Gleichwahrscheinlichkeit

$$\mathbb{P}(A) = \frac{\text{"Länge von } A\text{"}}{\text{"Länge von } S\text{"}}$$

$\forall$ = für alle

$\exists$ = es existiert

$\exists!$ = es existiert ein eindeutiger

$$\mathbb{P}(B \cap C) = \frac{\text{Länge von } (B \cap C)}{\text{Länge von } S} = \frac{0.1}{1}$$

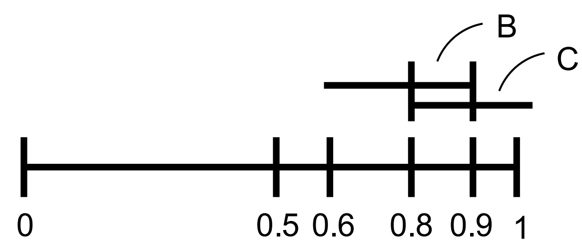
$$= 0.1$$

Intuitiv:

$P(A) = 0.4;\ P(B) = 0.3$ und $P(C) = 0.2.$ Warum?

$\rightarrow$ S ist ein Laplacescher stetiger Gleichwahrscheinlichkeitsraum: $P(A) = \dfrac{\text{Länge A}}{\text{Länge S}}$;

$$P[B \cap C] = \frac{\text{Länge } B \cap C}{\text{Länge S}} = \frac{0.1}{1} = 0.1$$



$$P[B \cup C] \overset{\text{Th.5}}{=} P(B) + P(C) - P[B \cap C] = 0.3 + 0.2 - 0.1 = 0.4$$

Bemerkung:

Die Intervalle können offen oder abgeschlossen sein. Es gilt:

$$B^+ = \{a \mid 0.6 \leq a \leq 0.9\} \rightarrow P\left(B^+\right) = 0.3 = P\left(B\right),$$

auch wenn B^+ zwei Elementarereignisse mehr hat;
aber die Wahrscheinlichkeit beider Randpunkte ist “vom Mass Null“.

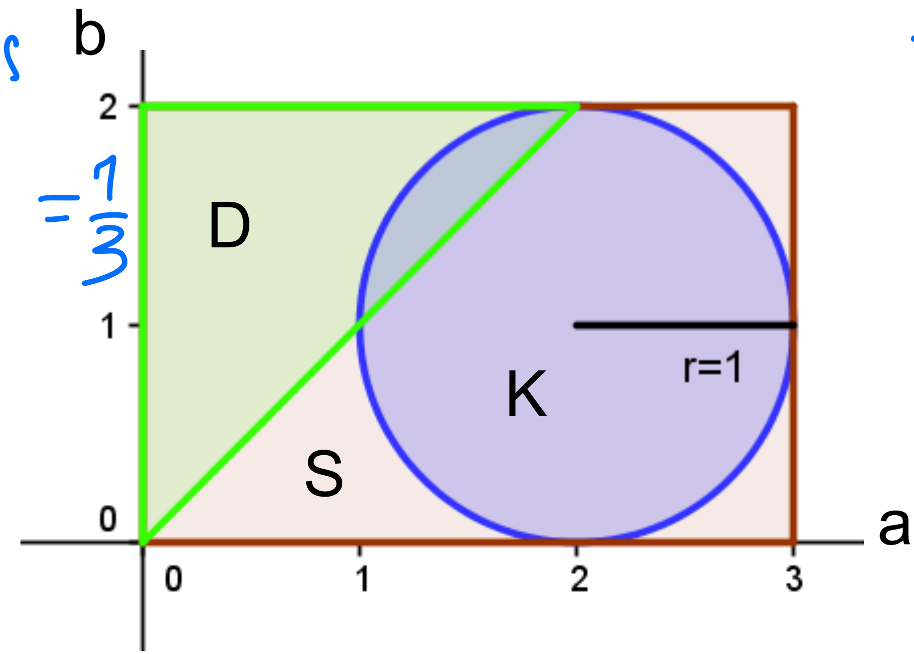
2. Definiere $S = \{(a,b) \mid 0 < a < 3 \text{ und } 0 < b < 2\}$

$$D = \{(a,b) \mid 0 < a < b < 2\}$$

$$K = \{(a,b) \mid (a-2)^2 + (b-1)^2 < 1\}$$

(Im Rechteck einbeschriebener Kreis)

$$\mathbb{P}(D) = \frac{\text{Fläche von } D}{\text{Fläche von } S} = \frac{2}{6} = \frac{1}{3}$$



$$\mathbb{P}(K) = \frac{\pi}{6} = 0.52$$

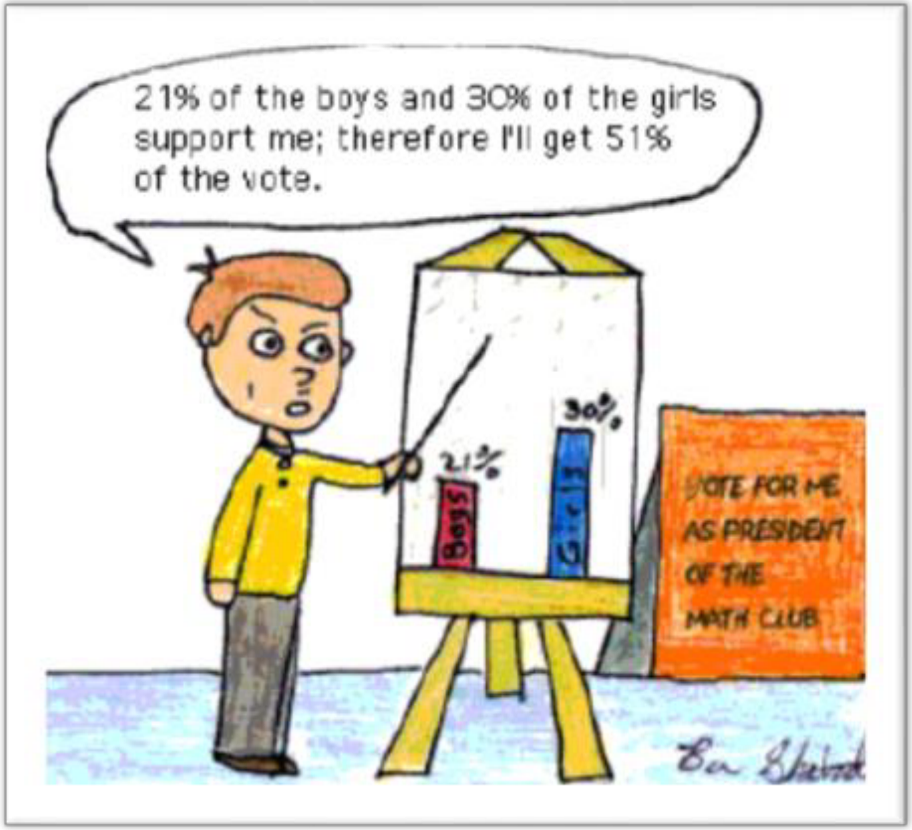
$$P(D) = \frac{\text{Fläche von D}}{\text{Fläche von S}} = \frac{2}{6} = \frac{1}{3}$$

$$P(K) = \frac{\text{Fläche von K}}{\text{Fläche von S}} = \frac{\pi}{6} \cong 0.5236$$

Universität St.Gallen

Kapitel 8.9

1.7. Bedingte Wahrscheinlichkeit und stochastische Unabhängigkeit



Beispiel 1.7.1:

Einfacher Würfelwurf:

Wie gross ist die Wahrscheinlichkeit eine “6“ zu bekommen?

$$\rightarrow P(\text{"6"}) = \frac{1}{6}$$

Und wenn wir erfahren, dass die gewürfelte Augenzahl eine gerade Zahl ist?

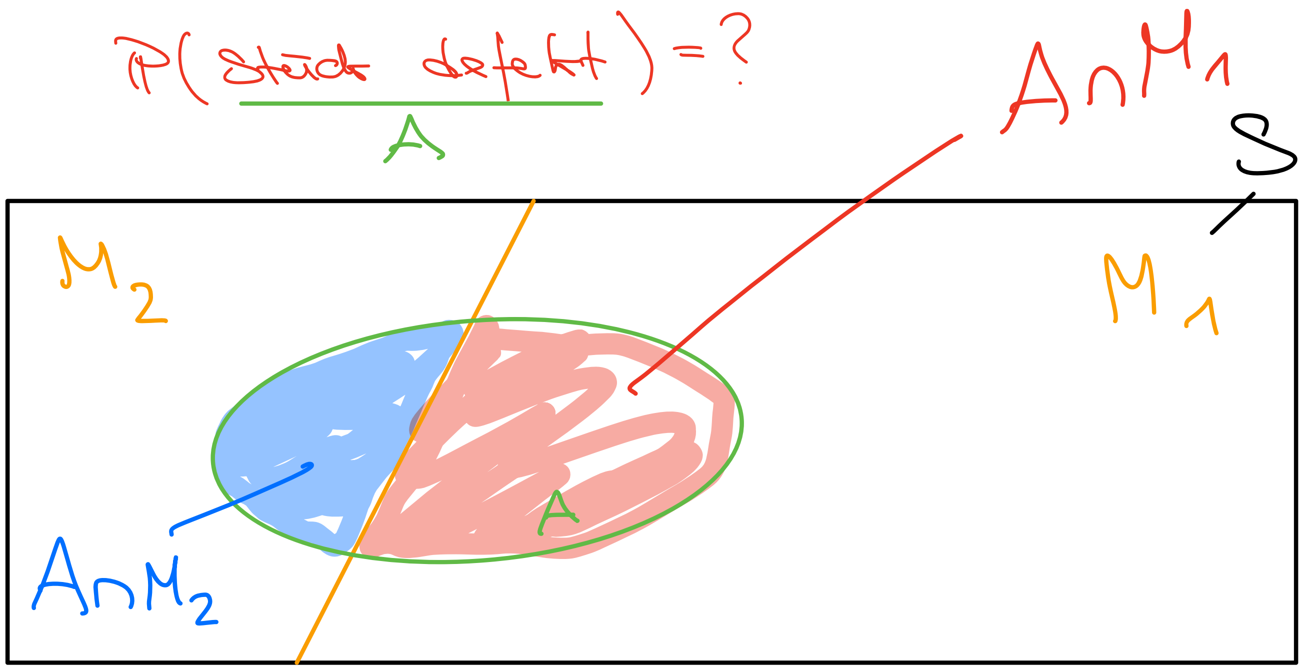
$$P(\text{"6"} \mid \text{Augenzahl ist gerade}) = \frac{1}{3}$$

Universität St.Gallen

Maschine	Produktion	Ausschuss (defekte Ware)
M_1	$2/3$	10%
M_2	$1/3$	7%

- $\mathbb{P}(\text{Stück defekt} \mid M_1) = 0.1$
- $\mathbb{P}(\text{Stück defekt} \mid M_2) = 0.07$
- $\mathbb{P}(\text{Stück auf } M_1 \text{ produziert}) = 2/3$
- $\mathbb{P}(\text{ " } M_2 \text{ " }) = 1/3$

$$\mathbb{P}(\underbrace{\text{Stück defekt}}_{A}) = ?$$



$$\mathbb{P}(A \cap M_i) = \mathbb{P}(A \mid M_i) \cdot \mathbb{P}(M_i) \quad , i = 1,2$$

$$\underset{\text{Axiom 3}}{\Longrightarrow} \quad \mathbb{P}(A) = \mathbb{P}(A \cap M_1) + \mathbb{P}(A \cap M_2)$$

Definition (Bedinge Wahrscheinlichkeit)

Seien A, B zwei Ereignisse eines gegebenen Wahrscheinlichkeitsraums. Die <u>bedingte Wahrscheinlichkeit</u> A unter der Bedingung B ist definiert als

$$\mathbb{P}(A \mid B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \quad \text{unter der Annahme } \mathbb{P}(B) > 0$$

<u>Bem</u>: Für $\mathbb{P}(B) = 0$ ist $\mathbb{P}(A \mid B)$ nicht definiert!

Definition (Bedinge Wahrscheinlichkeit)

Seien A und B zwei Ereignisse eines gegebenen Wahrscheinlichkeitsraums. Die bedingte Wahrscheinlichkeit A unter der Bedingung B ist definiert als

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ für } P(B) > 0$$

und undefiniert für $P(B)=0$

Universität St.Gallen

Beispiel 1.7.2:

Ein roter und ein grüner Würfel werden geworfen.
Wie gross ist die Laplacesche Wahrscheinlichkeit, dass mindestens einer der beiden Würfel eine Sechs zeigt unter der Bedingung, dass die Augensumme (strikt) grösser als neun ist?

$A =$ "mindestens eine 6"

$B =$ "Augensumme > 9"

<u>Gesucht</u>: $\mathbb{P}(A \mid B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$

(1,1)	(1,2)	......	......	......	(1,6)
(2,1)	(2,2)	......	......	......	(2,6)
(3,1)	(3,2)	......	......	......	(3,6)
(4,1)	(4,2)	......	......	(4,5)	(4,6)
(5,1)	(5,2)	......	......	(5,5)	(5,6)
(6,1)	(6,2)	......	(6,4)	(6,5)	(6,6)

B

$A =$ "mindestens eine 6"

$B =$ "Augensumme > 9"

<u>Gesucht:</u> $\mathbb{P}(A \mid B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$

$\mathbb{P}(A) = \dfrac{11}{36}$

$\mathbb{P}(A \cap B) = \dfrac{5}{36}$

$\mathbb{P}(B) = \dfrac{6}{36} = \dfrac{1}{6}$

$\Longrightarrow \mathbb{P}(A \mid B) = \dfrac{5/36}{6/36} = \underline{\underline{5/6}}$

$$
\begin{array}{cccccc}
(1,1) & (1,2) & \ldots\ldots & \ldots\ldots & \ldots\ldots & (1,6) \\
(2,1) & (2,2) & \ldots\ldots & \ldots\ldots & \ldots\ldots & (2,6) \\
(3,1) & (3,2) & \ldots\ldots & \ldots\ldots & \ldots\ldots & (3,6) \\
(4,1) & (4,2) & \ldots\ldots & \ldots\ldots & (4,5) & (4,6) \\
(5,1) & (5,2) & \ldots\ldots & \ldots\ldots & (5,5) & (5,6) \\
(6,1) & (6,2) & \ldots\ldots & (6,4) & (6,5) & (6,6)
\end{array}
$$

$$A = \text{"mindestens eine Sechs"}: P(A) = \frac{11}{36}$$

$$B = \text{"Augensumme} > 9\text{"}: P(B) = \frac{6}{36} \text{ und } P(A \cap B) = \frac{5}{36}$$

$$\Rightarrow P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{5/36}{6/36} = \frac{5}{6}$$

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

Theorem 7 (Multiplikationssatz)

Die Wahrscheinlichkeit, dass zwei Ereignisse A und B gleichzeitig eintreten, beträgt

$$\underline{P(A \cap B)} = P(A) \cdot P(B \mid A)$$

oder

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

Beispiel 1.7.3:

Eine Urne enthält 4 Kugeln, nämlich 3 rote und 1 blaue. Wie gross ist die Wahrscheinlichkeit, beim Ziehen von 2 Kugeln (ohne Zurücklegen), 2 rote zu erwischen?

$$\mathbb{P}(R_2 | R_1) = \frac{\mathbb{P}(R_2 \cap R_1)}{\mathbb{P}(R_1)}$$

3R 1B

R_i = i-the gezogene Kugel ist rot

$i = 1, 2$

$$\underbrace{\mathbb{P}(R_1 \cap R_2)}_{\text{gesucht}} = \underbrace{\mathbb{P}(R_2 | R_1)}_{2/3} \cdot \underbrace{\mathbb{P}(R_1)}_{3/4}$$

$$= 2/4 = 1/2$$

Beispiel 1.7.3:

Eine Urne enthält 4 Kugeln, nämlich 3 rote und 1 blaue. Wie gross ist die Wahrscheinlichkeit, beim Ziehen von 2 Kugeln (ohne Zurücklegen), 2 rote zu erwischen?

Sei R_i = “i-te gezogene Kugel ist rot“ für i=1, 2. Dann ist

$$P\left[R_1 \cap R_2\right] = P\left[R_1\right] \cdot P\left[R_2 \middle| R_1\right] = \frac{\cancel{3}}{\cancel{4}_2} \cdot \frac{\cancel{2}}{\cancel{3}} = {}^1\!/_2 .$$

Theorem 8 (Multiplikationssatz für unabhängige Ereignisse)

Sind zwei Ereignisse A und B unabhängig, dann ist

$$P(A \cap B) = P(A) \cdot P(B).$$

Bew: $\underline{A, B \text{ unabh. } \quad \mathbb{P}(A|B) = \mathbb{P}(A)}$

$$\underbrace{\mathbb{P}(A|B)}_{\mathbb{P}(A)} \overset{\text{Formel}}{=} \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \Longleftrightarrow \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A \cap B)$$

Universität St.Gallen

Beispiel 1.7.4: Zeigen Sie A, B sind unabhängig

Eine Münze wird zweimal geworfen.

$S = \{ZZ, ZK, KZ, KK\}$, Laplace-Modell.

Für die Ergebnisse: $A = \{\text{Kopf beim 1. Wurf}\} = \{KK, KZ\}$

$B = \{\text{Kopf beim 2. Wurf}\} = \{KK, ZK\}$

$\mathbb{P}(A) = 1/2$, $\mathbb{P}(B) = 1/2$

$(A \cap B) = \{KK\} \Rightarrow \mathbb{P}(A \cap B) = 1/4$

$\Rightarrow \underbrace{\mathbb{P}(A \cap B)}_{1/2} = \underbrace{\mathbb{P}(A)}_{1/2} \cdot \underbrace{\mathbb{P}(B)}_{1/2} \Rightarrow$ A, B unabh.

$\boxtimes$

Beispiel 1.7.4:

Eine Münze wird zweimal geworfen.

$S=\{ZZ, ZK, KZ, KK\}$, Laplace-Modell.

Für die Ergebnisse: $A=\{\text{Kopf beim 1. Wurf}\} = \{KK, KZ\}$

$B=\{\text{Kopf beim 2. Wurf}\} = \{KK, ZK\}$

ist $P(A)=P(B)=\frac{2}{4} = 1/2$ und

$P(A \cap B) = P(\{KK\}) = 1/4 = P(A) \cdot P(B)$.

Also sind A und B unabhängig (wie erwartet).

Universität St.Gallen

A, B unabhängig
B, C unabhängig

Merke:

Stochastische Unabhängigkeit ist keine transitive Relation!

Aus "A und B unabhängig" und "B und C unabhängig" folgt ***nicht*** "A und C unabhängig"!

$\Longrightarrow$ A, C nicht unabhängig

Beispiel:

Man würfle mit zwei Würfeln. Sei A das Ereignis *"1,2 oder 3 beim 1. Würfel"*, B *"4, 5 oder 6 beim 2. Würfel"* und C *"4,5 oder 6 beim 1. Würfel"*. Offensichtlich ist A und B sowie B und C unabhängig, *keinesfalls* aber ist A und C unabhängig.

Universität St.Gallen

Teil I: Wahrscheinlichkeitsrechnung

1. Grundlagen der Wahrscheinlichkeit
   1.1. Ereignisse, Ereignisraum und Ereignismenge
   1.2. Das Rechnen mit Ereignissen
   1.3. Der Wahrscheinlichkeitsbegriff
   1.4. Axiomatik der Wahrscheinlichkeitstheorie
   1.5. Wichtige Regeln der Wahrscheinlichkeitsrechnung
   1.6. Wahrscheinlichkeitsräume
   1.7. Bedingte Wahrscheinlichkeit und stochastische Unabhängigkeit
   1.8. Totale Wahrscheinlichkeit
   1.9. Das Bayes-Theorem

2. Elementare Kombinatorik
   2.1. Fakultät und Binomialkoeffizient
   2.2. Das Fundamentalprinzip der Kombinatorik
   2.3. Permutationen
   2.4. Kombinationen
   2.5. Variationen

3. Zufallsvariablen
   3.1. Die Verteilungsfunktion
   3.2. Diskrete Zufallsvariablen
   3.3. Stetige Zufallsvariablen
   3.4. Erwartungswerte von Zufallsvariablen
   3.5. Varianz
   3.6. Standardisieren

4. Stochastische Modelle und spezielle Verteilungen
   4.1. Gleichförmige Verteilung (diskret)
   4.2. Bernoulli-Verteilung (diskret)
   4.3. Binomialverteilung (diskret)
   4.4. Poisson-Verteilung (diskret)
   4.5. Rechteckverteilung (stetig)
   4.6. Exponentialverteilung (stetig)
   4.7. Normalverteilung (stetig)

5. Mehrdimensionale Zufallsvariablen
   5.1. Gemeinsame Verteilung und Randverteilungen
   5.2. Bedingte Verteilungen und stochastische Unabhängigkeit
   5.3. Kovarianz und Korrelationskoeffizient
   5.4. Summe von zwei oder mehreren Zufallsvariablen
6. Der zentrale Grenzwertsatz

Teil II: Statistik

7. Beschreibende/Deskriptive Statistik
   7.1. Häufigkeitsverteilung, Histogramm und Verteilungsfunktion
   7.2. Messzahlen zur Beschreibung statistischer Verteilungen
   7.3. Boxplot
   7.4. Quantile-Quantile Plot
   7.5. Streudiagramm

8. Schätzung unbekannter Parameter: Punktschätzung
   8.1. Intuitiv heuristische Ansätze für Schätzfunktionen
   8.2. Eigenschaften von Punktschätzungen
   8.3. Methoden zur Konstruktion von Schätzfunktionen

9. Intervallschätzungen-Konfidenzintervalle
   9.1. Konzept des Konfidenzintervalls
   9.2. Ableitung von Konfidenzintervallen (bei grossen Stichproben)
   9.3 Zusammenhang mit Hypothesentests

10. Hypothesentests
   10.1 Arten von Hypothesen
   10.2 Kritischer Bereich und Teststatistik
   10.3 Gütefunktion und Arten von Fehlern
   10.4 Der p-Wert

Teil III: Aufgaben

Universität St.Gallen

Kapitel 8.10

Bisher:

Ereignisraum, Ereignismenge, Wahrscheinlichkeits-funktion

Wahrscheinlichkeitsraum $(S, E, \mathbb{P})$

- $E \subset S$, so dass $\emptyset \in E$, $S \in E$
- $\mathbb{P}: E \to [0,1]$
- Ereignis $A \in E$ (Bsp $A = \{2,4,6\}$) $\mathbb{P}(A) = 1/2$

$\mathbb{P}(A)$: die W'keit dass Ereignis A eintritt

Bedingte Wahrscheinlichkeit:

$A, B \in E$ beliebige Ereignisse

$$\mathbb{P}(A|B) := \frac{\overbrace{\mathbb{P}(A \cap B)}^{\mathbb{P}(A,B)}}{\mathbb{P}(B)}, \quad \text{falls } \mathbb{P}(B) > 0$$

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

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

$\mathbb{P}(A), \mathbb{P}(B) > 0$

$$\Longrightarrow \mathbb{P}(A|B)\,\mathbb{P}(B) = \mathbb{P}(B|A)\,\mathbb{P}(A)$$

$$\Longrightarrow \boxed{\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\,\mathbb{P}(A)}{\mathbb{P}(B)}}$$

Satzes von Bayes

$\Longrightarrow$ bed. W'keit

$$\mathbb{P}(A \cap B) = \underbrace{\mathbb{P}(A|B)}_{\mathbb{P}(A)}\mathbb{P}(B)$$

$\Longleftarrow$ bed. W'keit

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A,B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)\cancel{\mathbb{P}(B)}}{\cancel{\mathbb{P}(B)}} = \mathbb{P}(A)$$

Zwei Ereignisse $A, B \in E$ heissen unabhängig, falls

$$\mathbb{P}(A|B) = \mathbb{P}(A)$$

$$\Longleftrightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$$

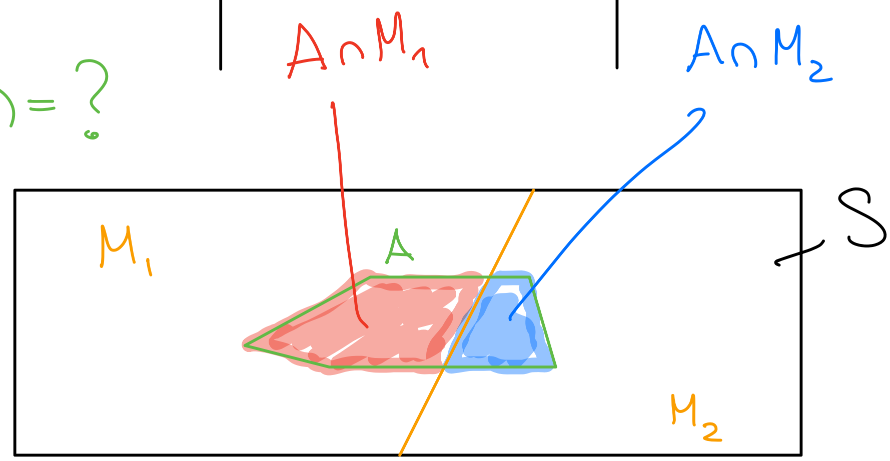
1.8. Totale Wahrscheinlichkeit

Beispiel 1.8.1:

Ein und derselbe Massenartikel werde auf zwei Maschinen gefertigt. Die schnellere Maschine M_1 hinterlässt 10% Ausschuss, produziert aber doppelt soviel wie die langsamere Maschine M_2, die aber nur einen Ausschussanteil von 7% aufweist. Wie gross ist die Wahrscheinlichkeit, dass ein zufällig aus der Gesamtproduktion gezogenes Einzelstück defekt ist?

Maschine	Produktion	Fehler W'keit
M_1	$2/3$	0.1
M_2	$1/3$	0.07

$\mathbb{P}(A) = ?$



A: Stück defekt

M_1: Stück wurde von Maschine 1 gefertigt

M_2: Stück wurde von Maschine 2 gefertigt

$$\Rightarrow \mathbb{P}(A) \overset{13}{=} \mathbb{P}(A \cap M_1) + \mathbb{P}(A \cap M_2)$$

$$= \underbrace{\mathbb{P}(A \mid M_1)}_{0.1}\underbrace{\mathbb{P}(M_1)}_{2/3} + \underbrace{\mathbb{P}(A \mid M_2)}_{0.07}\underbrace{\mathbb{P}(M_2)}_{1/3}$$

$$= \dots = 0.09$$

Gegeben sind also:

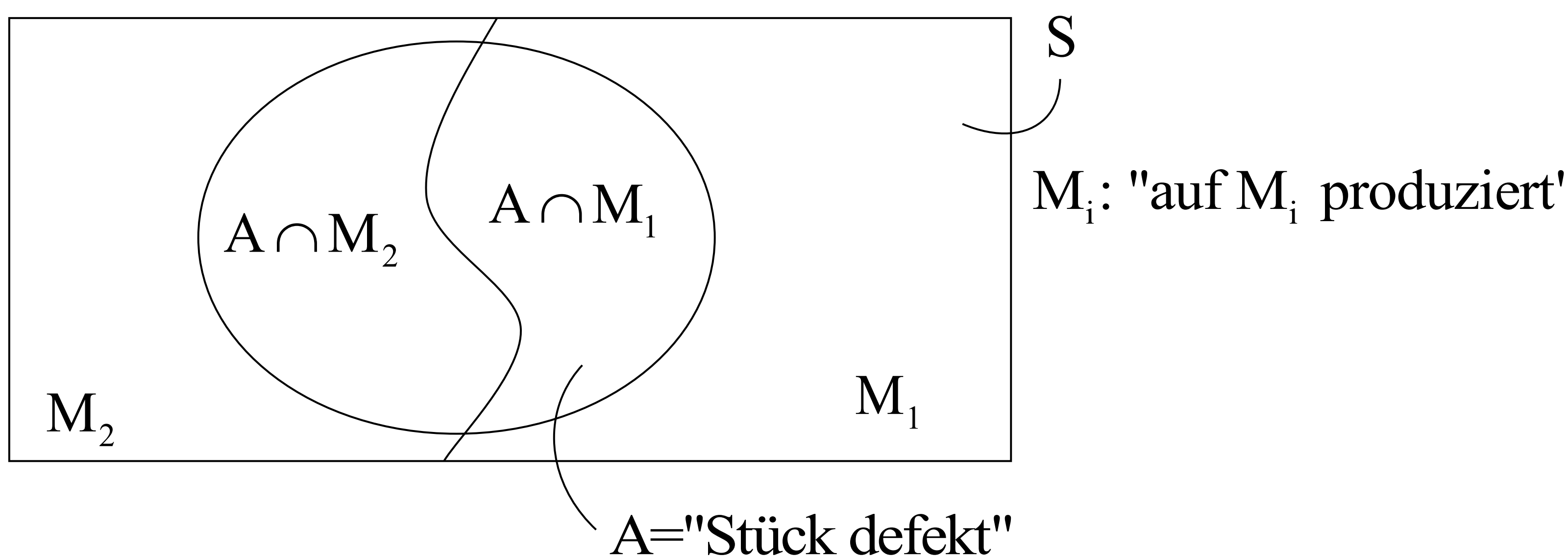
Maschine	Produktion	Ausschuss
M_1	2/3	10%
M_2	1/3	7%

$P(\text{Stück defekt} \mid M_1) = 0.1$; $P(\text{Stück defekt} \mid M_2) = 0.07$

$P(\text{Stück auf } M_1 \text{ produziert}) = 2/3$

$P(\text{Stück auf } M_2 \text{ produziert}) = 1/3$

$\rightarrow P(\text{Stück defekt}) = ?$



Wir haben:

$$P(A \cap M_i) = P(A|M_i) \cdot P(M_i), \quad i = 1,2$$

Nach Axiom 3:

$$\begin{aligned} P(A) &= P(A \cap M_1) + P(A \cap M_2) \\ &= P(A|M_1) \cdot P(M_1) + P(A|M_2) \cdot P(M_2) \\ &= 0.1 \cdot \tfrac{2}{3} + 0.07 \cdot \tfrac{1}{3} = 0.09 \end{aligned}$$

Universität St.Gallen

disjunkt

Definition (Partition)

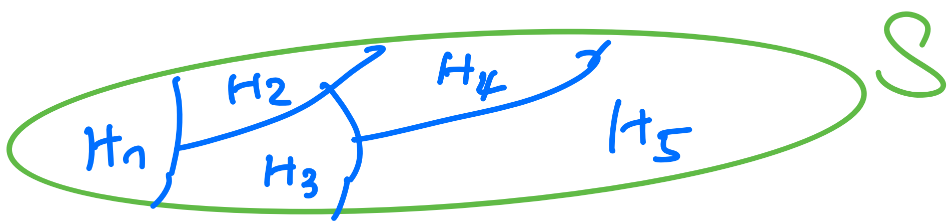
Beliebige n Ereignisse $H_1, H_2, \ldots, H_n$, die sich gegenseitig ausschliessen, aber zusammengenommen den Ereignisraum ganz ausfüllen, heissen eine ***Aufteilung*** oder ***Partition*** von S.

Eine Partition $H_1, H_2, \ldots, H_n$ erfüllt also

$$H_i \cap H_j = \emptyset \quad \text{für} \quad i \neq j$$

und

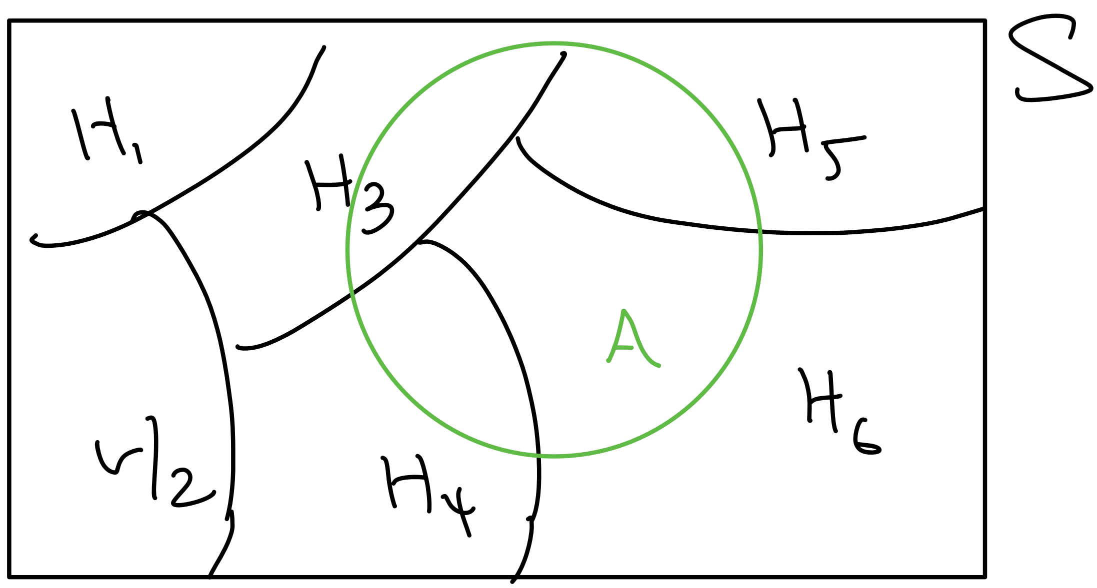
$$\bigcup_{i=1}^{n} H_i = H_1 \cup H_2 \cup \ldots \cup H_n = S.$$



Theorem: (Totale Wahrscheinlichkeit)

Seien $H_1, H_2, \dots, H_n$ eine Partition von S.
Dann gilt für jedes Ereignis $A \in E$

$$\mathbb{P}(A) = \sum_{i=1}^{n} \underbrace{\mathbb{P}(A \mid H_i)\,\mathbb{P}(H_i)}_{\mathbb{P}(A \cap H_i)}$$



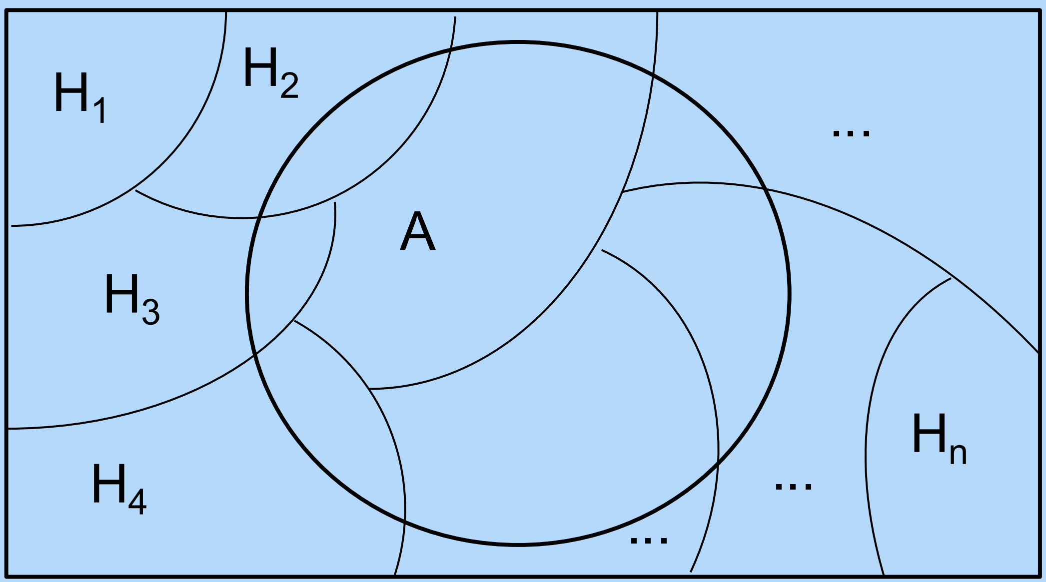
Theorem 6 (Totale Wahrscheinlichkeit)

Partition

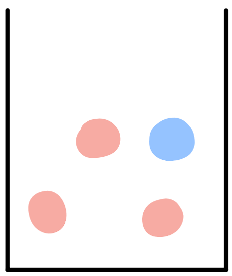
Seien $H_1, H_2, \ldots, H_n$ eine Aufteilung von S,
Dann gilt für jedes Ereignis $A \in E$

$$P(A) = \sum_{j=1}^{n} P(A \mid H_j) \cdot P(H_j)$$

$\mathbb{P}(A \cap H_j)$



Beispiel 1.8.2:



Ziehen ohne Zurücklegen:

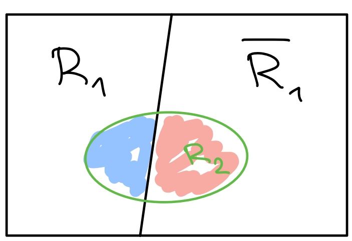
$$\mathbb{P}(\underbrace{\text{zweite gezogene Kugel ist rot}}_{=:R_2})$$

R_1: erste gezogene Kugel ist rot

$\overline{R_1}$: erste gezogene Kugel blau

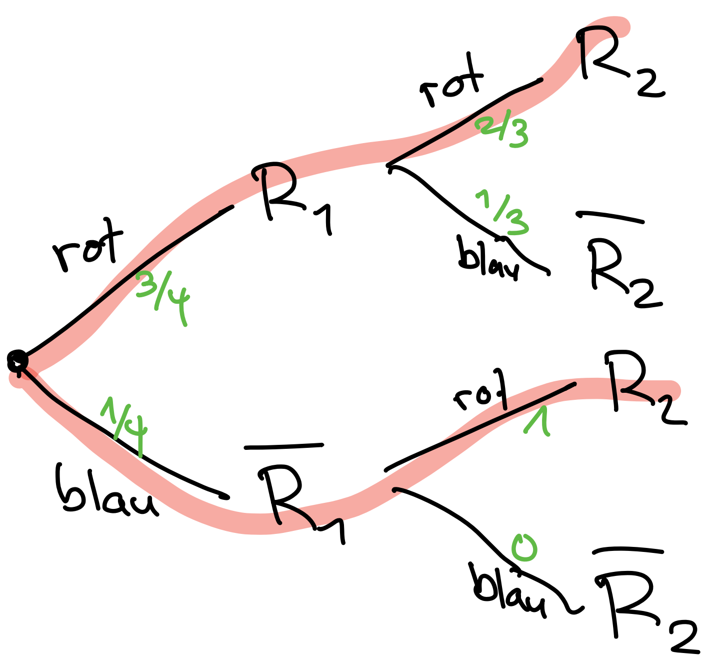
$$\mathbb{P}(R_2) \overset{\text{Tot. Wkeit}}{=} \mathbb{P}(R_2 \cap R_1) + \mathbb{P}(R_2 \cap \overline{R_1})$$

$$= \underbrace{\mathbb{P}(R_2 | R_1)}_{2/3} \underbrace{\mathbb{P}(R_1)}_{3/4} + \underbrace{\mathbb{P}(R_2 | \overline{R_1})}_{1} \underbrace{\mathbb{P}(\overline{R_1})}_{1/4}$$



$$= \frac{1}{2} + \frac{1}{4} = \underline{\underline{\frac{3}{4}}}$$

$$\frac{3}{4} \cdot \frac{2}{3} + \frac{1}{4} \cdot 1$$



Beispiel 1.8.2:

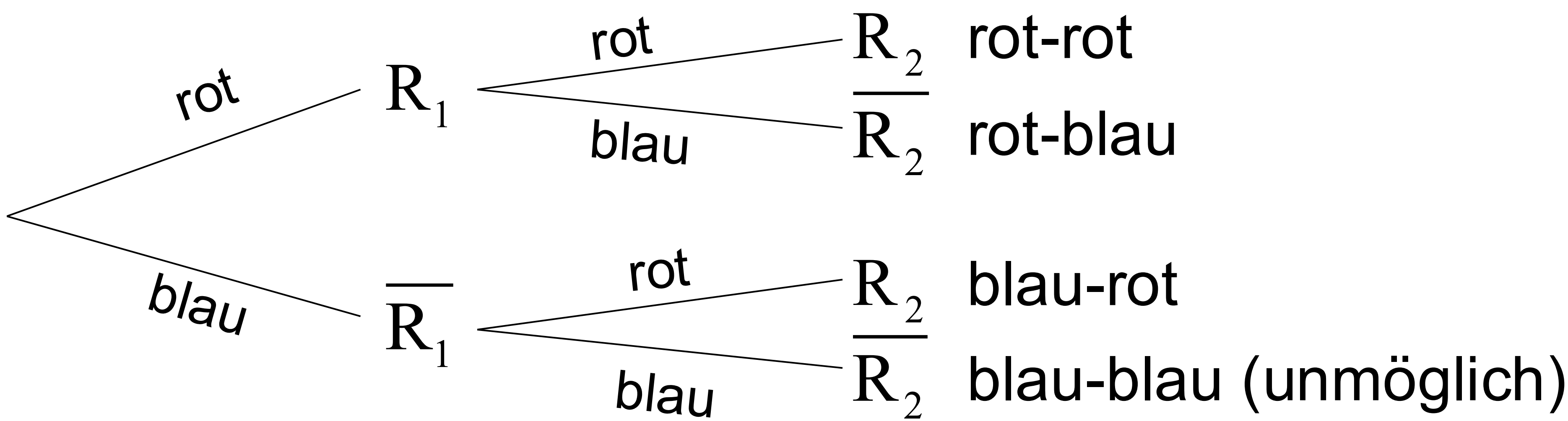
In Beispiel 1.7.3, der Urne mit drei roten und einer blauen Kugel bilden R_1 = {erste gezogene Kugel ist rot} und $\overline{R_1}$ = {erste gezogene Kugel ist blau} offenbar eine Aufteilung von S. Für R_2 = {zweite gezogene Kugel ist rot} erhalten wir also:

$$P(R_2) = P(R_2 | R_1) \cdot P(R_1) + P(R_2 | \overline{R_1}) \cdot P(\overline{R_1})$$

$$= \frac{2}{3} \cdot \frac{3}{4} + \frac{3}{3} \cdot \frac{1}{4} = 3/4$$

Universität St.Gallen

Bemerkung: Überlegungen wie im Satz der totalen Wahrscheinlichkeit werden oft in einem ***Baumdiagramm*** dargestellt.



Universität St.Gallen

Kapitel 8.11

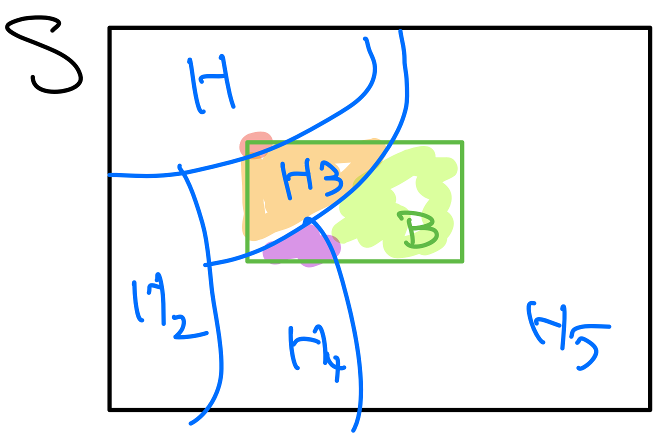
1.9. Das Bayes-Theorem

Bayes Theorem: Seien $H_1, \dots, H_n$ irgendeine Partition von S. Sei B ein Ereignis ($B \in E$) so dass $\mathbb{P}(B) > 0$. Dann gilt für jedes H_i

$$\mathbb{P}(H_i | B) = \frac{\mathbb{P}(B|H_i)\,\mathbb{P}(H_i)}{\sum_{j=1}^{n} \mathbb{P}(B|H_j)\,\mathbb{P}(H_j)}$$

"Beweis:"

$$\text{Nenner: } \sum_{j=1}^{n} \underbrace{\mathbb{P}(B|H_j)\,\mathbb{P}(H_j)}_{\mathbb{P}(B \cap H_j)} \overset{\text{Satz der tot. W'keit}}{=} \mathbb{P}(B)$$



$$\mathbb{P}(H_i|B) = \frac{\mathbb{P}(B|H_i)\,\mathbb{P}(H_i)}{\mathbb{P}(B)}$$

$$\Longleftrightarrow$$

$$\underbrace{\mathbb{P}(H_i|B)\,\mathbb{P}(B)}_{\mathbb{P}(H_i \cap B)} = \underbrace{\mathbb{P}(B|H_i)\,\mathbb{P}(H_i)}_{\mathbb{P}(B \cap H_i)}$$

1.9. Das Bayes-Theorem

Bayes-Theorem:

Sei $H_1,\ldots,H_n$ irgendeine Aufteilung von S und sei B ein Ereignis mit P(B) > 0, dann gilt für jedes H_i

$$P(H_i \mid B) = \frac{P(B \mid H_i)P(H_i)}{\sum_{j=1}^{n} P(B \mid H_j)P(H_j)}$$

Beispiel 1.9.1:

Im Beispiel 1.8.1 ist die Wahrscheinlichkeit, dass ein zufällig aus der Tagesproduktion herausgegriffenes Stück von Maschine M_1 produziert worden ist, a-priori

$$\mathbb{P}(\text{Stück auf M1 produziert}) = 2/3$$

$$\underbrace{\mathbb{P}(\text{Stück auf M1 produziert} \mid \text{Stück defekt})}_{\mathbb{P}(H_1 \mid B)} = \ ?$$

H_1: Stück wurde von M_1 gefertigt

H_2: Stück " " M_2 gefertigt $\quad (H_2 = \overline{H_1})$

$\Rightarrow H_1, H_2$ ist eine Partition

$$\boxed{P(H_i \mid B) = \frac{P(B \mid H_i)P(H_i)}{\sum_{j=1}^{n} P(B \mid H_j)P(H_j)}}$$

B: Stück defekt

Satz von Bayes

$$\mathbb{P}(H_1 \mid B) = \frac{\overset{0.1}{\mathbb{P}(B \mid H_1)}\overset{2/3}{\mathbb{P}(H_1)}}{\underset{0.1}{\mathbb{P}(B \mid H_1)}\underset{2/3}{\mathbb{P}(H_1)} + \underset{0.07}{\mathbb{P}(B \mid H_2)}\underset{1/3}{\mathbb{P}(H_2)}} = \underline{0.74}$$

Alternativ:

$$\mathbb{P}(H_1 \mid B) = \frac{\overset{0.1}{\mathbb{P}(B \mid H_1)}\overset{2/3}{\mathbb{P}(H_1)}}{\mathbb{P}(B)} = \underline{\underline{0.74}}$$

$\mathbb{P}(B) \to 0.09$

siehe einige Folien zuvor

Beispiel 1.9.1:

Im Beispiel 1.8.1 ist die Wahrscheinlichkeit, dass ein zufällig aus der Tagesproduktion herausgegriffenes Stück von Maschine M_1 produziert worden ist, a-priori

$$P(\text{Stück auf M1 produziert}) = \frac{2}{3} = 0.6\overline{6}.$$

Beobachtet man allerdings, dass das Stück defekt ist, wird man die Wahrscheinlichkeit sicher höher einschätzen wollen, da Maschine M_1 ja mehr Ausschuss produziert.

Beispiel 1.9.1 *(Fortsetzung):*

Nach der Bayes-Formel errechnet man

$$P\left(\text{Stück auf } M_1 \mid \text{Stück defekt}\right) = \frac{0.1 \cdot \frac{2}{3}}{0.1 \cdot \frac{2}{3} + 0.07 \cdot \frac{1}{3}}$$

$$= \frac{20}{27} = 0.741$$

In der Bayes-Statistik heissen $H_1, \ldots, H_n$ alternative ***Hypothesen***, $P(H_i)$ die ***a-priori-Wahrscheinlichkeit*** der i-ten Hypothese und $P(H_i|B)$ die ***a-posteriori-Wahrscheinlichkeit*** der i-ten Hypothese nach Kenntnis der Beobachtung B.

Beispiel 1.9.2:

Bei einer Krankheitsdiagnose sind die folgenden Angaben bekannt:

- Von den kranken Personen werden 90% durch die Untersuchung entdeckt.
- Von den gesunden Personen werden 99% durch die Untersuchung als gesund eingestuft.
- In der gesamten Bevölkerung sind 0.1% krank.

Beispiel 1.9.2 *(Fortsetzung)*:

Nun wird eine Person aus der Bevölkerung herausgegriffen, untersucht und als krank eingestuft. Wie wahrscheinlich ist es, dass das stimmt?

$\mathbb{P}(H_1 | B)$

1. Definiere Ereignisse:

H_1: Person ist krank

B: Person wird als krank diagnostiziert

$\mathbb{P}(B | H_1) = 0.9 \longrightarrow \mathbb{P}(\overline{B} | H_1) = 0.1$

$\mathbb{P}(\overline{B} | \overline{H_1}) = 0.99 \longrightarrow \mathbb{P}(B | \overline{H_1}) = 0.01$

$\mathbb{P}(H_1) = 0.001 \quad , \quad \mathbb{P}(\overline{H_1}) = 0.999$

$$\mathbb{P}(B) \overset{\text{Satz d. tot. W'keit}}{=} \underbrace{\mathbb{P}(B | H_1)}_{0.9} \underbrace{\mathbb{P}(H_1)}_{0.001} + \underbrace{\mathbb{P}(B | \overline{H_1})}_{0.01} \underbrace{\mathbb{P}(\overline{H_1})}_{0.999}$$

$$= 0.01089$$

$$\mathbb{P}(H_1 | B) \overset{\text{Bayes}}{=} \frac{\overset{0.9}{\mathbb{P}(B|H_1)} \overset{0.001}{\mathbb{P}(H_1)}}{\underset{0.01089}{\mathbb{P}(B)}} = 0.083$$

$\longrightarrow$ kein zuverlässiger Test

Beispiel 1.9.2 *(Fortsetzung):*

Nun wird eine Person aus der Bevölkerung herausgegriffen, untersucht und als krank eingestuft. Wie wahrscheinlich ist es, dass das stimmt?

Sei: $H_1 = \{\text{Person ist krank}\};$

$B = \{\text{Person wird als krank diagnostiziert}\}$

$$\rightarrow \quad P\left(B \middle| H_1\right) = 0.9\ ;\ P\left(\overline{B} \middle| H_1\right) = 0.1\ ;$$

$$P\left(\overline{B} \middle| \overline{H_1}\right) = 0.99\ ;\ P\left(B \middle| \overline{H_1}\right) = 0.01$$

***Beispiel 1.9.2** (Fortsetzung):*

a-priori-
Wahrscheinlichkeit

$$\rightarrow \qquad P(H_1) = 0.001 \quad \text{und} \quad P\left(\overline{H_1}\right) = 0.999$$

Satz der totalen
Wahrscheinlichkeit

$$\Rightarrow \qquad P(B) = P(B|H_1) \cdot P(H_1) + P\left(B\middle|\overline{H_1}\right) \cdot P\left(\overline{H_1}\right)$$

$$= 0.9 \cdot 0.001 + 0.01 \cdot 0.999 = 0.01089 \quad \text{und}$$

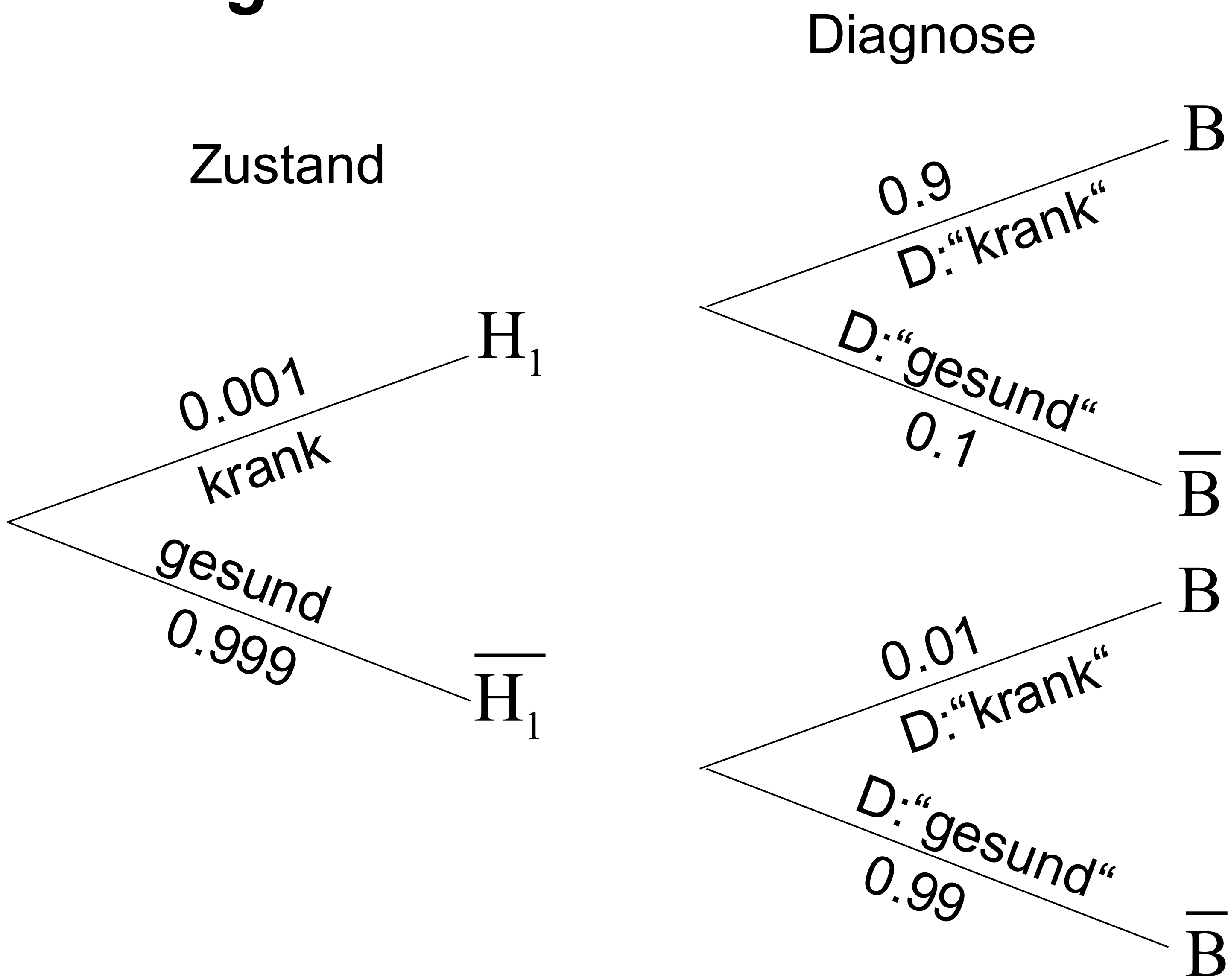
Bayes-Theorem

$$\Rightarrow \qquad P(H_1|B) = \frac{P(B|H_1) \cdot P(H_1)}{P(B)} = \frac{0.9 \cdot 0.001}{0.01089} = 0.0826,$$

d.h. die Vertrauenswürdigkeit einer

Krank-Diagnose ist nur 8.3%.

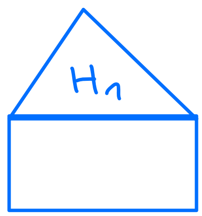
Baumdiagramm:



Universität St.Gallen

Simpson's Paradox

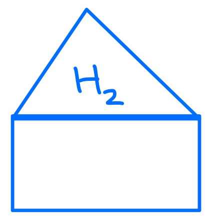
Hospital 1



$\mathbb{P}(S \mid A, H_1)$

	Success
Treatment A	$\frac{81}{87} = 0.93$
Treatment B	$\frac{234}{270} = 0.87$

Hospital 2



$\mathbb{P}(S \mid A, H_2)$

	Success
Treatment A	$\frac{192}{263} = 0.73$
Treatment B	$\frac{55}{80} = 0.69$

$\longrightarrow$ Für beide Spitäler scheint Behandlung A "besser"

$$\mathbb{P}(S \mid A) = \frac{81 + 192}{87 + 263} = \frac{273}{350} = 0.78$$

$$\mathbb{P}(S \mid B) = \frac{234 + 55}{270 + 80} = \frac{289}{350} = 0.83$$

Intuitive Erklärung ?

Kapitel 7.1

2. Elementare Kombinatorik

2.1. Fakultät und Binomialkoeffizient

Definition: Das Symbol $n!$ für $n \in \mathbb{N}$ ist definiert als

$$n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 1$$

heisst n-Fakultät

wir setzen $0! := 1$

Definition (Binomialkoeffizient)

$$\binom{n}{k} := \frac{n!}{k!\,(n-k)!}$$

ist der Binomialkoeffizient für ganze Zahlen $n > 0$, $k \geq 0$ $(n \geq k)$

$\binom{n}{k}$ "k aus n", beschreibt wie viele Möglichkeiten es gibt aus n Objekten k auszuwählen (ohne Zurücklegen ohne Reihenfolge)

Bsp: $n = 5$ Personen, A, B, C, D, E

Wie viele 2-er Teams lassen sich bilden

(A, B), (A, C), (A, D), (A, E), (B, C)
(B, D), (B, E), (C, D), (C, E), (D, E)

10 Möglichkeiten

$$\binom{5}{2} := \frac{5!}{2!\;3!} = \frac{5 \cdot \overset{2}{\cancel{4}} \cdot \cancel{3 \cdot 2 \cdot 1}}{\cancel{2} \cdot 1 \quad \cancel{3 \cdot 2 \cdot 1}} = \underline{\underline{10}}$$

2.1. Fakultät und Binomialkoeffizient

Definition

Das Symbol n! bezeichnet das Produkt der natürlichen Zahlen von 1 bis n

$$n! = 1 \cdot 2 \cdot 3 \cdot \underset{\ldots}{\cancel{L}} \cdot (n\text{-}1) \cdot n$$

und heisst ***n – Fakultät.***

Zusätzlich ist festgelegt: 0!= 1

Definition

Der **Binomialkoeffizient** $\binom{n}{k}$ ist für ganze n > 0, ganze k ≥ 0 und n ≥ k definiert als

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

$\binom{n}{k}$ gibt an, wieviele Möglichkeiten es gibt aus n objekten genau k auszuwählen

(ohne Reihenfolge, ohne Zurücklegen)

<u>Beispiel</u>: $n=5$ (Personen) A, B, C, D, E

<u>Frage</u>: Wie viele verschiedene 2-er Teams lassen sich bilden?

$$\binom{5}{2} = \frac{5!}{2!\,(5-2)!} = \frac{5\cdot \overset{2}{\cancel{4}}\cdot \cancel{3\cdot 2\cdot 1}}{\cancel{2}\cdot 1\ \cancel{3\cdot 2\cdot 1}} = 10$$

(A, B), (A, C), (A, D), (A, E), (B, C), (B, D), (B, E)
(C, D), (C, E) (D, E)

2.2. Das Fundamentalprinzip der Kombinatorik

"Wenn ein Sachverhalt auf n_1 Arten erfüllt werden kann und ein zweiter Sachverhalt unabhängig davon auf n_2 Arten, so ist die Gesamtzahl der Möglichkeiten, gleichzeitig beide Sachverhalte zu erfüllen, gerade gleich dem Produkt $n_1 \cdot n_2$."

2.3. Permutationen

Def: Gegeben eine Menge von n Elementen.

Jede Zusammenstellung dieser n Elemente in irgendeiner Reihenfolge heisst Permutation dieser n Elemente

Bsp: $\{A, B, C\}$

ABC, ACB, BAC, BCA,
CAB, CBA

Q: Was ist die Anzahl der Permutationen von n Elementen?

$n!$

2.3. Permutationen

Definition

Gegeben sei eine Menge mit n Elementen.
Jede Zusammenstellung all dieser Elemente in irgendeiner Reihenfolge heisst eine **Permutation** dieser n-Elemente.

\# Permutationen : $n!$

Bsp: A, B, C $\Rightarrow$ ABC, ACB, BAC, BCA, CAB, CBA

Universität St.Gallen

2.4. Kombinationen

Def: Gegeben sei eine Menge mit n verschiedenen Elementen.

Jede Zusammenstellung von k Elementen daraus heisst Kombination k-ter Ordnung

- Kombination mit Berücksichtigung der Anordnung $\dfrac{n!}{(n-k)!}$

- Kombination ohne Berücksichtigung der Anordnung $\dfrac{n!}{(n-k)!\,k!}$ $\binom{n}{k}$

2.4. Kombinationen

Definition

Gegeben sei eine Menge mit n verschiedenen Elementen. Jede Zusammenstellung von K Elementen daraus heisst Kombination K-ter Ordnung aus diesen Elementen.

- Kombinationen mit Berücksichtigung der Anordnung $\frac{n!}{(n-k)!}$
- Kombinationen ohne Berücksichtigung der Anordnung $\binom{n}{k}$

2.5. Variationen

Def: Gegeben sei eine Menge mit n Elementen.

Wieviele Sequenzen der Länge m kann man mit n Symbolen bilden

$\Longrightarrow$ Diese Anzahl heisst Anzahl der Variationen (mit Wiederholung) und ist

$$\underset{\text{1. Stelle}}{n} \cdot \underset{\text{2. Stelle}}{n} \cdot \ldots \cdot \underset{m^{\text{te}}\text{ Stelle}}{n} = n^m$$

2.5. Variationen

Definition

Wie viele Sequenzen der Länge m kann man mit n Symbolen bilden?

Diese Anzahl heisst die Anzahl der **Variationen** (mit Wiederholung) und ist gegeben durch n^m .

3. Zufallsvariablen

Definition: Gegeben sei ein Wahrscheinlichkeitsraum $(S, E, \mathbb{P})$.

Ereignisraum → S; Ereignismenge → E; Wahrscheinlichkeits fkt. → $\mathbb{P}$

Eine Funktion

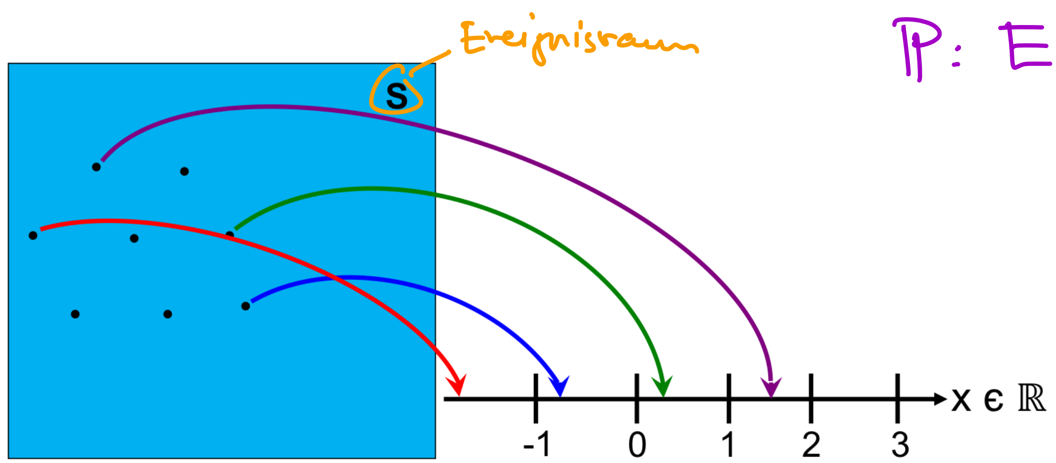
$$X: S \longrightarrow \mathbb{R}$$
$$s \longmapsto X(s)$$

$S = \{s_1, s_2, s_3, \ldots$

die jedem Elementarereignis $s \in S$ eine reelle Zahl $X(s)$ zuordnet, heisst Zufallsvariable, wenn für jedes $r \in \mathbb{R}$ gilt, dass $A_r \in E$, wobei

$$A_r := \{ s \in S : X(s) \leq r \}.$$

$\mathbb{P}: E \longrightarrow [0,1]$

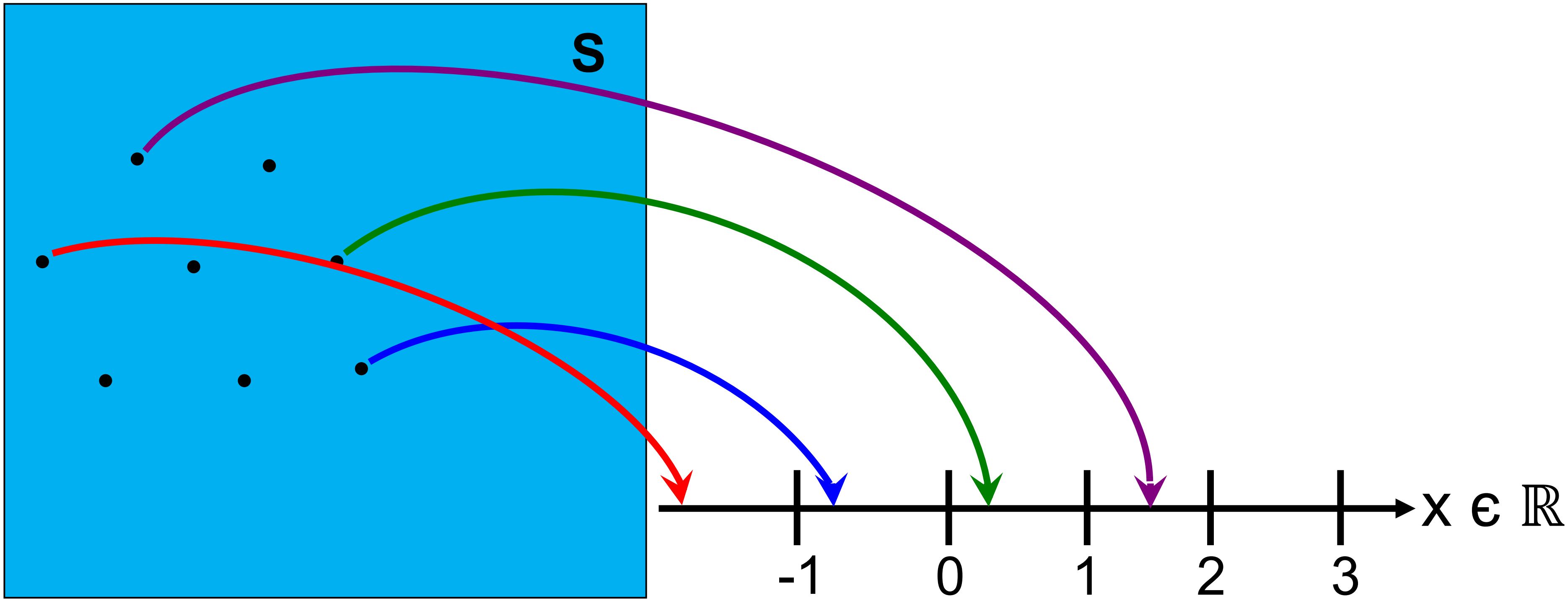


Definition

Gegeben sei ein Wahrscheinlichkeitsraum $[S, E, P(\cdot)]$. Eine Funktion

$$X: \quad S \longrightarrow \mathbb{R}$$
$$e \longrightarrow X(e) \in \mathbb{R},$$

die jedem Elementarereignis e eine reelle Zahl $X(e)$ zuordnet, heisst **Zufallsvariable**, wenn dabei zu jedem reellen r ein Ereignis $A_r \in E$ gehört, mit $A_r = \{e \mid X(e) \le r\}$.


S
-1
0
1
2
3
x ϵ ℝ

Beispiel 1:

Würfel $\quad S = \{1, 2, 3, 4, 5, 6\}$

X: gewürfelte Augenzahl

$X(s) = s$ für alle $s \in S$

Beispiel 2: (Werfen einer Münze)

$S = \{K, Z\}$

Definiere: $X(Z) := 0$

$X(K) := 1$

Beispiel 2: (Werfen einer Münze)

$S = \{K, Z\}$

Definiere: $X(Z) := 0$

$X(K) := 1$

Definition

Gegeben sei ein Wahrscheinlichkeitsraum $[S, E, P(\cdot)]$. Eine Funktion

$X: S \longrightarrow \mathbb{R}$

$e \longrightarrow X(e) \in \mathbb{R}$,

die jedem Elementarereignis e eine reelle Zahl X(e) zuordnet, heisst **Zufallsvariable**, wenn dabei zu jedem reellen r ein Ereignis $A_r \in E$ gehört, mit $A_r = \{e \mid X(e) \leq r\}$.

$E = \{\emptyset, \{K, Z\}, K, Z\}$ $\quad P: E \longrightarrow [0,1]$

Sei $r \in \mathbb{R}$

Behauptung

$$A_r = \{s \in S : X(s) \leq r\} \in E$$

- Für $-\infty < r < 0 \Longrightarrow A_r = \emptyset \in E$
- Für $0 \leq r < 1 \Longrightarrow A_r = \{Z\} \in E$
- Für $1 \leq r < \infty \Longrightarrow A_r = \{Z, K\} \in E$

Beispiel 3.0.1:

Beim Würfeln mit einem Würfel ist $S = \{1, 2, \ldots, 6\}$.

Die geworfene Augenzahl ist eine Zufallsvariable, beschrieben durch die Funktion

$$X(s) = s. \qquad \forall s \in S$$

Beispiel 3.0.2:

$S = \{Zahl, Kopf\}$

Eine Münze werde einmal geworfen. X bezeichne die Anzahl der Köpfe. X kann nur zwei Werte annehmen: $X(\text{Zahl}) = 0$ und $X(\text{Kopf}) = 1.$

Die Ereignismenge E hat die vier Ereignisse:

$$E = \{\emptyset, \text{Zahl}, \text{Kopf}, S\}.$$

Es gilt:

$$\text{für } -\infty < r < 0 \quad \text{ist} \quad A_r = \emptyset$$

$$\text{für } 0 \leq r < 1 \quad \text{ist} \quad A_r = \{\text{Zahl}\}$$

$$\text{für } 1 \leq r < \infty \quad \text{ist} \quad A_r = S$$

Beispiel 3.0.3:

Zwei regelmässige Würfel werden geworfen. Der Ereignisraum besteht aus 36 Elementarereignissen: $S = \{(i, j) \mid i=1,\ldots,6,\ j=1,\ldots,6\}$.

Mehrere Zufallsvariablen lassen sich hier bilden:

X = "Summe der Augenzahlen":

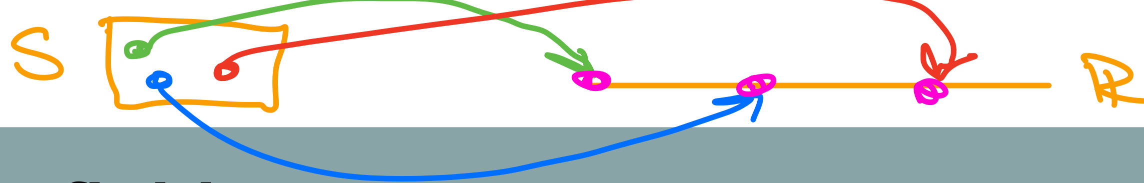
$$X(i, j) = i+j = x, \quad x = 2, 3, \ldots, 12.$$

Y = "Differenz der Augenzahlen":

$$Y(i, j) = |i-j| = y, \quad y = 0, 1, \ldots, 5.$$

Ziel später

Ausblick: $\mathbb{P}(X \leq k) = ?$



Definition

Sei W die Menge, die gerade jene reellen Zahlen enthält, die durch die Abbildung erreicht werden.
W heisst **Wertevorrat**. $W \subseteq \mathbb{R}$

Hat der Wertevorrat $W \subset \mathbb{R}$ einer Zufallsvariablen X endlich viele oder abzählbar unendlich viele Werte heisst sie **diskret**.

Besteht der Wertevorrat $W \subset \mathbb{R}$ einer Zufallsvariablen X aus der ganzen reellen Achse oder aus Teilintervallen so heisst sie **stetig**
(überabzählbar unendlich viele Werte).

Beispiel 3.0.5: Umsatz bei unsicherer Auftragslage

Der Umsatz eines Unternehmens ist bestimmt durch das Auftragsvolumen weniger Grossaufträge.

Für das nächste Jahr hofft die Geschäftsleitung, drei Grossaufträge A, B und C zu erhalten, wobei sie die Erfolgsaussichten bei der Akquisition unterschiedlich einschätzt (Aufträge stochastisch unabhängig).

Was ist die W'keit, dass die Firma Auftrag C bekommt, Aufträge A und B jedoch nicht?

Auftrag	Volumen [Mio CHF]	W'keit
A	10	0.8
B	14	0.5
C	24	0.75

$$\mathbb{P}(\bar{A} \cap \bar{B} \cap C) \overset{\text{Unabhängigkeit}}{=} \underbrace{\mathbb{P}(\bar{A})}_{0.2} \cdot \underbrace{\mathbb{P}(\bar{B})}_{0.5} \cdot \underbrace{\mathbb{P}(C)}_{0.75}$$

$$= \underline{0.075}$$

$S = \{-, A, B, C, AB, AC, CB, ABC\}$

Zufallsvariable X: Umsatz

$$\left.\begin{aligned} X(-) &= 0 \\ X(A) &= 10 \\ X(B) &= 14 \\ &\vdots \\ X(ABC) &= 48 \end{aligned}\right\}$$

Wertevorrat

$W = \{0, 10, 14, \dots, 48\}$

diskrete ZV.

$$\mathbb{P}(X = 24) = \mathbb{P}(\{C, AB\}) = \mathbb{P}(C) + \mathbb{P}(AB)$$

$$= \dots\dots \quad = \underline{\underline{0.175}}$$

siehe ok

$$\mathbb{P}(C) = \mathbb{P}(\bar{A} \cap \bar{B} \cap C) = 0.075$$

$$\mathbb{P}(AB) = \mathbb{P}(A \cap B \cap \bar{C}) = 0.8 \cdot 0.5 \cdot 0.25$$

Beispiel 3.0.5 *(Fortsetzung):*

Auftrag	Auftragsvolumen [Mio. CHF]	Wahrscheinlichkeit des Auftrags
A	10	0.8
B	14	0.5
C	24	0.75

Der Umsatz ist dann eine Zufallsvariable X, die von der unsicheren Auftragslage abhängt.

$$P\left(\bar{A} \cap \bar{B} \cap C\right) = (1-0.8)\cdot(1-0.5)\cdot 0.75 = 0.075$$

Wertevorrat von X ?

X : Umsatz

Auftragslage (e_i)	**Umsatz ($X(e_i)$)**	**$P(\{e_i\})$**
–	0	0.025
A	10	0.1
B	14	0.025
C	24	0.075
AB	24	0.1
AC	34	0.3
BC	38	0.075
ABC	48	0.3
		$\Sigma = 1$

$$W = \{0, 10, 14, 24, 34, 38, 48\}$$

$$\rightarrow P[X=24] = P[\{C, AB\}] = 0.175.$$

Universität St.Gallen

Teil I: Wahrscheinlichkeitsrechnung

Teil II: Statistik

Teil III: Aufgaben

Universität St.Gallen

Kapitel 9.1

3.1. Verteilungsfunktion

Definition

Die Funktion $F(x) = P(X \leq x) = P(\{e|X(e) \leq x\})$, die jeder reellen Zahl x die Wahrscheinlichkeit zuordnet, mit der die Zufallsvariable X einen Wert $X \leq x$ annimmt, heisst **Verteilungsfunktion** von X.

Def. Sei X eine Zufallsvariable.
Dann nennen wir die Funktion

$F: \mathbb{R} \to [0,1]$ definiert als

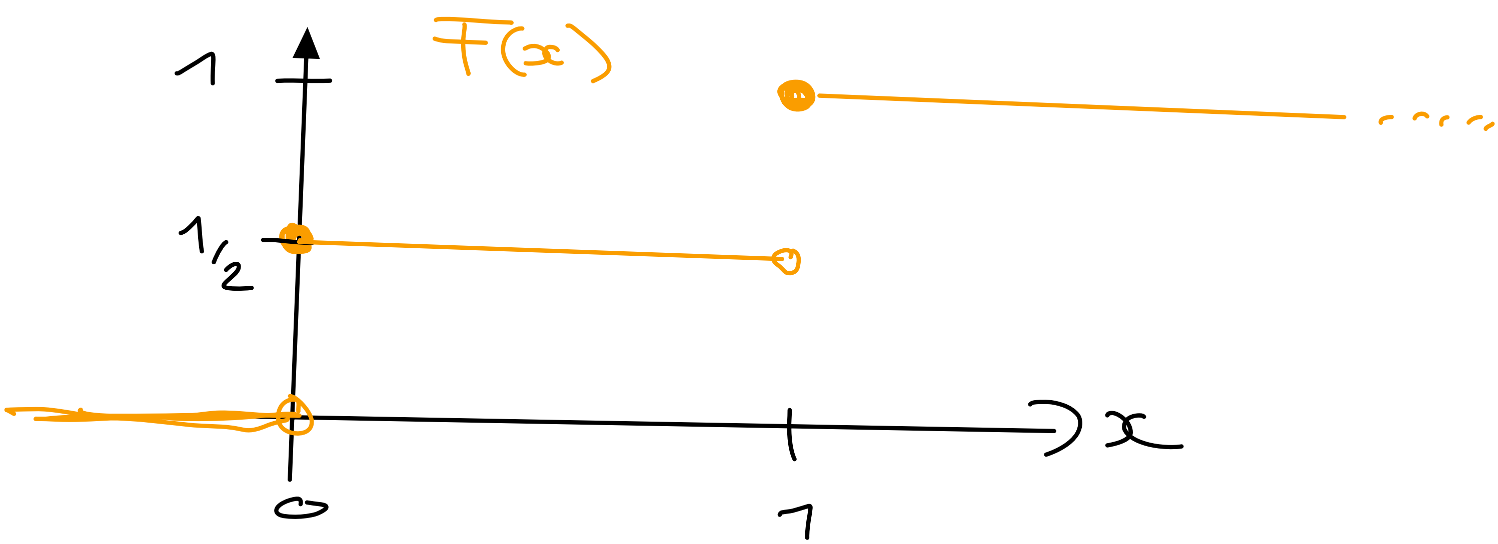
$$F(x) = \mathbb{P}(X \le x) \quad \mathbb{P}(\{s \in S : X(s) \le x\})$$

heisst Verteilungsfunktion von X

Bsp: $X :=$ # Köpfe beim einfachen Münzwurf

$\mathbb{P}(X=0) = 1/2$, $\mathbb{P}(X=1) = 1/2$

$$F(x) = \begin{cases} 0 & , x < 0 \\ 1/2 & , 0 \le x < 1 \\ 1 & , x \ge 1 \end{cases}$$



W'keitsraum $(S, E, \mathbb{P})$

Wiederholung:

Zufallsvariable $X : S \to \mathbb{R}$

$$W = \{ w \in \mathbb{R} : \exists s \in S,\ X(s) = w \}$$

Wertevorrat:

W endlich o. abzählbar $\Rightarrow X$ diskret

W unabzählbar $\Rightarrow X$ stetig

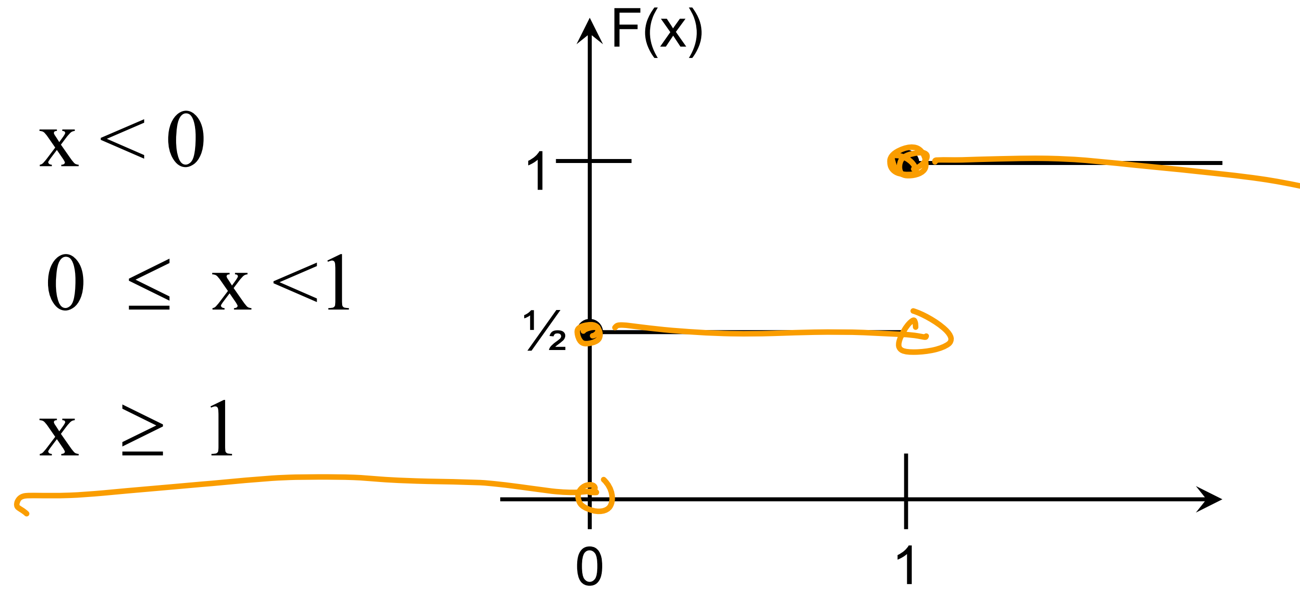
Verteilungsfunktion: $F_X : \mathbb{R} \to [0,1]$

$$F_X(x) = \underbrace{\mathbb{P}(X \le x)}_{\mathbb{P}(\{s \in S : X(s) \le x\})}$$

Beispiel 3.1.1:

X = “Anzahl der Köpfe beim einfachen Münzwurf“.
X kann die beiden Werte 0 oder 1 annehmen. Wenn die Wahrscheinlichkeit, Kopf zu werfen, 0.5 beträgt, dann gilt:

$$F(x) = \begin{cases} 0 & , \text{ für } x < 0 \\ \frac{1}{2} & , \text{ für } 0 \leq x < 1 \\ 1 & , \text{ für } x \geq 1 \end{cases}$$



Y diskrete Z.V.

$W = \{0, 1, 2, \ldots, 5\}$

Beispiel 3.1.2:

Zwei Würfel werden geworfen. Y sei die Differenz der Augenzahlen.

$S = \{(1,1), (1,2) \ldots \}$

Dann:

Y = 0 ,	wenn (i, j) = (1,1); (2,2); (3,3); ...; (6,6)	:	6 Paare
Y = 1 ,	wenn (i, j) = (1,2); (2,1); ...	:	10 Paare
Y = 2 ,	wenn (i, j) = (1,3); (3,1); (2,4); ...	:	8 Paare
Y = 3 ,	wenn (i, j) = (1,4); (4,1); ...	:	6 Paare
Y = 4 ,	wenn (i, j) = (1,5); (5,1); (2,6); (6,2)	:	4 Paare
Y = 5 ,	wenn (i, j) = (1,6); (6,1)	:	2 Paare
			36 Paare

$W = \{w \in \mathbb{R} : \exists s \in S, Y(s) = w\}$

Beispiel 3.1.2 *(Fortsetzung):*

Also:

$$P[Y=0]=\frac{6}{36}=\tfrac{1}{6}; \qquad P[Y=3]=\frac{6}{36}=\tfrac{1}{6};$$

$$P[Y=1]=\frac{10}{36}=\frac{5}{18}; \qquad P[Y=4]=\frac{4}{36}=\tfrac{1}{9};$$

$$P[Y=2]=\frac{8}{36}=\frac{2}{9}; \qquad P[Y=5]=\frac{2}{36}=\tfrac{1}{18};$$

$\mathbb{P}(Y \leq y)$ for $1 \leq y < 2$

$= \mathbb{P}(Y=0)$

$+ \mathbb{P}(Y=1)$

$y = 1.2$

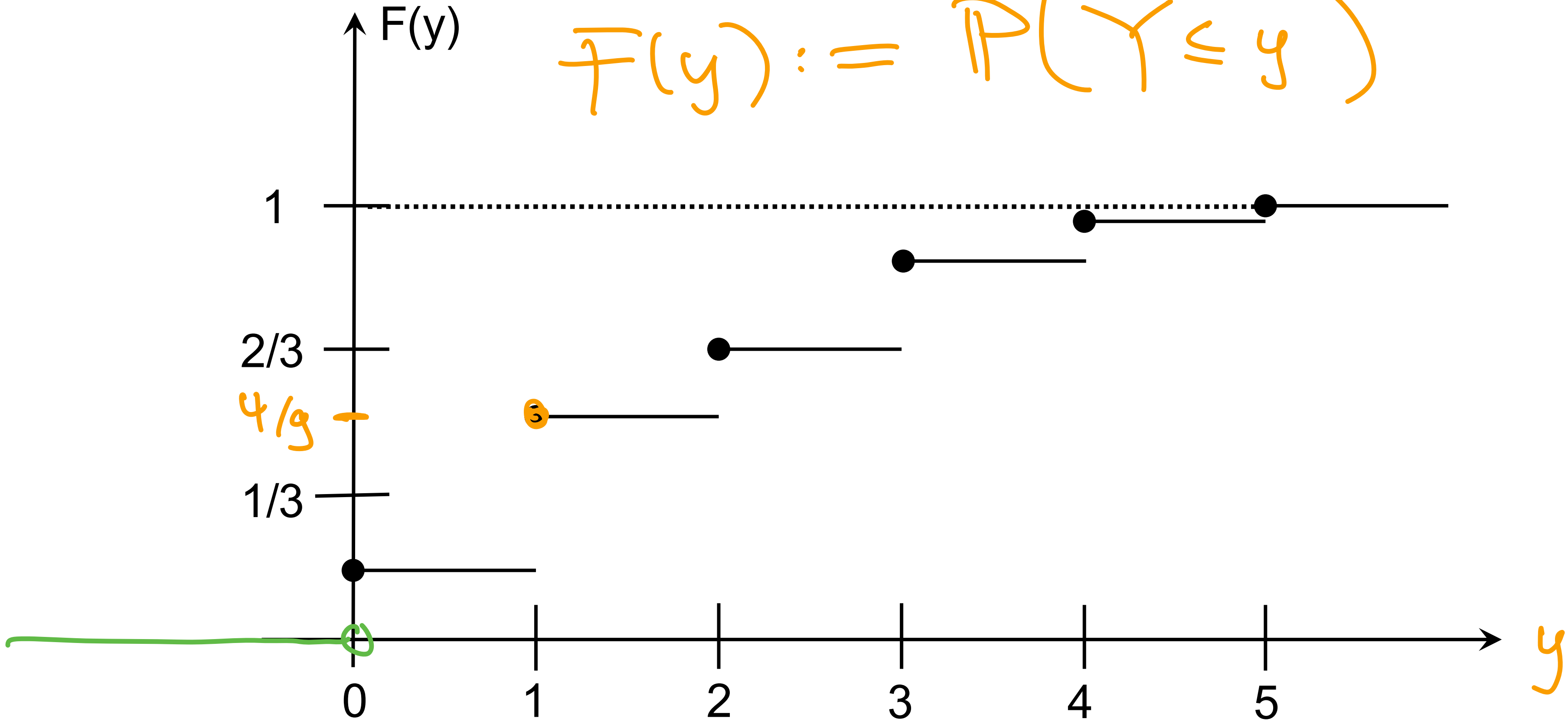
$\mathbb{P}(Y \leq y)$, $y < 0$

$$\Rightarrow F_Y(y) = \begin{cases} 0 & , \quad y < 0 \\ 1/6 & , \quad 0 \leq y < 1 \\ 4/9 & , \quad 1 \leq y < 2 \\ 2/3 & , \quad 2 \leq y < 3 \\ 5/6 & , \quad 3 \leq y < 4 \\ 17/18 & , \quad 4 \leq y < 5 \\ 1 & , \quad y \geq 5 \end{cases}$$

$\mathbb{P}(Y \leq 1.2)$

$F(y)$ $\qquad F(1) = 4/9$

$F(y) := \mathbb{P}(Y \leq y)$



$F(y)$

Eigenschaften der Verteilungsfunktion:

(1) F(x) ist an jeder Stelle x zumindest rechtsseitig stetig: $\lim_{\Delta x \to 0} F(x + \Delta x) = F(x)$.

(2) F(x) ist überall monoton steigend:

$$F(a) \leq F(b) \quad \text{für} \quad a < b.$$

(3) F(x) hat die Grenzwerte:

$$\lim_{x \to -\infty} F(x) = 0 \quad \text{und} \quad \lim_{x \to +\infty} F(x) = 1.$$

Universität St.Gallen

Kapitel 9.2

endlich oder abzählbarer Wertebereich

3.2. Diskrete Zufallsvariablen

Definition

Ist X eine **diskrete** Zufallsvariable, dann heisst die Funktion f(x) = P [X=x] die **Massenfunktion** der Zufallsvariablen X.

Natürlich hat sie nur an den Stellen $x=x_i$, die zum Wertevorrat W gehören, positive Werte ~~pi~~ p_i = $P(X = x_i)$. Dazwischen ist sie Null:

$$f(x) = P(X = x) = \begin{cases} p_i & x = x_i \in W, \\ 0 & sonst. \end{cases}$$

1 $p_i = f(x_i) \geq 0 \quad \forall i$

2 $\sum_i p_i = \sum_i P(X=x_i) = 1$

3 $p_i \leq 1$

Jede Massenfunktion f hat die Eigenschaften:

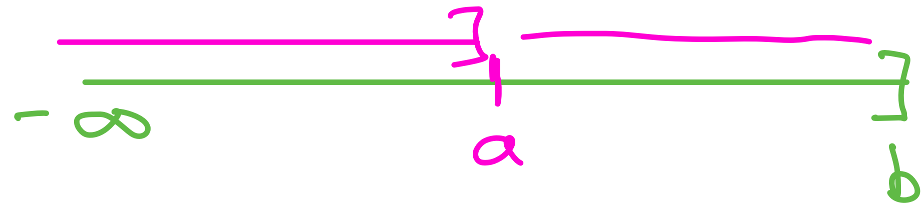
(1) $f(x_i) \geq 0$ (Wahrscheinlichkeit nicht negativ!)

(2) $\sum_{\text{alle } i} f(x_i) = 1$ (Sicheres Ereignis hat Wahrscheinlichkeit 1)

Aus (1) und (2) folgt unmittelbar (3) $f(x_i) \leq 1$.

$a < b$

Merke:



Für Intervall-Wahrscheinlichkeiten gilt die folgende Grundformel:

$$P(a<X \le b) \overset{(*)}{=} F(b) - F(a) = \sum_{a<x_i \le b} p(x_i)$$

Diese ist bei diskreten Zufallsvariablen zu modifizieren:

$$F(b) = F(a) + P(a<X\le b)$$

$$P(a < X < b) = F(b) - F(a) - f(b)$$

$$P(a \le X \le b) = F(b) - F(a) + f(a)$$

$$P\left(X \in (a,b]\right)$$

$$P(a \le X < b) = F(b) - F(a) + f(a) - f(b)$$

Begründung von (*): $(-\infty, b] = (-\infty, a] \cup (a,b] \Rightarrow P(X \in (-\infty, b]) = P(X \in (-\infty, a]) + \cdot$

$F(b) = P(X \le b)$ $\qquad$ $P(X \le a) = F(a)$

Universität St.Gallen

3.3. Stetige Zufallsvariablen

Definition

Ist X eine stetige Zufallsvariable mit der Verteilungsfunktion F, so heisst die erste Ableitung von F nach x

$$f(x) = \frac{d}{dx} F(x)$$

die **(Wahrscheinlichkeits)dichtefunktion** von X.

Fundamentalsatz der Analysis $\Longleftrightarrow$ $F(x) = \int_{-\infty}^{x} f(s)\, ds$

Jede Dichtefunktion hat die Eigenschaften:

(1) $\boxed{f(x) \geq 0}$ (Verteilungsfunktion monoton steigend!)

(2) $\boxed{\int_{-\infty}^{\infty} f(x)dx = 1}$ (Die Fläche unter jeder Dichte-funktion ist genau 1.)

Universität St.Gallen

Beispiel 3.3.1:

Eine stetige Zufallsvariable X sei charakterisiert durch die Verteilungsfunktion

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{1}{27}(x-3)^3 + 1 & , \quad 0 \le x < 3 \\ 1 & , \quad x \ge 3 \end{cases}$$

$$f(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{1\,\cancel{3}}{9\,\cancel{27}} \cdot (x-3)^2 & , \quad 0 \le x < 3 \\ 0 & , \quad x \ge 0 \end{cases}$$

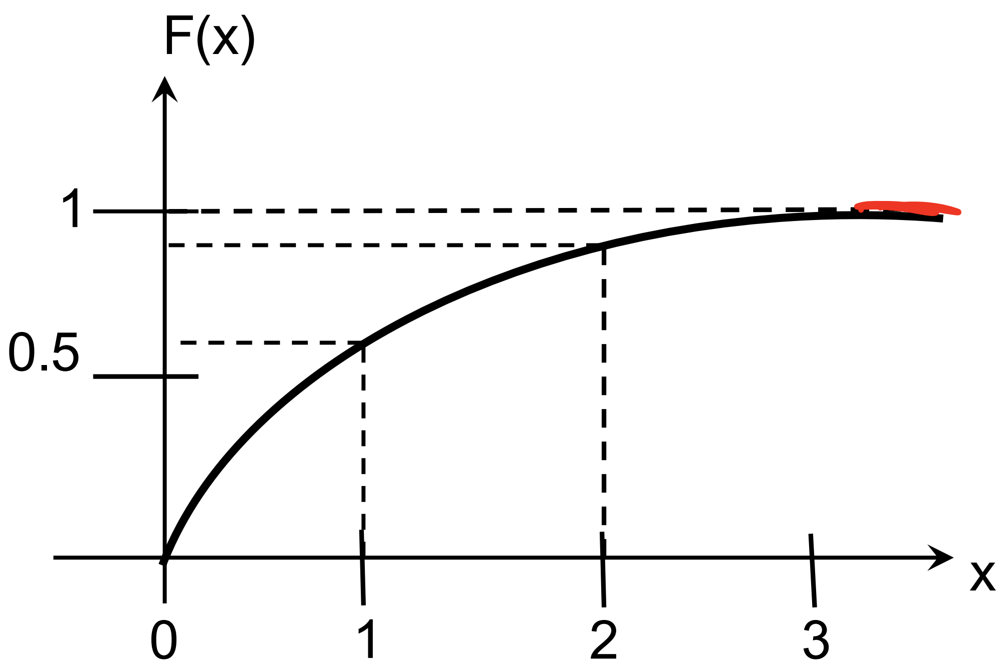
Beispiel 3.3.1 *(Fortsetzung):*

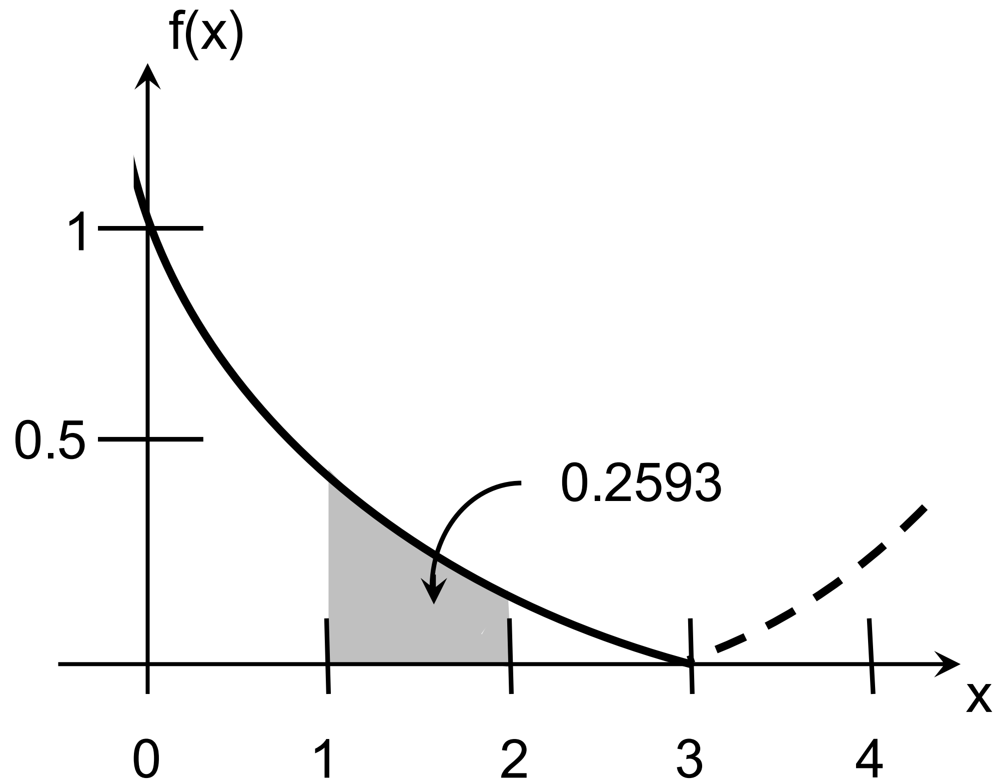
Berechnung der Dichtefunktion:

(F(·) in allen drei Teilstücken ableiten)

$$f(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{1}{9}(x-3)^2 & , \quad 0 \leq x < 3 \\ 0 & , \quad x \geq 3 \end{cases}$$

Beispiel 3.3.1 *(Fortsetzung):*





Bei stetigen ZV X gilt:

$\mathbb{P}(X=c)=0$ für alle c

Universität St.Gallen

Wie gross ist $P(1 \leq X \leq 2) = F(2) - F(1) =$

$$\frac{-1}{27} + \cancel{1} - \left(\frac{(-2)^3}{27} + \cancel{1}\right) = \frac{7}{27} = 0.2593$$

f(x)

oder $\quad P(1 \leq X \leq 2) = \int_1^2 \frac{1}{9}(x\text{-}3)^2 \, dx =$

$$\left[\frac{1}{27}(x-3)^3\right]_1^2 = \frac{-1}{27} - \frac{-8}{27} = \frac{7}{27}.$$

Beide Rechnungen führen selbstverständlich zum gleichen Ergebnis. Die Information enthalten in $F(\cdot)$ und $f(\cdot)$ ist dieselbe!

Sei X eine stetige Z.V.

und sei $a < b$ beliebig. Dann gilt

$$\mathbb{P}(\underbrace{a \leq X \leq b}_{X \in [a,b]}) = \int_a^b f(x)\,dx$$

wobei $f(\cdot)$ die Dichte-Funktion von X ist

Dichtefunktion einer uniform verteilten Z.V.

Beispiel 3.3.2: Wartezeit an der S-Bahn Station

An einer S-Bahn Station fahren die Züge im 12-Minuten Takt. Ein Fahrgast, der den Fahrplan ignoriert, erscheint zu einem zufälligen Zeitpunkt an der Station.

Dichtefunktion $f(x) = \begin{cases} 1/12, & x \in [0,12] \\ 0, & \text{sonst} \end{cases}$

X = Wartezeit

$F(x) = \int_{-\infty}^{x} f(y)\, dy \qquad x \in [0,12]$

Zufallsvariable mit Wertevorrat W = [0,12] (Minuten)

Frage: (1) $F(x) = \begin{cases} \frac{x}{12}, & x \in [0,12] \\ 0, & x \notin [0,12] \end{cases}$

$F(x) = \int_0^x \frac{1}{12}\, dy = \frac{1}{12} \int_0^x 1\, dy = \frac{x}{12}$

$y \Big|_0^x$

(2) Berechne $\underbrace{\mathbb{P}(10 \leq X \leq 15)}_{\mathbb{P}(10 \leq X \leq 12)} = F(15) - F(10)$

$= \int_{10}^{12} f(x)\, dx$

uniform Verteilung

Dichtefunktion: $f(x) = \begin{cases} \frac{1}{12} & , \quad \text{für} \quad x \in [0, 12] \\ 0 & , \quad \text{sonst.} \end{cases}$

Verteilungsfunktion:

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{x}{12} & , \quad 0 \le x \le 12 \\ 1 & , \quad x > 12 \end{cases} \longleftarrow \int_0^x \frac{1}{12} du = \left[\frac{u}{12}\right]_0^x = \frac{x}{12}$$

Dann: $P(10 < X < 15) = F(15) - F(10) = 1 - \frac{10}{12} = 0.1667$

$$P(X > 9) = 1 - F(9) = 1 - \frac{9}{12} = 0.25$$

Universität St.Gallen

Kapitel 9.4

Definition: Sei X eine Z.V., sei f ihre Massenfunktion (Dichtefunktion). Der Erwartungswert von X ist definiert als

(1) Falls X diskret $\quad \mathbb{E}[X] := \sum_i x_i \underbrace{f(x_i)}_{=p_i}$

(2) Falls X stetig $\quad \mathbb{E}[X] := \int_{-\infty}^{\infty} x f(x)\,dx$

Satz: (Zentraleigenschaft)

$$\mathbb{E}\left[\underbrace{X - \mathbb{E}[X]}_{Y}\right] = 0$$

, Y zentrierte Form von X

Bew: Übungsaufgabe

3.4. Erwartungswerte von Zufallsvariablen

Definition

Sei X eine Zufallsvariable und f ihre Massen- bzw. Dichtefunktion.

Ihr Erwartungswert ist definiert als

$$E[X] = \sum_{\text{alle } j} x_j f(x_j) = \sum_{\text{alle } j} x_j p_j, \qquad \text{falls X diskret}$$

$$E[X] = \int_{-\infty}^{\infty} x f(x)\,dx, \qquad \text{falls X stetig.}$$

Der Erwartungswert ist als **Lageparameter** der Verteilung anzusehen.

Satz:

Der Erwartungswert der Abweichungen jeder Zufallsvariablen X von ihrem Erwartungswert μ_x ist Null, das heisst

$$E[X - \mu_x] = 0 \quad \text{(Zentraleigenschaft).}$$

Universität St.Gallen

Beispiel 3.4.1:

$$E[X] = \sum_{i=1}^{2} x_i f(x_i)$$

$$= 1 \cdot \underbrace{f(1)}_{1/3} + 2 \cdot \underbrace{f(2)}_{2/3}$$

X Zufallsvariable mit Massenfunktion

$$f(x) = \begin{cases} x/3 & , \quad x = 1, 2 \\ 0 & , \quad \text{sonst.} \end{cases}$$

$$\Rightarrow \quad E[X] = 1 \cdot \frac{1}{3} + 2 \cdot \frac{2}{3} = 5/3$$

diskrete Z.V.

Beispiel 3.4.2:

Absolute

Y = "Differenz der Augenzahlen zweier Würfel"

Wir erhalten:

Y	0	1	2	3	4	5
$f(y_i)$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{8}{36}$	$\frac{6}{36}$	$\frac{4}{36}$	$\frac{2}{36}$

$$\Rightarrow \quad E[Y] = 0 \cdot \frac{1}{6} + 1 \cdot \frac{5}{18} + 2 \cdot \frac{2}{9} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{9} + 5 \cdot \frac{1}{18}$$

$$= \frac{5+8+9+8+5}{18} = \frac{35}{18}$$

Bsp: Sei X stetig mit Dichtefunktion

$$f(x) = \begin{cases} c \cdot x & , \ 1 \leq x \leq 3 \\ 0 & , \ \text{sonst} \end{cases} \qquad c = 1/4$$

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_1^3 c \cdot x\,dx = c \int_1^3 x\,dx \qquad \left.\frac{x^2}{2}\right|_1^3$$

$$= c\left(\underbrace{\frac{9}{2} - \frac{1}{2}}_{4}\right) \overset{!}{=} 1$$

$$\Rightarrow \underline{\underline{c = 1/4}}$$

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x)\,dx = \int_1^3 x \cdot \frac{1}{4} \cdot x\,dx$$

$$= \frac{1}{4} \int_1^3 x^2\,dx \qquad \left.\frac{x^3}{3}\right|_1^3$$

$$= 1/4 \left(\frac{27}{3} - \frac{1}{3}\right) = \frac{26}{12} = \underline{\underline{\frac{13}{6}}}$$

Spezialfall

$g(x) = x$

- Sei X Zufallsvariable
- Sei $g: \mathbb{R} \to \mathbb{R}$

Ziel: $\mathbb{E}[g(X)] = ?$

$p_i = f(x_i)$

Formel:

- X diskret $\Rightarrow \mathbb{E}[g(X)] = \sum_i g(x_i)\, p_i$
- X stetig $\Rightarrow \mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx$

Beispiel 3.4.3:

X stetige Zufallsvariable mit Dichtefunktion

$$f(x) = \begin{cases} \frac{1}{4}\,x & , \quad 1 \le x \le 3 \\ 0 & , \quad \text{sonst.} \end{cases}$$

$$\Rightarrow \; E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\,dx = \int_{1}^{3} \frac{1}{4} x^2 dx = \frac{1}{12} x^3 \Big|_{1}^{3}$$

$$= \frac{27-1}{12} = \frac{26}{12} = \frac{13}{6}.$$

Erwartungswert einer Funktion von Zufallsvariablen

Zuweilen gilt es, anstelle von E[X] mit gegebener Massen- oder Dichtefunktion f(x) den Erwartungswert einer Funktion g(X) zu berechnen:

$$E[g(X)] = \sum_{\text{alle } j} g(x_j) \cdot p_j \qquad \text{, falls X diskret}$$

$$E[g(X)] = \int_{\mathbb{R}} g(x) \cdot f(x)dx \qquad \text{, falls X stetig.}$$

$$g(\mathbb{E}[X]) \neq \mathbb{E}[g(X)] = g(0)\cdot 0.35 + g(1)\cdot 0.4 + g(2)\,0.15 + g(3)\cdot 0.1 = \dots = \underline{\underline{2.5}}$$

Beispiel 3.4.4:

Beim Betrieb einer Produktionsanlage treten Störfälle auf. Aus Beobachtungen in der Vergangenheit hält man für X="die Anzahl der Störfälle pro Tag" folgende Wahrscheinlichkeitsverteilung für korrekt:

Q: Was sind die erwarteten Kosten?

X	0	1	2	3
$f(x)$	0.35	0.4	0.15	0.1

Massenfunktion

Kosten als Funktion der # Störfälle

$$g(x) := 5 - \frac{4}{x+1}$$ in tausend CHF

Beispiel 3.4.4 *(Fortsetzung):*

Mit der Behebung der Störfälle sind Kosten

gemäss $g(x) = 5 - \frac{4}{x+1}$ (in tausend CHF)

verbunden.

Würde man

$$E[X] = \cancel{0 \cdot 0.35} + 1 \cdot 0.4 + 2 \cdot 0.15 + 3 \cdot 0.1 = 1$$

in die Kostenfunktion einsetzen, erhielte man

erwartete Kosten in Höhe von

$$g(E[X]) = 5 - \frac{4}{1+1} = 3.$$

Beispiel 3.4.4 *(Fortsetzung):*

Aber:

Korrekt müssten wir jedoch die erwarteten Kosten als

$$E[g(X)]$$

$$= g(0)\cdot 0.35 \; + \; g(1)\cdot 0.4 \; + \; g(2)\cdot 0.15 \; + \; g(3)\cdot 0.1$$

$$= 1\cdot 0.35 + 3\cdot 0.4 + \frac{11}{3}\cdot 0.15 + 4\cdot 0.1$$

$$= 2.5$$

berechnen.

Beispiel 3.4.5:

X	0	1	2	3
f_x	0.1	0.3	0.2	0.4

$$Y=(X-2)^2 \rightarrow f_y? \qquad W_y=\{0, 1, 4\}$$

Y	0	1	4
f_y	0.2	0.7	0.1

$$\rightarrow E[Y] = 0.7 + 0.4 = 1.1$$

Satz (Eigenschaften von EW)

Sei X eine Z.V., seien $a, b \in \mathbb{R}$ Konstanten. Dann gilt dass

$$\mathbb{E}[\underbrace{aX+b}] = \underline{a\mathbb{E}[X]+b}$$

Linearität des EW.

$g(X)$ für $\underline{g(x) = ax+b}$

Lineare Funktion

$g(\mathbb{E}[X])$

Bsp: $b = -\mathbb{E}[X]$, $a = 1$

$$\overset{\text{Satz}}{\Longrightarrow} \mathbb{E}\big[X - \mathbb{E}[X]\big] = \mathbb{E}[X] - \mathbb{E}[X]$$

$$= 0$$

Beispiel Sei X eine Z.V. mit Dichte

$$f_X(x) = \begin{cases} e^{-x}, & x \geq 0 \\ 0 & \end{cases}$$

Dichtefkt. einer exponentiell verteilten Z.V.

, $Y := 2X + 1$

Frage: $\mathbb{E}[Y] = ?$

Satz

$$\mathbb{E}[Y] = \mathbb{E}[\underbrace{2}_{a}X + \underbrace{1}_{b}] \stackrel{\text{Satz}}{=} 2\,\mathbb{E}[X] + 1$$

$$\mathbb{E}[X] = \int_0^\infty x \, e^{-x} \, dx$$

$$= \left(-x e^{-x}\right)_0^\infty - \int_0^\infty 1 \cdot (-e^{-x})\, dx$$

$$= \ldots = 1$$

☒ VL

Bisher:

Massenfunktion

$P(X=x_j)$
$=f(x_j)$

Definition

Sei X eine Zufallsvariable und f ihre Massen- bzw. Dichtefunktion.

Ihr Erwartungswert ist definiert als

$$E[X] = \sum_{\text{alle } j} x_j f(x_j) = \sum_{\text{alle } j} x_j p_j, \qquad \text{falls X diskret}$$

$$E[X] = \int_{-\infty}^{\infty} x f(x)\,dx, \qquad \text{falls X stetig.}$$

Wahrscheinlichkeitsdichtefunktion
(Dichtefunktion)

$g: \mathbb{R} \to \mathbb{R}$

$$\mathbb{E}[g(X)] = \sum_j g(x_j) f(x_j) = \sum_j f(x_j) p_j \quad : \quad X \text{ diskret}$$

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\,dx \quad : \quad X \text{ stetig}$$

Dichtefunktion $f(x) \geq 0 \quad \forall x \in \mathbb{R}$

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

Rechnen mit Erwartungswerten

Satz

Wenn X eine Zufallsvariable mit Erwartungswert E[X] ist und a, b reelle Konstanten sind, dann hat

$Y = aX + b$

den Erwartungswert

$E[Y] = E[aX + b] = a \cdot E[X] + b.$

Universität St.Gallen

Beh: $\mathbb{E}[aX+b] = a\mathbb{E}[X] + b$

Bew: (Diskreter Fall)

① $\mathbb{E}[aX] = a\,\mathbb{E}[X]$

- $\mathbb{E}[aX] \overset{\text{Def}}{=} \sum_j (ax_j) f(x_j) = a \sum_j x_j f(x_j) = a\mathbb{E}[X]$

② $\mathbb{E}[X+Z] = \mathbb{E}[X] + \mathbb{E}[Z]$

- $\mathbb{E}[X+Z] = \sum_j \sum_i (x_j + z_i) f_{X,Z}(x_j, z_i)$

$$= \sum_j \sum_i x_j f_{X,Z}(x_j, z_i) + \sum_j \sum_i z_i f_{X,Z}(x_j, z_i)$$

$$= \sum_j x_j f_X(x_j) + \sum_i \sum_j z_i f_{X,Z}(x_j, z_i)$$

$$= \mathbb{E}[X] + \mathbb{E}[Z]$$

$$\Rightarrow \mathbb{E}[aX+b] \overset{②}{=} \underbrace{\mathbb{E}[aX]}_{\overset{①}{=} a\mathbb{E}[X]} + b$$

$\blacksquare$

Beispiel 3.4.6:

X stetig

Sei X eine Zufallsvariable mit Dichtefunktion

$$f_x(x) = \begin{cases} e^{-x}, & x \geq 0 \\ 0, & \text{sonst} \end{cases} \qquad \text{und} \quad Y = \underbrace{2}_{a}X + \underbrace{1}_{b}$$

$$E[Y] = ? \qquad E[2x+1] \overset{\text{Satz}}{=} 2 \cdot E[X] + 1 = 3$$

$$\hookrightarrow E[X] \overset{\text{Def}}{=} \int_0^{+\infty} x\underbrace{e^{-x}}_{g(x)}\, dx \underset{\text{P.I.}}{=} \underbrace{\cancel{\left(-x.e^{-x}\right)_0^{+\infty}}}_{0} + \int_0^{+\infty} 1 \cdot e^{-x} dx = -e^{-x} \Big|_0^{\infty} = 1$$

direkt!

$$f_y = ?$$

$$G(x) = -e^{-x}$$

Transformation von Zufallsvariablen

(A) Sei X eine stetige Z.V., sei $Y := g(X)$

Sei $f_X(\cdot)$ die Dichtefunktion von X

<u>Frage</u>: Was ist die Dichtefkt. von Y, $f_Y(\cdot)$?

<u>Satz</u>: (Dichtetransformation)

$$\boxed{f_Y(y) = f_X\left(g^{-1}(y)\right) \cdot \left| \frac{d\, g^{-1}(y)}{dy} \right|}$$

<u>Bsp</u>: Sei $g(x) = 2x + 1$, Was ist f_Y für $Y = g(X)$?

$$f_X(x) = \begin{cases} e^{-x}, & x \geq 0 \\ 0, & \text{o.w.} \end{cases}$$

$$g(x) = y \iff 2x + 1 = y$$
$$\iff 2x = y - 1$$
$$\iff x = \frac{y-1}{2} \implies g^{-1}(y) = \frac{y-1}{2}$$

$$\implies \frac{d\, g^{-1}(y)}{dy} = \left(\frac{y-1}{2}\right)' = \frac{1}{2}$$

$y \geq 1$

$\Updownarrow$

$$f_Y(y) \overset{\text{Satz}}{=} \begin{cases} e^{-\left(\frac{y-1}{2}\right)} \cdot \frac{1}{2}, & \text{falls } \frac{y-1}{2} \geq 0 \\ 0, & \text{sonst} \end{cases}$$

(B) Sei X diskret

$g: \mathbb{R} \to \mathbb{R}$, $Y := g(X)$

Massenfunktion $f_X(x) = \mathbb{P}(X = x)$

$f_Y(y) = \mathbb{P}(Y = y)$

$$f_Y(y) \overset{\text{Def}}{=} \mathbb{P}(Y=y) = \mathbb{P}(g(X) = y)$$

$$= \sum_{x \in g^{-1}(y)} f_X(x)$$

$\{x \in \mathbb{R} : \underbrace{g^{-1}(y) = x}_{g(x) = y}\}$

Satz

Unter geeigneten Voraussetzungen

(Dichtetransformation):

$$f_y(y) = f_x\left(g^{-1}(y)\right) \cdot \left|\frac{dg^{-1}(y)}{dy}\right|$$

In Beispiel 3.4.6:

$$\rightarrow y = g(x) = 2x + 1 \quad \rightarrow \quad g^{-1}(y) = \frac{y-1}{2}$$

$$\Rightarrow \quad f_y(y) = e^{-\frac{y-1}{2}} \cdot \frac{1}{2}, \quad \text{falls} \quad \frac{y-1}{2} \geq 0 \quad \Leftrightarrow \quad y \geq 1$$

$$\cancel{\Leftrightarrow y \geq 1}$$

Universität St.Gallen

Ist dies eine Dichtefunktion?

$$\frac{1}{2}\int_1^\infty e^{-\frac{y-1}{2}}\,dy = \frac{-1}{\cancel{2}}\cdot \cancel{2}\; e^{-\frac{y-1}{2}}\Bigg|_1^\infty = 1 \quad \checkmark$$

$$\text{Also:} \rightarrow E[Y] = \frac{1}{2}\int_1^\infty y\; e^{-\frac{y-1}{2}}\,dy \underset{\uparrow \text{P.I.}}{=} -\frac{1}{2}\cdot 2\;\cdot y \cdot e^{-\frac{y-1}{2}}\Bigg|_1^\infty$$

$$+\int_1^\infty e^{-\frac{y-1}{2}}\,dy = 1+2\;\cdot\left(-e^{-\frac{y-1}{2}}\right)\Bigg|_1^\infty = 3$$

Beispiel 3.4.7: Sei X eine Zufallsvariable mit

X	-1	0	1.5	2
f(x)	0.3	0.1	0.4	0.2

$$E[X] = (-1)\,0.3 + 0\cdot 0.1 + 1.5\cdot 0.4 + 2\cdot 0.2 = 0.7$$

$$E[X+3] = E[X]+3 = 3.7$$

$$E[4X] = 4\cdot E[X] = 2.8$$

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

Beispiel 3.4.7: Sei X eine Zufallsvariable mit

X	-1	0	1.5	2
$f(x)$	0.3	0.1	0.4	0.2

$$\Rightarrow \; E[X] = (-1)\cdot 0.3 + 0\cdot 0.1 + 1.5\cdot 0.4 + 2\cdot 0.2 = 0.7$$

$$E[X+3] = E[X] + 3 = 3.7$$

$$E[4X] = 4\,E[X] = 4\cdot 0.7 = 2.8$$

X : Gewinn

Beispiel 3.4.8:

$$E[X] = 10 \cdot \frac{1}{3} + 20 \cdot \frac{1}{3} + 40 \cdot \frac{1}{6} + 80 \cdot \frac{1}{6}$$

$$= \underline{\underline{30}}$$

Spieler 1 verspricht Spieler 2, ihm beim Würfelspiel die folgende Gewinne auszuzahlen:

10 Rappen, falls 1 oder 2 gewürfelt wird;
20 Rappen, bei 3 oder 4;
40 Rappen, bei 5; und
80 Rappen, bei 6.

Wie viel muss Spieler 2 vor jeder Runde an Spieler 1 bezahlen, damit das Spiel **fair** ist?

***Beispiel 3.4.8** (Fortsetzung):*

"**Faires Spiel**" soll dabei heissen, dass der Einsatz gleich dem "durchschnittlichen Gewinn" sein soll.

$\longrightarrow$

$X=x$	10	20	40	80
$f(x)$	$2/6$	$2/6$	$1/6$	$1/6$

Gewinn

Idealisierter durchschnittlicher Gewinn bei unendlich vielen Spielrunden ist

$$E[X] = 10 \cdot \frac{2}{6} + 20 \cdot \frac{2}{6} + 40 \cdot \frac{1}{6} + 80 \cdot \frac{1}{6} = 30,$$

d.h. der "faire Einsatz" in diesem Sinn ist 30 Rappen.

(→ Casino-Spiele sind nicht fair!)

Kapitel 9.5

3.5. Varianz

① Sei X eine Z.V. mit Erwartungswert $\mathbb{E}[X] =: \mu_X$. Dann heisst

$$\underbrace{\mathrm{Var}(X)}_{=:\sigma_X^2} := \mathbb{E}\left[(X - \underbrace{\mu_X}_{\mathbb{E}[X]})^2\right]$$

Varianz von X (angenommen das Integral existiert).

② Die Standartabweichung von X ist $\sigma_X := \sqrt{\mathrm{Var}(X)}$

Bemerkung:

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

Begründung:

$$\mathrm{Var}(X) = \mathbb{E}\left[(X-\mu_X)^2\right] = \mathbb{E}\left[X^2 - 2X\mu_X + \mu_X^2\right]$$

$$= \mathbb{E}[X^2] - 2\mu_X \underbrace{\mathbb{E}[X]}_{\mu_X} + \underbrace{\mathbb{E}[\mu_X^2]}_{\mu_X^2}$$

$$= \mathbb{E}[X^2] - \mu_X^2$$

3.5. Varianz

$g(x) = (x - \mu_x)^2$

Definition

Sei X eine Zufallsvariable und μ_x ihr Erwartungswert, dann heisst

$\mu_x = E[X]$

$$\mathrm{Var}(X) = \sigma_x^2 = V(X) = E[\underbrace{(X - \mu_x)^2}^{g(X)}]$$

Varianz der Zufallsvariablen X, falls die Summe bzw. das Integral existiert.

Die positive Wurzel aus der Varianz $\sigma_x = +\sqrt{V(X)}$ heisst ***Standardabweichung***.

Universität St.Gallen

Die so definierte Varianz einer Zufallsvariablen ist als ein **Streuungsparameter** der Verteilung anzusehen.

$$\mathrm{Var}(X) = \sigma_x^2 = \sum_{\text{alle } j} \underbrace{(x_j - \mu_x)^2}_{g(x_j)} p_j, \qquad \text{falls X diskret}$$

$$\mathrm{Var}(X) = \sigma_x^2 = \int_{\mathbb{R}} \underbrace{(x-\mu_x)^2}_{g(x)} f(x)dx, \qquad \text{falls X stetig}$$

Beispiel 3.5.1:

Sei X eine diskrete Zufallsvariable:

X = "Anzahl der Köpfe bei zwei Münzwürfen"

Dann:

X	0	1	2
f(x)	1/4	1/2	1/4

$$\mu_X = \mathbb{E}[X] = 0 \cdot 1/4 + 1 \cdot 1/2 + 2 \cdot 1/4 = 1$$

$$\underline{\underline{Var(X)}} = \mathbb{E}\left[(X - \underset{1}{\mu_X})^2\right] = (0-1)^2 \tfrac{1}{4} + (1-1)^2 1/2 + (2-1) 1/4$$

$$= 1/4 + 1/4 = \underline{\underline{1/2}}$$

$$\mathbb{E}[X^2] = 0^2 \cdot 1/4 + 1^2 \cdot 1/2 + 2^2 \cdot 1/4$$

$$= 1/2 + 1 = 3/2$$

$$Var(X) = \mathbb{E}[X^2] - \mu_X^2 = 3/2 - 1^2 = \underline{\underline{1/2}}$$

Beispiel 3.5.1:

Sei X eine diskrete Zufallsvariable:

X = “Anzahl der Köpfe bei zwei Münzwürfen“

Dann:

X	0	1	2
$f(x)$	$1/4$	$1/2$	$1/4$

und $\mu_x = 1$

$$\sigma^2_x = (0\text{-}1)^2 \cdot 1/4 + (1-1)^2 \cdot \frac{1}{2} + (2-1)^2 \cdot 1/4 = 1/2$$

Universität St.Gallen

Beispiel 3.5.2:

Sei wie vorher Y = “Differenz der Augenzahlen zweier Würfel“. (Siehe Beispiel 3.4.2)

Wir haben schon berechnet, dass

$$E[Y] = \frac{35}{18} \cdot \quad \text{Nun ist } V(Y) = \sigma_y^2 = 2.05247$$

Universität St.Gallen

Beispiel 3.5.3:

Sei X eine stetige Zufallsvariable mit Dichtefunktion

Var(X) = ?

$$f(x)=\begin{cases} c\left(x-\frac{1}{2}x^2\right), & 0<x<2 \\ 0 \quad , & \text{sonst.}\end{cases}$$

1. Bestimme c:

$\int_0^2 c\left(x-\frac{1}{2}x^2\right)dx \overset{!}{=} 1$

$\Rightarrow \ldots \Rightarrow c = 3/2$

2. Berechne: $\mathbb{E}[X]$

$\mathbb{E}[X] = \int_0^2 x\cdot f(x)\,dx = \int_0^2 \frac{3}{2}\cdot x\left(x-\frac{1}{2}x^2\right)dx$

$= \frac{3}{2}\underbrace{\int_0^2 x^2\,dx}_{\frac{x^3}{3}\Big|_0^2 = \frac{8}{3}} - \frac{3}{4}\underbrace{\int_0^2 x^3\,dx}_{\frac{x^4}{4}\Big|_0^2 = \frac{16}{4} = 4} = 1$

$$\mathrm{Var}(X) = \int_0^2 (x-\mu_X)^2 f(x)\,dx$$

$$= \frac{3}{2}\int_0^2 \underbrace{(x-1)^2}_{x^2-2x+1}\left(x-\frac{1}{2}x^2\right)dx$$

$$= \frac{3}{2}\left(\int_0^2 \left(x^3-\frac{1}{2}x^4-2x^2+x^3 + x-\frac{1}{2}x^2\right)dx\right)$$

$$= \frac{3}{2}\left(\int_0^2 \left(-\frac{1}{2}x^4+2x^3-\frac{5}{2}x^2+x\right)dx\right)$$

$$= \ldots \underline{\underline{1/5}}$$

Beispiel 3.5.3 (Fortsetzung):

Wie gross ist die Konstante c ?

$$c \cdot \int_0^2 \left(x - \frac{1}{2} x^2 \right) dx = c \left(\frac{x^2}{2} \Big|_0^2 - \frac{x^3}{6} \Big|_0^2 \right) = c \left(2 - \frac{8}{6} \right) = c \, \frac{2}{3} \overset{!}{=} 1$$

$$\Rightarrow \quad c = {}^{3}\!/_{2}$$

Beispiel 3.5.3 *(Fortsetzung):*

Dann:

$$E\left[X\right] = \frac{3}{2}\cdot\int_0^2\left(x^2 - \frac{1}{2}x^3\right)dx = \frac{3}{2}\cdot\left(\frac{x^3}{3}\bigg|_0^2 - \frac{x^4}{8}\bigg|_0^2\right) = \frac{3}{2}\cdot\left(\frac{8}{3} - 2\right) = 1$$

und

$$V\left(X\right) = \frac{3}{2}\int_0^2\left(x\text{-}1\right)^2\cdot\left(x\text{-}\frac{1}{2}x^2\right)dx = \frac{3}{2}\int_0^2\left(x\text{-}\frac{5}{2}x^2 + 2x^3 - \frac{1}{2}x^4\right)dx$$

$$= \frac{3}{2}\left(2\text{-}\frac{40}{6} + \frac{16}{2} - \frac{32}{10}\right) = 1/5; \quad \sigma_x = \sqrt{1/5} = 0.4472.$$

Rechenregeln für Varianzen

Satz: Sei X eine Z.V. mit Varianz $\mathrm{Var}(X)$. Seien $a, b \in \mathbb{R}$. Dann gilt

$$\mathrm{Var}(\underbrace{a \cdot X + b}_{=: Y}) = a^2 \, \mathrm{Var}(X)$$

Bew: **IDEE:** $\mathrm{Var}(Y) = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2$

$$\mathrm{Var}(aX+b) = \mathbb{E}[\underbrace{(aX+b)^2}_{a^2X^2 + 2abX + b^2}] - \underbrace{\mathbb{E}[aX+b]^2}_{\underbrace{(a\,\mathbb{E}[X]+b)^2}_{a^2\,\mathbb{E}[X]^2 + \cancel{2ab\,\mathbb{E}[X]} + \cancel{b^2}}}$$

$$a^2\,\mathbb{E}[X^2] + \cancel{2ab\,\mathbb{E}[X]} + \cancel{b^2}$$

$$= a^2 \left(\mathbb{E}[X^2] - \mathbb{E}[X]^2 \right)$$

$$= a^2 \, \mathrm{Var}(X)$$

QED

Rechenregeln für Varianzen

Satz:

Wenn X eine Zufallsvariable mit der Varianz V(X) ist und a und b reelle Konstanten sind, dann hat $Y = aX+b$ die Varianz $V(Y) = a^2 \cdot V(X)$ und die Standardabweichung $\sigma_y = |a| \cdot \sigma_x$.

Bew: $\mathrm{Var}(aX+b) = \mathbb{E}\left[\left(aX+b-\mathbb{E}[aX+b]\right)^2\right]$

$= \mathbb{E}\left[\left(aX+\cancel{b}-a\mathbb{E}[X]-\cancel{b}\right)^2\right] = a^2\,\mathbb{E}\left[\left(X-\mathbb{E}[X]\right)^2\right]$

$\boxtimes$

Beispiel 3.5.4:

Gegeben sei die Zufallsvariable X mit

X	6060	6100	6140
$f_X(x)$	0.2	0.3	0.5

Ges: Var(X)

Definiere:

$$\boxed{Y := \frac{X - 6100}{40}} \leftrightarrow \boxed{X = 40Y + 6100}$$

standardisierte Z.V.

Y	−1	0	1
f_Y	0.2	0.3	0.5

$$\mathbb{E}[Y] = -1 \cdot 0.2 + 1 \cdot 0.5 = 0.3$$

$$Var(Y) = (-1-0.3)^2 \, 0.2 + (0-0.3)^2 \cdot 0.3 + (1-0.3)^2 \, 0.7$$

$$= 0.61$$

$$Var(X) = Var(40Y + 6100) = 40^2 \cdot \underbrace{Var(Y)}_{0.61}$$

$$= \underline{\underline{976}}$$

Beispiel 3.5.4:

Gegeben sei die Zufallsvariable X mit

X	6060	6100	6140
$f(x)$	0.2	0.3	0.5

Um die Varianz V(X) zu berechnen, nehmen wir zuerst eine Transformation $Y = \dfrac{X-6100}{40}$,

deren Verteilung offensichtlich

Y	-1	0	1
$f(y)$	0.2	0.3	0.5

ist.

Beispiel 3.5.4 *(Fortsetzung):*

Wir erhalten:

$$E[Y] = 0.3$$

$$V(Y) = (-1-0.3)^2 \cdot 0.2 + (0-0.3)^2 \cdot 0.3$$

$$+ \quad (1-0.3)^2 \cdot 0.5 \;=\; 0.61$$

$$\Rightarrow V(X) = V(6100 + 40Y) = 40^2 \cdot V(Y)$$

$$= 40^2 \cdot 0.61 \;=\; 976$$

Satz (Vereinfachte Berechnung der Varianz)

Vereinfachte Berechnung: $V(X) = E[X^2] - \mu_x^2$.

Beispiel 3.5.5:

Sei wie vorher Y = "Differenz der Augenzahlen zweier Würfel" (siehe Beispiel 3.5.2).

Schon berechnet: $E[Y] = \frac{35}{18}$, $V(Y) = 2.05247$

Nun: $E[Y^2] = 0^2 \cdot 1/6 + 1^2 \cdot \frac{10}{36} + ... = \frac{210}{36}$

$$\Rightarrow V(Y) = \frac{210}{36} - \left(\frac{70}{36}\right)^2 = 5.8\overline{3} - 3.78086 = 2.05247. \checkmark$$

Universität St.Gallen

Satz von Steiner: Sei X eine Z.V. mit $\mathbb{E}[X] = \mu$ und sei d eine beliebige reelle Zahl. Dann gilt:

$$\mathrm{Var}(X) = \mathbb{E}\left[(X-d)^2\right] - (\mu - d)^2$$

Bew: Übungsaufgabe

Steinerscher Satz:

Ist X eine Zufallsvariable mit E[X] = μ und ist d eine reelle Zahl, dann gilt

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

<u>Bew</u>: $\mathbb{E}[X^2 - 2dX + \cancel{d^2}] - (\mu^2 - 2\mu d + \cancel{d^2})$

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

$\boxtimes$

Universität St.Gallen

Kapitel 9.6

Definition (Standardisieren)

Sei X eine Z.V. mit $\mu := \mathbb{E}[X]$ und $\sigma^2 := \operatorname{Var}(X) > 0$. Die transformierte Z.V.

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

heisst ==standardisiert==.

$$\mathbb{E}[Z] = \mathbb{E}\left[\frac{X}{\sigma} - \frac{\mu}{\sigma}\right] = \underbrace{\mathbb{E}\left[\frac{X}{\sigma}\right]}_{\mathbb{E}[aX]} - \underbrace{\mathbb{E}\left[\frac{\mu}{\sigma}\right]}_{\mu/\sigma}$$

$a := 1/\sigma$

$$= \frac{\overset{=\mu}{\mathbb{E}[X]}}{\sigma} - \frac{\mu}{\sigma}$$

$$= 0$$

$a = 1/\sigma$

$b = -\frac{\mu}{\sigma}$

$$\operatorname{Var}(Z) = \operatorname{Var}\left(\frac{X}{\sigma} - \frac{\mu}{\sigma}\right)$$

$$= \operatorname{Var}(aX + b)$$

$$= a^2 \operatorname{Var}(X)$$

$$= \frac{1}{\sigma^2}\operatorname{Var}(X) = 1$$

3.6. Standardisieren

Definition

Sei X eine Zufallsvariable mit $\mu_X = \mu$ und $\sigma_X = \sigma > 0$. Dann heisst die transformierte Zufallsvariable

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

Jede standardisierte Zufallsvariable hat Erwartungswert 0 und Varianz 1:

$$X \xrightarrow{\text{Verschiebung}} Y = X - \mu \xrightarrow{\text{Streckung}} Z = \frac{1}{\sigma} Y = \frac{X - \mu}{\sigma}$$

$E[X] = \mu$	$E[Y] = 0$	$E[Z] = 0$
$V(X) = \sigma^2$	$V(Y) = \sigma^2$	$V(Z) = 1$

<u>Beispiel:</u> $X = \#$ Köpfe bei zwei Münzwürfen

$$\mathbb{E}[X] = 1 \quad , \quad Var(X) = \ldots = 1/2$$

Standardisierung: $Z = \frac{X-\mu}{\sigma} = \frac{X-1}{\sqrt{1/2}} = \sqrt{2}\,(X-1)$

z	$-\sqrt{2}$	0	$\sqrt{2}$
$f_Z(z)$	$1/4$	$1/2$	$1/4$

$$\mathbb{E}[Z] = -\sqrt{2}\cdot 1/4 + 0\cdot 1/2 + \sqrt{2}\cdot 1/4 = 0 \quad \checkmark$$

$$Var(Z) = (\sqrt{2})^2 \cdot 1/4 + 0^2 \cdot 1/2 + (\sqrt{2})^2 \cdot 1/4$$

$$= 2/4 + 2/4 = \underline{1} \quad \checkmark$$

Beispiel 3.6.1:

X = “Anzahl Köpfe bei zwei Münzwürfen“

Man kann berechnen: E(X) = 1 , V(X) = ½

Standardisierung: $Z=\dfrac{X-\mu}{\sigma}=\dfrac{X-1}{\sqrt{1/2}}$

Wir erhalten für Z

Z	$-\sqrt{2}$	0	$\sqrt{2}$
$f(z)$	$1/4$	$1/2$	$1/4$

und $E[Z]=0,\quad V(Z)=1.$

Universität St.Gallen

Kolloquium
Wed 22.10, 8:15
01-014

Kapitel 11

Universität St.Gallen

Gestern:

Varianz

- $V(X) = \underbrace{\mathrm{Var}(X)}_{\sigma_X^2} := \mathbb{E}\left[(X - \mathbb{E}[X])^2 \right]$

$$= \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

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

Aufgabe: Seien X, Y zwei unabhängige ZV

zeige $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

Bew: $\mathrm{Var}(X+Y) = \mathbb{E}[\underbrace{(X+Y)^2}_{X^2+2XY+Y^2}] - (\underbrace{\mathbb{E}[X+Y]}_{\mathbb{E}[X]+\mathbb{E}[Y]})^2$

da X, Y unabhängig

$$= \underbrace{\mathbb{E}[X^2+2XY+Y^2]}_{\mathbb{E}[X^2] + \cancel{2\,\mathbb{E}[XY]} + \mathbb{E}[Y^2]} - \left(\mathbb{E}[X]^2 + \cancel{2\,\mathbb{E}[X]\mathbb{E}[Y]} + \mathbb{E}[Y]^2 \right)$$

$\mathbb{E}[XY] = \cancel{\mathbb{E}[X]\mathbb{E}[Y]}$

$$= \mathbb{E}[X^2] - \mathbb{E}[X]^2 + \mathbb{E}[Y^2] - \mathbb{E}[Y]^2$$

$$= \mathrm{Var}(X) + \mathrm{Var}(Y)$$

Regel: $\mathrm{Var}(aX+b) = a^2\,\mathrm{Var}(X)$

stetig & diskret

Standardisieren: X Z.V.

$$Z := \frac{X - \mathbb{E}[X]}{\sqrt{V(X)}}$$

$\Rightarrow \mathbb{E}[Z] = 0$ & $V(Z) = 1$

4. Stochastische Modelle und spezielle Verteilungen

Einige Verteilungen haben in der Statistik und ihrer Anwendung eine fundamentale Bedeutung.

Jede der speziellen Verteilungen stellt eine ganze ***Familie*** von Verteilungen dar.

Die einzelnen Mitglieder erhält man durch Festlegung ihrer ***Parameter***.

Teil I: Wahrscheinlichkeitsrechnung

1. Grundlagen der Wahrscheinlichkeit
   - 1.1. Ereignisse, Ereignisraum und Ereignismenge
   - 1.2. Das Rechnen mit Ereignissen
   - 1.3. Der Wahrscheinlichkeitsbegriff
   - 1.4. Axiomatik der Wahrscheinlichkeitstheorie
   - 1.5. Wichtige Regeln der Wahrscheinlichkeitsrechnung
   - 1.6. Wahrscheinlichkeitsräume
   - 1.7. Bedingte Wahrscheinlichkeit und stochastische Unabhängigkeit
   - 1.8. Totale Wahrscheinlichkeit
   - 1.9. Das Bayes-Theorem

2. Elementare Kombinatorik
   - 2.1. Fakultät und Binomialkoeffizient
   - 2.2. Das Fundamentalprinzip der Kombinatorik
   - 2.3. Permutationen
   - 2.4. Kombinationen
   - 2.5. Variationen

3. Zufallsvariablen
   - 3.1. Die Verteilungsfunktion
   - 3.2. Diskrete Zufallsvariablen
   - 3.3. Stetige Zufallsvariablen
   - 3.4. Erwartungswerte von Zufallsvariablen
   - 3.5. Varianz
   - 3.6. Standardisieren

4. Stochastische Modelle und spezielle Verteilungen
   - 4.1. Gleichförmige Verteilung (diskret)
   - 4.2. Bernoulli-Verteilung (diskret)
   - 4.3. Binomialverteilung (diskret)
   - 4.4. Poisson-Verteilung (diskret)
   - 4.5. Rechteckverteilung (stetig)
   - 4.6. Exponentialverteilung (stetig)
   - 4.7. Normalverteilung (stetig)

5. Mehrdimensionale Zufallsvariablen
   - 5.1. Gemeinsame Verteilung und Randverteilungen
   - 5.2. Bedingte Verteilungen und stochastische Unabhängigkeit
   - 5.3. Kovarianz und Korrelationskoeffizient
   - 5.4. Summe von zwei oder mehreren Zufallsvariablen

6. Der zentrale Grenzwertsatz

Teil II: Statistik

7. Beschreibende/Deskriptive Statistik
   - 7.1. Häufigkeitsverteilung, Histogramm und Verteilungsfunktion
   - 7.2. Messzahlen zur Beschreibung statistischer Verteilungen
   - 7.3. Boxplot
   - 7.4. Quantile-Quantile Plot
   - 7.5. Streudiagramm

8. Schätzung unbekannter Parameter: Punktschätzung
   - 8.1. Intuitiv heuristische Ansätze für Schätzfunktionen
   - 8.2. Eigenschaften von Punktschätzungen
   - 8.3. Methoden zur Konstruktion von Schätzfunktionen

9. Intervallschätzungen-Konfidenzintervalle
   - 9.1. Konzept des Konfidenzintervalls
   - 9.2. Ableitung von Konfidenzintervallen (bei grossen Stichproben)
   - 9.3 Zusammenhang mit Hypothesentests

10. Hypothesentests
   - 10.1 Arten von Hypothesen
   - 10.2 Kritischer Bereich und Teststatistik
   - 10.3 Gütefunktion und Arten von Fehlern
   - 10.4 Der p-Wert

Teil III: Aufgaben

Universität St.Gallen

Kapitel 11.1

Gleichverteilung

$\mathbb{P}(X = x_3) = \frac{1}{m}$

4.1. Gleichförmige Verteilung (diskret)

$f(x_3) = \frac{1}{m}$

Die Zufallsvariable X habe m (also endlich viele) Ausprägungen, wobei alle Ausprägungen mit der gleichen Wahrscheinlichkeit vorkommen:

Massenfunktion

X	x_1	x_2	x_3	…	x_{m-1}	x_m
f(X)	1/m	1/m	1/m	….	1/m	1/m

Notation: $f_{GL}(x; m)$

m: Parameter, # Werte die X annehmen kann

Parameter der Familie

$\Rightarrow f_{GL}(x; m) = \frac{1}{m}$

Beispiel (Würfel)

$$X = \text{Augenzahl} \to \{x_1 = 1, x_2 = 2, x_3 = 3, \dots, x_6 = 6\}$$

$\Longrightarrow$ Massenfunktion $f_{GL}(x;6) = \frac{1}{6}$, fall $x \in \{1,2,\dots,6\}$

$$\Longrightarrow f_{GL}(x;6) = \begin{cases} \frac{1}{6}, & \text{für } x \in \{1,2,\dots,6\} \\ 0, & \text{sonst} \end{cases}$$

Verteilungsfunktion: $F: \mathbb{R} \to [0,1]$

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

falls $x \in \{1,2,\dots,6\}$

$$f_{GL}(x;6) = \frac{1}{2}$$

Beispiel 4.1.1:

Für m = 6 ist es die Massenfunktion von

X = "Augenzahl beim klassischen Würfelwurf":

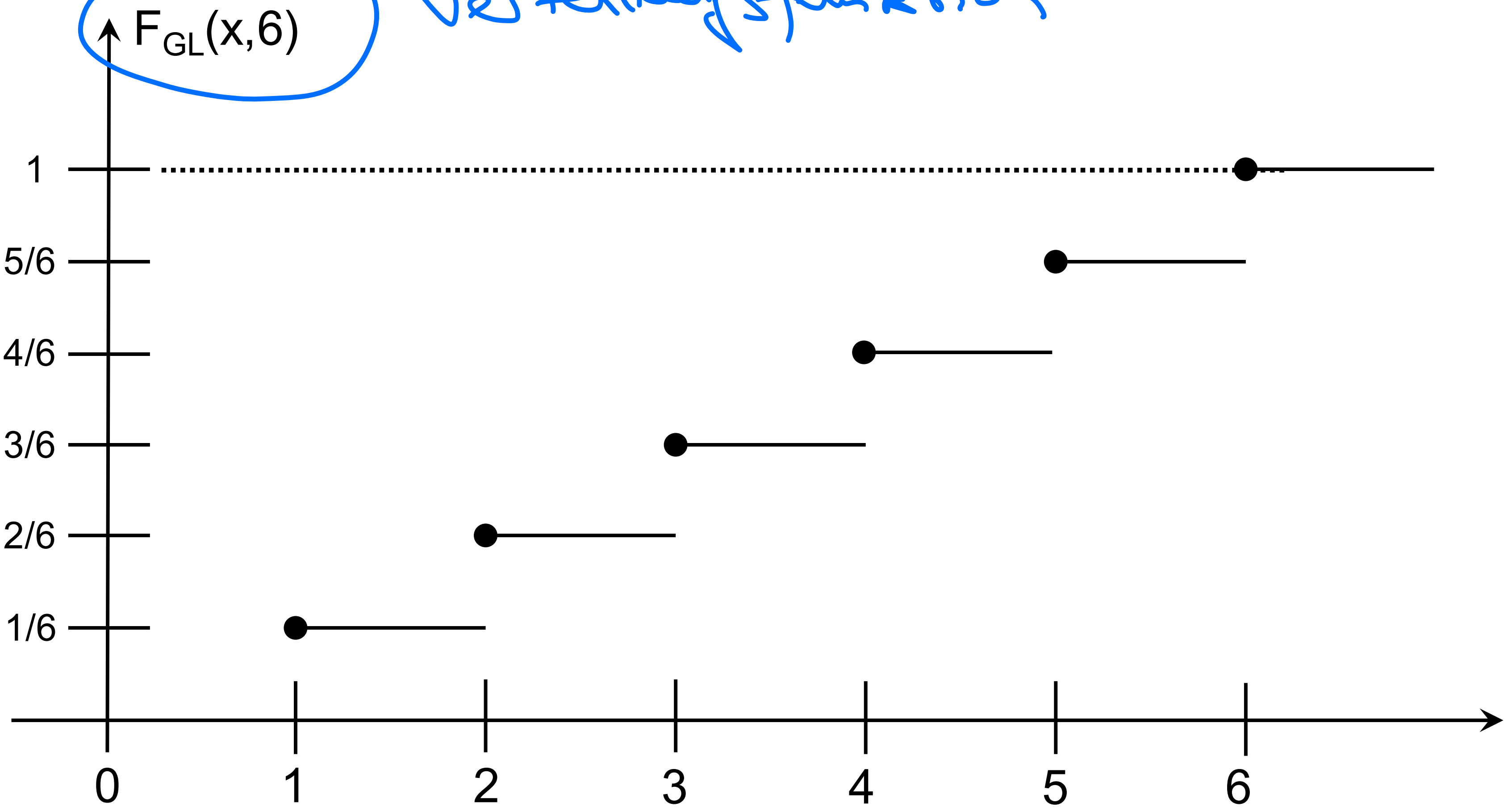
$$f_{GL}(x;6) = \begin{cases} \frac{1}{6}, & x = 1, 2, \ldots, 6 \\ 0, & \text{sonst} \end{cases}$$

Universität St.Gallen

Beispiel 4.1.1 *(Fortsetzung):*

Verteilungsfunktion?

Verteilungsfunktion



Würfelbeispiel

$$\mathbb{E}[X] = \sum_{i=1}^{6} x_i \, \mathbb{P}(X = x_i)$$

$$= \sum_{i=1}^{6} i \cdot \frac{1}{6} = \frac{1}{\cancel{6}} \sum_{i=1}^{6} i = \frac{6+1}{2}$$

$$\frac{\cancel{6}(6+1)}{2}$$

$$= 3.5$$

$$\sum_{i=1}^{m} i = \frac{m(m+1)}{2}$$

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

$$(3.5)^2$$

$$\frac{m(m+1)(2m+1)}{6}$$

$$\mathbb{E}[X^2] = \sum_{i=1}^{m} i^2 \cdot \frac{1}{m} = \frac{1}{m} \sum_{i=1}^{m} i^2$$

$$\overset{m=6}{=} \dots = \frac{91}{6}$$

$$\Rightarrow \mathrm{Var}(X) = \frac{91}{6} - (3.5)^2 = \dots = 2.9\dots$$

Beispiel: Gleichverteilung auf $1, 2, \ldots, m$

$$\Rightarrow f_{GL}(x, m) = \begin{cases} \frac{1}{m}, & \text{falls } x \in \{1, 2, \ldots, m\} \\ 0, & \text{sonst} \end{cases}$$

$$\Rightarrow \boxed{\begin{aligned} \mathbb{E}[X] &= \frac{1+m}{2} \\ Var(X) &= \frac{m^2-1}{12} \end{aligned}}$$

Bsp 3: Sei X = Summe der Augenzahlen von 2 Würfeln

X	x_1 2	x_2 3	x_3 4	5	$\ldots$	x_{11} 12
$f(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$			

$\longrightarrow$ X ist nicht Gleichverteilt!

Beispiel 4.1.1 *(Fortsetzung):*

Erwartungswert?

$$E[X] = \sum_{i=1}^{m} x_i \cdot p_i = \sum_{i=1}^{m} x_i \cdot P[X = x_i] = \sum_{i=1}^{m} x_i \cdot \frac{1}{m} = \frac{1}{m} \cdot \sum_{i=1}^{m} x_i$$

$\rightarrow$ Bsp: $E[X] = \frac{1}{6} \cdot \sum_{i=1}^{6} i = \frac{1}{\cancel{6}} \cdot \frac{\cancel{6} \cdot (6+1)}{2} = 3.5$

$$\sum_{i=1}^{k} i^2 = \frac{k(k+1)(2k+1)}{6}$$

Beispiel 4.1.1 *(Fortsetzung):*

Varianz?

$$E\left[X^2\right] = \frac{1}{m} \cdot \sum_{i=1}^{m} x_i^2 = \frac{1}{6} \cdot \sum_{i=1}^{6} i^2 = \frac{1}{6} \cdot \frac{(6+1)(6 \cdot 2+1) \cdot \cancel{6}}{\cancel{6}} = \frac{7 \cdot 13}{6} = \frac{91}{6}$$

$$V(X) = E\left[X^2\right] - E[X]^2 = \frac{1}{m} \cdot \sum_{i=1}^{m} x_i^2 - \left(\frac{1}{m} \cdot \sum_{i=1}^{m} x_i\right)^2$$

$$= \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{91}{6} - \frac{49}{4} = \frac{182-147}{12} = \frac{35}{12} = 2.9167$$

$$\sigma_X = \sqrt{V(X)} = \sqrt{2.9167} = 1.7078$$

Merke:

Die Summe von unabhängigen diskret gleichverteilten Zufallsvariablen ist nicht diskret gleichverteilt (keine reproduktive Eigenschaft).

Bsp: Summe der Augenzahlen beim Würfeln mit zwei oder mehr Würfeln.

$$\begin{array}{ccccc} 2 & 3 & 4 & K & 12 \\ \frac{1}{36} & \frac{2}{36} & \frac{3}{36} & K & \end{array} \quad \leftarrow \text{Keine Gleichverteilung!}$$

$$\text{Kein Laplace-Modell}$$

Beispiel: Gleichverteilung auf

$1, 2, \ldots, n$

$$\boxed{X \sim GL(n)}$$

$$f_{GL}(x; n) = \frac{1}{n}$$

Parameter

$$X \sim f_{GL}(\cdot\,; n)$$

Summenformel

$$\mathbb{E}[X] = \sum_{i=1}^{n} i \cdot \underbrace{\frac{1}{n}}_{\mathbb{P}(X=i)} = \frac{1}{n} \overbrace{\sum_{i=1}^{n} i}^{n \cdot \frac{n+1}{2}} = \frac{1+n}{2}$$

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \underbrace{\mathbb{E}[X]^2}_{\frac{(1+n)^2}{4}}$$

$$\cdot\ \mathbb{E}[X^2] = \sum_{i=1}^{n} i^2\, \mathbb{P}(X=i) = \sum_{i=1}^{n} i^2 \frac{1}{n} = \frac{1}{n} \overbrace{\sum_{i=1}^{n} i^2}^{\frac{n(n+1)(2n+1)}{6}}$$

$$= \frac{(n+1)(2n+1)}{6}$$

$$\Rightarrow \mathrm{Var}(X) = \frac{2(n+1)(2n+1)}{12} - \frac{3(1+n)^2}{12}$$

$$= \frac{4n^2+6n+2-3n^2-6n-3}{12} = \frac{n^2-1}{12}$$

Kapitel 11.2

4.2. Bernoulli-Verteilung (diskret)

Die Zufallsvariable X habe nur die beiden Ausprägungen 0 und 1.

⟶ Zufallsexperiment mit zwei Ausgängen A und $\bar{A}$, wobei A als **Erfolg** bezeichnet werde und $\bar{A}$ als **Misserfolg**. Das Ereignis A trete mit einer bestimmten **Erfolgswahrscheinlichkeit** (Parameter) p auf.

$p \in [0,1]$

Solche Experimente heissen

Bernoulli-Experimente.

$$\mathbb{P}(X=1) =: p$$

Erfolgswahrscheinlichkeit einer Bernoulli Zufallsvariable

$$\mathbb{P}(X=0) = (1-p) =: q$$

$$f_{Ber}(x;p) = \begin{cases} 1-p, & \text{falls } x=0 \\ p, & \text{falls } x=1 \\ 0, & \text{sonst} \end{cases}$$

Massenfkt. einer Bernoulli-verteilten Z.V.

Sei $\boxed{X \sim Ber(p)}$

Sei X eine Z.V. mit Massenfunktion $f_{Ber}(\cdot\,;p)$

$$\mathbb{E}[X] = 0 \cdot \underbrace{\mathbb{P}(X=0)}_{1-p} + 1 \cdot \underbrace{\mathbb{P}(X=1)}_{p} = p$$

$$Var(X) = \underline{\mathbb{E}[X^2]} - \underbrace{\mathbb{E}[X]^2}_{p^2} = p(1-p)$$

$$0^2\, \mathbb{P}(X=0) + 1^2\, \mathbb{P}(X=1) = p$$

Definition

Eine diskrete Zufallsvariable X mit der Massenfunktion

$$f_{Be}(x;p) = \begin{cases} 1-p, & x=0 \\ p, & x=1 \\ 0, & sonst \end{cases}$$

heisst **Bernoulli-Massenfunktion.**

Beispiel 4.2.1:

Um beim “Mensch ärgere dich nicht”-Spiel "herauskommen" zu können, muss man bei drei Würfen mindestens eine Sechs würfeln (“Erfolg”).

$\rightarrow$ Erfolgswahrscheinlichkeit?

$$p = 1 - P[\text{"keine Sechs in drei Würfen"}] = 1 - \left(\frac{5}{6}\right)^3 = \frac{216 - 125}{216} = \frac{91}{216}.$$

$$\Rightarrow \quad E[X] = \frac{91}{216} = 0.4213,$$

$$V(X) = \frac{91}{216} \cdot \left(1 - \frac{91}{216}\right) = 0.2438$$

$$\sigma_X = 0.4938$$

Kapitel11.3

4.3. Binomialverteilung (diskret)

Seien $Y_i \sim Ber(p)$, für $i=1,\dots,n$

unabhängig.

$$X := \sum_{i=1}^{n} Y_i$$

$$f_X(k) = \mathbb{P}(X=k) = \mathbb{P}\left(\sum_{i=1}^{n} Y_i = k\right)$$

$$= \underbrace{\binom{n}{k}}_{\text{\# Möglichkeiten } k \text{ aus } n \text{ auszuwählen}} p^k (1-p)^{n-k}$$

Erfolgsw'keit

\# Versuche

$$f_{Bin}(x; p, n) = \binom{n}{k} p^k (1-p)^{n-k}$$

für $x \in \{0,1,2,\dots,n\}$, $n \in \mathbb{N}$, $p \in [0,1]$

Sei $X \sim Bin(n,p)$

$$\mathbb{E}[X] = \sum_{x=0}^{n} x \cdot f_{Bin}(x; p, n)$$

$$= \mathbb{E}\left[\sum_{i=1}^{n} Y_i\right] = \sum_{i=1}^{n} \underbrace{\mathbb{E}[Y_i]}_{p} = n \cdot p$$

Y_i sind unabhängig

$$\underline{Var(X)} = Var\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} \underbrace{Var(Y_i)}_{p(1-p)} = \underline{np(1-p)}$$

Beispiel: Urne mit Zurücklegen

- Urne hat: 10 schwarze Kugel
  20 weisse Kugeln

- wir ziehen 4 Kugeln
- Sei $X = \#$ der gezogenen Kugeln die schwarz sind

$$\mathbb{P}(X=3) = ?$$

Definiere: $Y_i = \begin{cases} 1, & \text{falls Kugel } i \text{ schwarz ist} \\ 0, & \text{falls Kugel } i \text{ weiss ist} \end{cases}$

für $i = 1,2,3,4$

$$\mathbb{P}(Y_i = 1) = p = \frac{1}{3}$$

$\Longrightarrow$ Y_i ist Bernoulli verteilt mit Parameter p.

$$X = \sum_{i=1}^{4} Y_i$$

$$\Longrightarrow X \sim \mathrm{Bin}\left(n=4, p=\frac{1}{3}\right)$$

$$\mathbb{P}(X=3) = f_{Bin}\left(3\,;\, p=\frac{1}{3},\, n=4\right)$$

$$= \binom{4}{3}\left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^1 = \ldots$$

$$f_{Bin}(x\,;\,p,n) = \binom{n}{k} p^k (1-p)^{n-k}$$

Beispiel:

- Angenommen 35% der Wahlberechtigten stimmen links
- Es stimmen zufällig $n = 12$ Wähler ab

$\mathbb{P}$("Linkswähler die Mehrheit haben")

$$Y_i := \begin{cases} 1, & \text{falls Person } i \text{ Links stimmt} \\ 0, & \text{sonst} \end{cases}$$

$$X := \sum_{i=1}^{n} Y_i \longrightarrow \#\ \text{Linkswähler}$$

$$\mathbb{P}(X > 6) = ?$$

$n = 12$, $p = 0.35$

$$\mathbb{P}(X > 6) = \underbrace{\mathbb{P}(X=7)}_{f_{Bin}(x=7;\, n=12,\, p=0.35)} + \mathbb{P}(X=8) + \dots + \mathbb{P}(X=12)$$

4.3. Binomialverteilung (diskret)

Definition

Eine diskrete Zufallsvariable X mit der Massenfunktion

$$f_{Bi}\left(x;p,n\right)=\binom{n}{x}\cdot p^{x}\cdot\left(1-p\right)^{n-x},\qquad \begin{array}{c} x=0,1,2,\ldots,n \\ 0\leq p\leq 1 \end{array}$$

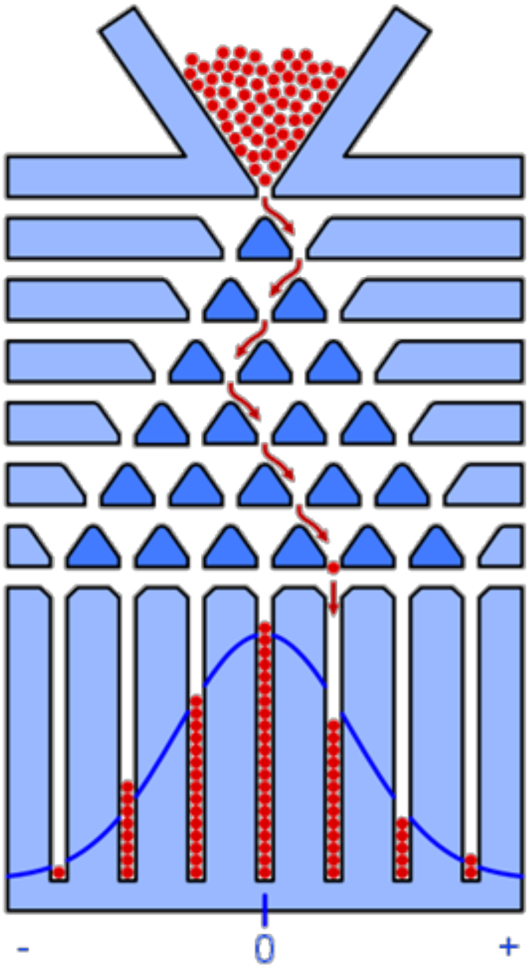
heisst **binomialverteilt**.

Herleitung der Binomialverteilung

Seien Y_i, i=1,…n, i.i.d Zufallvariablen mit Bernoulliverteilung mit Parameter p, dann ist die Summe X,

$$X = \sum_{i=1}^{n} Y_i$$

binomialverteilt mit Parametern n und p.



Beispiel 4.3.1: *Urnen-Modell mit Zurücklegen (Laplace-Wahrscheinlichkeit)*

In einer Urne befinden sich 10 schwarze und 20 weisse Kugeln. Daraus soll eine Zufallsstichprobe vom Umfang n=4 gezogen werden. Nachdem man die Farbe der einzelnen Kugeln notiert hat, wird sie sogleich wieder in die Urne zurückgelegt.

Beispiel 4.3.1 *(Fortsetzung):*

Nun: Sei X = "Anzahl der schwarzen Kugeln in der Stichprobe":

→ X kann die Werte: x = 0, 1, 2, 3, 4 annehmen

→ Erfolgswahrscheinlichkeit: $p = \frac{10}{30} = 1/3$; n = 4

→ $P[X=0] = \binom{4}{0} p^0 (1-p)^{4-0} = \frac{16}{81}$;

$P[X=1] = \binom{4}{1} p^1 (1-p)^{4-1} = \frac{32}{81}$;

Beispiel 4.3.1 *(Fortsetzung):*

Also:

Massenfunktion von X: $f_{Bi}\left(x; \frac{1}{3}, 4\right)$

$$\mathrm{E}[\mathrm{X}] = n \cdot p = 4 \cdot \frac{1}{3} = \frac{4}{3}$$

$$\mathrm{V}(\mathrm{X}) = n \cdot p \cdot (1 - p) = 4 \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{8}{9}$$

Beispiel 4.3.2:

Man stelle sich vor, 35% der Wahlberechtigten wollen links wählen. Mit einer Zufallsstichprobe vom Umfang $n=12$ soll die Wahlabsicht befragt werden.

Wie wird das Stichprobenergebnis ausfallen?

Beispiel 4.3.2 *(Fortsetzung):*

Erwartete "Linkswähler": $\mu = 12 \cdot 0.35 = 4.2$

Standardabweichung: $\sigma = \sqrt{12 \cdot 0.35 \cdot 0.65} = 1.6523$

$$\mathrm{P}\left(\text{"Linkswähler in der Stichprobe haben die Mehrheit"}\right)$$

$$= \mathrm{P}[X>6] = \sum_{x=7}^{12} \mathrm{P}[X=x] = \sum_{x=7}^{12} f_{Bi}\left(x;\, 0.35,\, 12\right) = 0.0846$$

Kapitel 11,5

4.4. Poisson-Verteilung (diskret)

Def: Eine Zufallsvariable mit Massenfunktion

$$f_{Poi}(x;\lambda) = \frac{\lambda^x}{x!} e^{-\lambda}$$

heisst Poisson-verteilt mit Parameter λ

$x \in \{0, 1, 2, 3, \ldots\}$

$\lambda \geq 0$

$X \sim Poi(\lambda)$

Übung: Zeigen Sie

$$\sum_{x=0}^{\infty} f_{Poi}(x;\lambda) = 1 \text{ für alle } \lambda$$

Übungsaufgabe:

$\mathbb{E}[X] = \ldots = \lambda$

$Var(X) = \ldots = \lambda$

Beispiel: (Roulette)

Ein Spieler setzt immer auf die Zahl 17.

Wie gross ist die W'keit dass er/sie genau 8 Spiele aus 200 Spielen gewinnt

$$X \sim \text{Bin}\left(n=200,\ p=\tfrac{1}{37}\right)$$

$$\mathbb{P}(X=8) = f_{\text{Bin}}(8;n,p) = \ldots = 0.0814$$

$$\mathbb{E}[X] = np = \frac{200}{37} \simeq 5.40\ldots$$

Wähle $\lambda = np = 5.40\ldots$

Angenommen $\tilde{X} \sim \text{Poi}(\lambda)$

$$\Longrightarrow \mathbb{E}[\tilde{X}] = \lambda$$

$$\mathbb{P}(\tilde{X}=8) = f_{\text{Poi}}(x=8;\lambda=5.4) = 0.0812$$

4.4. Poisson-Verteilung (diskret)

Definition

Die Massenfunktion der Poisson-Verteilung lautet

$$f_{Po}\left(x;\lambda\right) = \frac{\lambda^{x}}{x!} \cdot e^{-\lambda},$$

x = 0, 1, 2, ... (abzählbar unendliche Werte)

Universität St.Gallen

Beispiel 4.4.1: *Roulette*

Ein Spieler gehorcht der Eingebung, die Zahl 17 bringe ihm heute Glück. Er setzt beim Roulette unentwegt auf die Zahl 17. Wie gross ist die Wahrscheinlichkeit, dass er dabei genau 8 Mal bei 200 Spielen gewinnt?

→ Erfolgswahrscheinlichkeit:

$p = P[\text{Zahl } 17] = 1/37\,; \quad n=200$

Beispiel 4.4.1 *(Fortsetzung):*

Benutzen wir die Poisson-Verteilung $\lambda = np = 5.4054$,

damit finden wir: $f_{Po}(8\,;\lambda) = \dfrac{5.4054^8}{8!} e^{-5.4054} = 0.0812$

Bemerkung: Hätten wir die richtige Binomial-verteilung genommen:

$$f_{Bi}\left(8\,;\,\frac{1}{37}\,,\,200\right) = 0.0814.$$

p Erfolg

Beispiel 4.4.2:

Der professionelle Minigolfspieler E. Inlocher rühmt sich damit, dass er auch auf der schwierigsten Bahn nur zu 20% mit dem ersten Schlag nicht einlochen kann.

Heute sind zahlreiche Journalisten auf der Anlage erschienen. Der Spieler will mit einer Serie von 50 Versuchen zeigen, dass er wirklich so treffsicher ist.

Beispiel 4.4.2 *(Fortsetzung)*

Berechnen Sie die Wahrscheinlichkeit, dass Herr Inlocher nur in 38 von 50 Versuchen mit dem ersten Schlag ins Loch trifft: (a) exakt; (b) mit einer geeigneten Approximation.

a) X = "# Versuche nicht ins Loch" : $f_{Bi}(x\,;\,0.2\,,\,50)$

$$P[X=12] = 0.1033$$

b) Benützen Sie die Poisson-Verteilung: $\lambda = np = 10$

$$P[X=12] \approx P(\underset{\substack{\uparrow \\ \text{Poiss.}}}{Y} = 12) = 0.0948$$

Approximationsfehler: gross ! p = 0.2 zu gross!

Kapitel 11.7

Diskrete Verteilungen : (Wiederholung)

1. Gleichförmige Verteilung :

$X \sim GL(m)$

Parameter $m \in \mathbb{N}$

$X \in \{1, 2, \dots, m\}$

Massenfunktion: $f_{GL}(x; m) = \begin{cases} \frac{1}{m}, & x \in \{1, 2, \dots, m\} \\ 0, & \text{sonst} \end{cases}$

$\mathbb{E}[X] = \frac{n+1}{2}$, $Var(X) = \frac{n^2 - 1}{12}$

2. Bernoulli Verteilung :

$X \sim Ber(p)$

Erfolgsw'keit $p \in [0,1]$

$X \in \{0, 1\}$, $\mathbb{P}(X=1) = p$, $\mathbb{P}(X=0) = 1-p$

Massenfunktion: $f_{Ber}(x; p) = \begin{cases} 1-p, & x = 0 \\ p, & x = 1 \\ 0, & \text{sonst} \end{cases}$

$\mathbb{E}[X] = p$, $Var(X) = p(1-p)$

3. Binomial Verteilung :

#Versuche

Erfolgs-w'keit

$X \sim Bin\underbrace{(n, p)}_{\text{2 Parameter}}$

$X \in \{0, 1, \dots, n\}$

Massenfunktion: $f_{Bin}(x; p, n) = \binom{n}{x} p^x (1-p)^{n-x}$, für $x \in \{0, 1, \dots, n\}$

$\mathbb{E}[X] = n \cdot p$, $Var(X) = n\,p(1-p)$

4. Poisson Verteilung:

$X \sim \mathrm{Poi}(\lambda)$ — $\lambda \geq 0$ Parameter

$X \in \{0, 1, 2, \ldots\}$ ∞-viele Werte

Massenfunktion: $f_{\mathrm{Poi}}(x; \lambda) = \frac{\lambda^x}{x!} e^{-\lambda}$, für $x \in \{0, 1, 2, \ldots\}$

$\mathbb{E}[X] = \lambda$, $\mathrm{Var}(X) = \lambda$

Bsp: X: # Tore bei Fussballweltmeisterschaft

$X \sim \mathrm{Poi}(\lambda = 2.5)$

$\Longrightarrow \mathbb{P}(X=0) = 0.08$
$\mathbb{P}(X=1) = 0.20$
$\mathbb{P}(X=2) = 0.26$
$\mathbb{P}(X=3) = 0.21$
$\mathbb{P}(X=4) = 0.13$
$\vdots$

Bsp:

X: # Kaugummis auf einzelnem Pflasterstein

$X \sim \mathrm{Poi}(\lambda)$

Heute: Stetige Verteilungen

$a \leq b \quad , a,b \in \mathbb{R}$

4.5. Rechteckverteilung (stetig)

Definition

$X \in [a,b]$

Eine stetige Zufallsvariable X mit der Wahrscheinlichkeitsdichtefunktion

$$f_{\text{Re}}(x) = \begin{cases} \frac{1}{b-a}, & \textit{für } a \leq x \leq b \\ 0, & \textit{sonst} \end{cases}$$

heisst **Rechteckverteilt**.

Die Dichtefunktion ist im relevanten Bereich [a, b] **konstant**.

$\mathbb{E}[X] = \frac{1}{2}(a+b) \quad , \quad Var(X) = \frac{1}{12}(b-a)^2$

Universität St.Gallen

$$\mathbb{E}[X] = \int_{\mathbb{R}} x \cdot f_{Re}(x)\, dx = \int_a^b x \cdot \frac{1}{b-a}\, dx$$

$$= \frac{1}{b-a} \int_a^b x\, dx = \frac{1}{b-a} \left. \frac{x^2}{2} \right|_a^b \qquad \frac{b^2}{2} - \frac{a^2}{2}$$

$$= \frac{1}{\cancel{b-a}} \cdot \frac{1}{2} \cdot \underbrace{(b^2-a^2)}_{(\cancel{b-a})(b+a)} = \frac{a+b}{2}$$

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \qquad \frac{(a+b)^2}{4}$$

$$\frac{b^3}{3} - \frac{a^3}{3}$$

$$\mathbb{E}[X^2] = \frac{1}{b-a} \int_a^b x^2\, dx = \frac{1}{b-a} \cdot \left. \frac{x^3}{3} \right|_a^b$$

$$= \frac{1}{b-a} \cdot \frac{1}{3} (b^3 - a^3)$$

$$\Rightarrow \mathrm{Var}(X) = \frac{1}{3} \frac{1}{b-a} (b^3 - a^3) - \frac{(a+b)^2}{4}$$

$$= \dots \dots$$ $\Rightarrow$ siehe Notizen

$$= \frac{1}{12}(b-a)^2$$

Verteilungsfunktion

$$F_{Re}(x) = \mathbb{P}(X \leq x) = \int_a^x f_{Re}(y)\, dy \qquad f_{Re}(y) = \frac{1}{b-a}$$

$$= \frac{1}{b-a} \int_a^x 1\, dy = \frac{1}{b-a} \left. y \right|_a^x \qquad x-a$$

$$= \frac{x-a}{b-a}$$

$$\frac{d}{dx} F_{Re}(x) = \frac{1}{b-a} = f_{Re}(x)$$

$$\Rightarrow F_{Re}(x) = \begin{cases} 0 & , x < a \\ \frac{x-a}{b-a} & , x \in [a,b] \\ 1 & , x > b \end{cases}$$

$$\mathbb{E}[X] = \int_a^b x f_{Re}(x)\,dx = \frac{1}{b-a}\int_a^b x\,dx = \frac{1}{b-a}\frac{x^2}{2}\Big|_a^b$$

$$= \frac{1}{b-a}\left(\frac{b^2}{2}-\frac{a^2}{2}\right) = \frac{1}{2}\frac{1}{\cancel{b-a}}\underbrace{(b^2-a^2)}_{(\cancel{b-a})(b+a)} = \frac{1}{2}(a+b)$$

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

$$\mathbb{E}[X^2] = \frac{1}{b-a}\int_a^b x^2\,dx = \frac{1}{b-a}\frac{x^3}{3}\Big|_a^b$$

$$= \frac{1}{b-a}\cdot\frac{1}{3}(b^3-a^3)$$

$$\Rightarrow Var(X) = \frac{b^3-a^3}{3(b-a)} - \frac{(a+b)^2(b-a)}{4(b-a)}$$

$$= \frac{4b^3-4a^3-3\overset{a^2+2ab+b^2}{(a+b)^2}(b-a)}{12(b-a)}$$

$$= \frac{1}{12(b-a)}\left(4b^3-4a^3-3a^2b-6ab^2-3b^3 + 3a^3+6a^2b+3ab^2\right)$$

$$= \frac{1}{12(b-a)}\underbrace{\left(b^3-a^3+3a^2b^2-3ab^2\right)}_{(b-a)^3}$$

$$= \frac{(b-a)^2}{12}$$

Verteilungsfunktion:

$$F_{Re}(x) = \mathbb{P}(X \le x)$$

$$= \int_a^x f_{Re}(x)\,dx = \int_a^x \frac{1}{b-a}\,dx$$

$$= \frac{x-a}{b-a}$$

$$\Rightarrow F_{Re}(x) = \begin{cases} 0, & x<a \\ \frac{x-a}{b-a}, & x\in[a,b] \\ 1, & x>b \end{cases}$$

$$F(t) = \frac{t-a}{b-a} = \frac{t}{30}$$

$T \sim Re(a=0, b=30)$

Beispiel 4.5.1:

Zwischen Mitternacht und sechs Uhr morgens kommt der Bus gemäss Fahrplan jede halbe Stunde. Wie gross ist die Wahrscheinlichkeit, dass ein Fahrgast länger als 10 Minuten warten muss?

$a=0, b=30$

T = "Zeit zum nächsten Bus" ,

$F_{Re}(x) = \frac{x}{30}$

$$\text{Zufallsvariable mit } f_{Re}(t) = \begin{cases} \frac{1}{30}, & 0 \le t \le 30, \\ 0, & \text{sonst.} \end{cases}$$

$$\int_{10}^{30} f_{Re}(t)\,dt$$

$$\text{Also: } P[T \geq 10] = 1 - P[T \le 10] = 1 - F_{Re}(10) = 1 - \frac{10}{30} = \frac{2}{3}$$

$F_{Re}(10) = \frac{10}{30}$

$$\text{Ausserdem: } \underbrace{E[T]}_{\frac{30}{2}} = 15\,\text{min} \quad \text{und} \quad \underbrace{V(T)}_{\frac{30^2}{12} = \frac{10 \cdot 30}{4} = \frac{5 \cdot 15}{1} = \underline{\underline{75}}} = 75\,\text{min}^2$$

Verteilungsfunktion: $T \sim \mathrm{Re}(a,b)$

$$F(t) = \mathbb{P}(T \leq t) = \frac{t-a}{b-a}$$

Beispiel 4.5.2: *Wartezeit an der S-Bahn Station*

An einer S-Bahn Station fahren die Züge im 12-Minuten Takt. Ein Fahrgast, der den Fahrplan ignoriert, erscheint zu einem zufälligen Zeitpunkt an der Station.

X = Wartezeit

Zufallsvariable mit Wertevorrat W = [0,12] (Minuten)

Universität St.Gallen

Kapitel 11.8

$X \sim Ex(\lambda)$

4.6. Exponentialverteilung (stetig)

Definition

$X \in [0, \infty)$

Eine stetige Zufallsvariable X mit der Wahrscheinlichkeitsdichtefunktion

$$f_{Ex}(x;\lambda) = \lambda \cdot e^{-\lambda x}, \qquad 0 \leq x < \infty$$

$$\lambda > 0$$

heisst **exponentialverteilt**. mit Parameter $\lambda > 0$.

$\mathbb{E}[X] = 1/\lambda \quad , \quad Var(X) = 1/\lambda^2$

Frage: $\mathbb{E}[X] = ?$ $\quad F_{Ex}(x) = ?$

$$E[X] = \int_{\mathbb{R}} x\, f_{Ex}(x;\lambda)\, dx = \int_0^\infty x\, \lambda \cdot e^{-\lambda x}\, dx$$

$$= \lambda \int_0^\infty x\, e^{-\lambda x}\, dx$$

$$= \lambda \left(\underbrace{x\left(-\frac{1}{\lambda} e^{-\lambda x}\right)\Big|_0^\infty}_{=0} - \int_0^\infty 1 \cdot \left(-\frac{1}{\lambda} e^{-\lambda x}\right) dx \right)$$

$$= \cancel{\lambda} \cdot \frac{1}{\cancel{\lambda}} \int_0^\infty e^{-\lambda x}\, dx = \left(-\frac{1}{\lambda} e^{-\lambda x}\right)\Big|_0^\infty$$

$$= \frac{1}{\lambda}$$

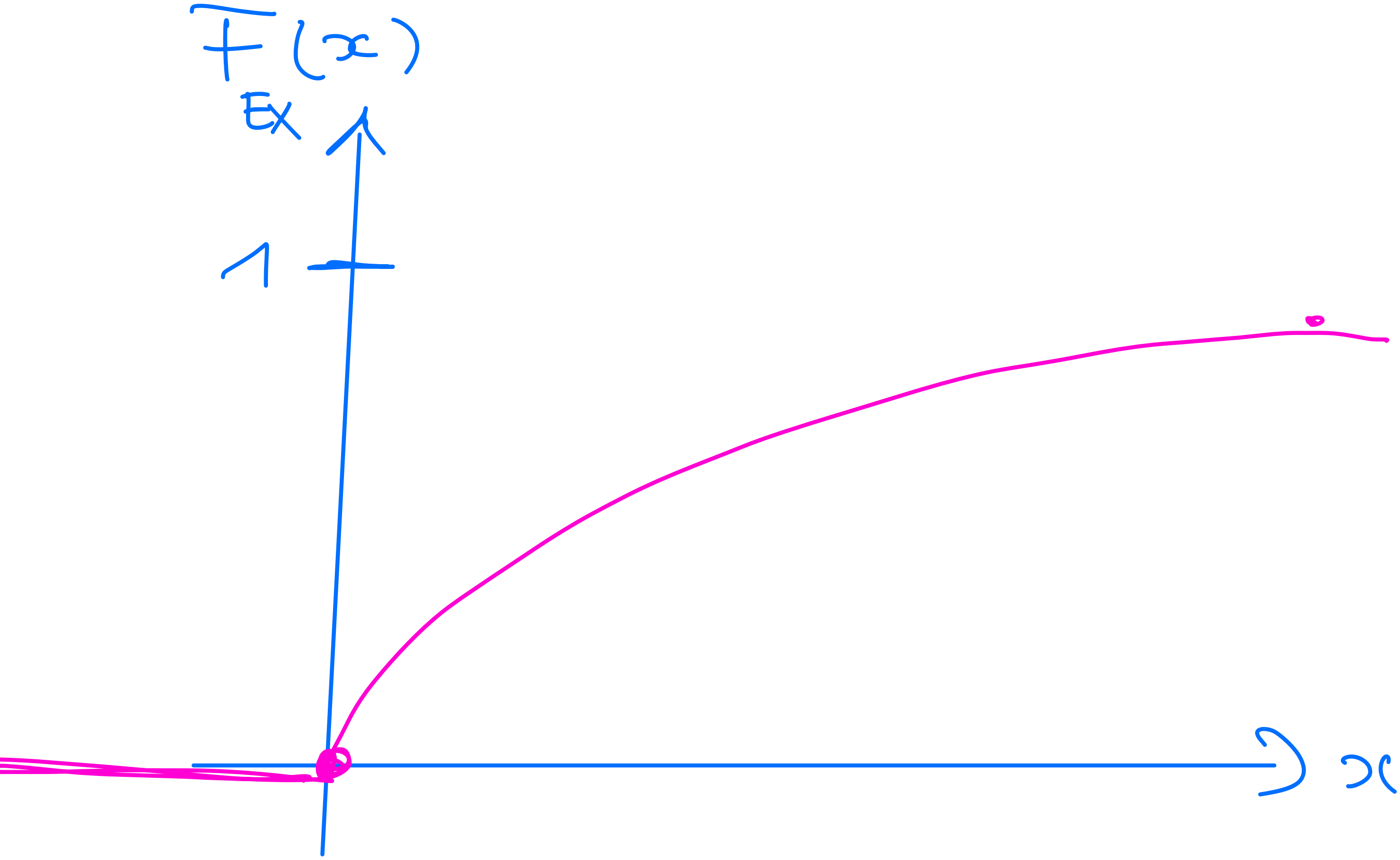
$$E[X] = 1/\lambda$$

$$F_{Ex}(x) = \underbrace{\int_0^x f_{Ex}(y;\lambda)\, dy}_{\mathbb{P}(X \leq x)}$$

$$= \int_0^x \lambda e^{-\lambda y}\, dy = -e^{-\lambda y}\Big|_0^x$$

$$= -e^{-\lambda x} + 1$$

$$= 1 - e^{-\lambda x}$$



Beh: Sei $X \sim Ex(\lambda) \Rightarrow \mathbb{E}[X] = 1/\lambda$

Bew:
$$\mathbb{E}[X] = \int_0^\infty x \lambda e^{-\lambda x} \, dx$$
$$= \lambda \int_0^\infty \underset{\downarrow}{x} \, \underset{\uparrow}{e^{-\lambda x}} \, dx$$
$$= \lambda \Big(\underbrace{x \left(-\tfrac{1}{\lambda} e^{-\lambda x}\right) \Big|_0^\infty}_{0} + \frac{1}{\lambda} \int_0^\infty e^{-\lambda x} dx \Big)$$
$$= \int_0^\infty e^{-\lambda x} \, dx = \left(-\frac{1}{\lambda}\right) e^{-\lambda x} \Big|_0^\infty$$
$$= \frac{1}{\lambda}$$
$\blacksquare$

$$\underline{F_{Ex}(x)} = \mathbb{P}(X \leq x) = \int_0^x \lambda e^{-\lambda y} \, dy$$
$$= -e^{-\lambda y} \Big|_0^x = -e^{-\lambda x} + 1$$
$$= \underline{1 - e^{-\lambda x}}$$

Beispiel 4.6.1: X = Lebensdauer der Glühbirne

Nach Angaben des Herstellers beträgt die mittlere Lebensdauer seiner neuen 100-Watt-Glühbirne 5000 Stunden. Als geeignetes Modell für die Approximation der Verteilung der Lebensdauer X verwenden wir:

$\lambda = 5000$

$$f_{Ex}(x) = \frac{1}{5000} e^{-\frac{x}{5000}}$$

$$\boxed{\lambda = \frac{1}{5000}}$$

Bemerkung: $$\lambda = \frac{1}{E[X]}$$

$$E[X] = \frac{1}{\lambda}$$

Universität St.Gallen

Beispiel 4.6.1 *(Fortsetzung):*

Wie gross ist die Wahrscheinlichkeit, dass eine Glühbirne:

(1) weniger als halb so lange brennt?

$$\mathbb{P}(X \leq 2500) = \int_0^{2500} f_{Ex}(x;\lambda)\,dx$$

$$= \underbrace{\underbrace{\underbrace{\mathbb{P}(X \leq 2500)}_{F_{Ex}(2500)}}_{1-e^{-\frac{2500}{5000}}}}_{1-e^{-1/2}} \approx 0.39$$

(2) mehr als doppelt so lange brennt?

$$\mathbb{P}(X \geq 10'000) = 1 - \mathbb{P}(X \leq 10'000)$$

$$= 1 - F_{Ex}(10'000) = 0.13$$

Beispiel 4.6.1 *(Fortsetzung):*

Wie gross ist die Wahrscheinlichkeit, dass eine Glühbirne:

(1) weniger als halb so lange brennt?

$$P\left[X \le 2500\right] = F_{Ex}\left(2500\right) = 1 - e^{-\frac{1}{2}} = 0.3935$$

(2) mehr als doppelt so lange brennt?

$$P\left[X > 10000\right] = 1 - F_{Ex}\left(10000\right) = 1 - \left(1 - e^{-2}\right) = 0.1353$$

Teil I: Wahrscheinlichkeitsrechnung

$X \in \mathbb{R}$, $X \sim \mathcal{N}(0,1)$

4.7. Normalverteilung (stetig)

Definition (Standard-Normalverteilung)

Eine stetige Zufallsvariable X mit der Dichtefunktion

$$f_X(x) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{x^2}{2}}, \quad \text{für} \quad -\infty < x < \infty$$

heisst **standardnormalverteilt**.

$\mathbb{E}[X] = 0$, $\mathrm{Var}(X) = 1$

Defition: $X \in \mathbb{R}$, $X \sim \mathcal{N}(\mu, \sigma^2)$

mit Parameter $\mu \in \mathbb{R}$, $\sigma > 0$

μ: Lageparameter

σ: Streuungsparameter

Dichtefunktion:

$$f_X(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Aufgabe: $\mathbb{E}[X] = \mu$, $\mathrm{Var}(X) = \sigma^2$

Allgemeine Normalverteilung

$X \sim \mathcal{N}(\mu, \sigma^2)$

Definition

Eine stetige Zufallsvariable X mit der Dichtefunktion

$$f_N\left(x;\, \mu, \sigma^2\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad -\infty < x < \infty$$

$\mu \in$ ¡

$\sigma > 0$

heisst **normalverteilt**.

$\mathbb{E}[X] = \mu$

Es wurde ein Lageparameter μ und eine Streuungsparameters $\sigma > 0$ eingeführt.

$\mathrm{Var}(X) = \sigma^2$

Eigenschaft: Seien $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ unabhängig. Definiere

$$\overline{X} := \frac{1}{n} \sum_{i=1}^{n} X_i$$

$$\Longrightarrow \overline{X} \sim \mathcal{N}\Big(\underbrace{\frac{1}{n}\sum_{i=1}^{n} \mu_i}_{\overline{\mu}}, \underbrace{\frac{1}{n^2}\sum_{i=1}^{n} \sigma_i^2}_{\overline{\sigma}^2}\Big)$$

Eigenschaft der Normalverteilung:

Seien X_i für $i = 1, \dots, n$ Zufallsvariablen mit $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, die unabhängig sind. Dann gilt für

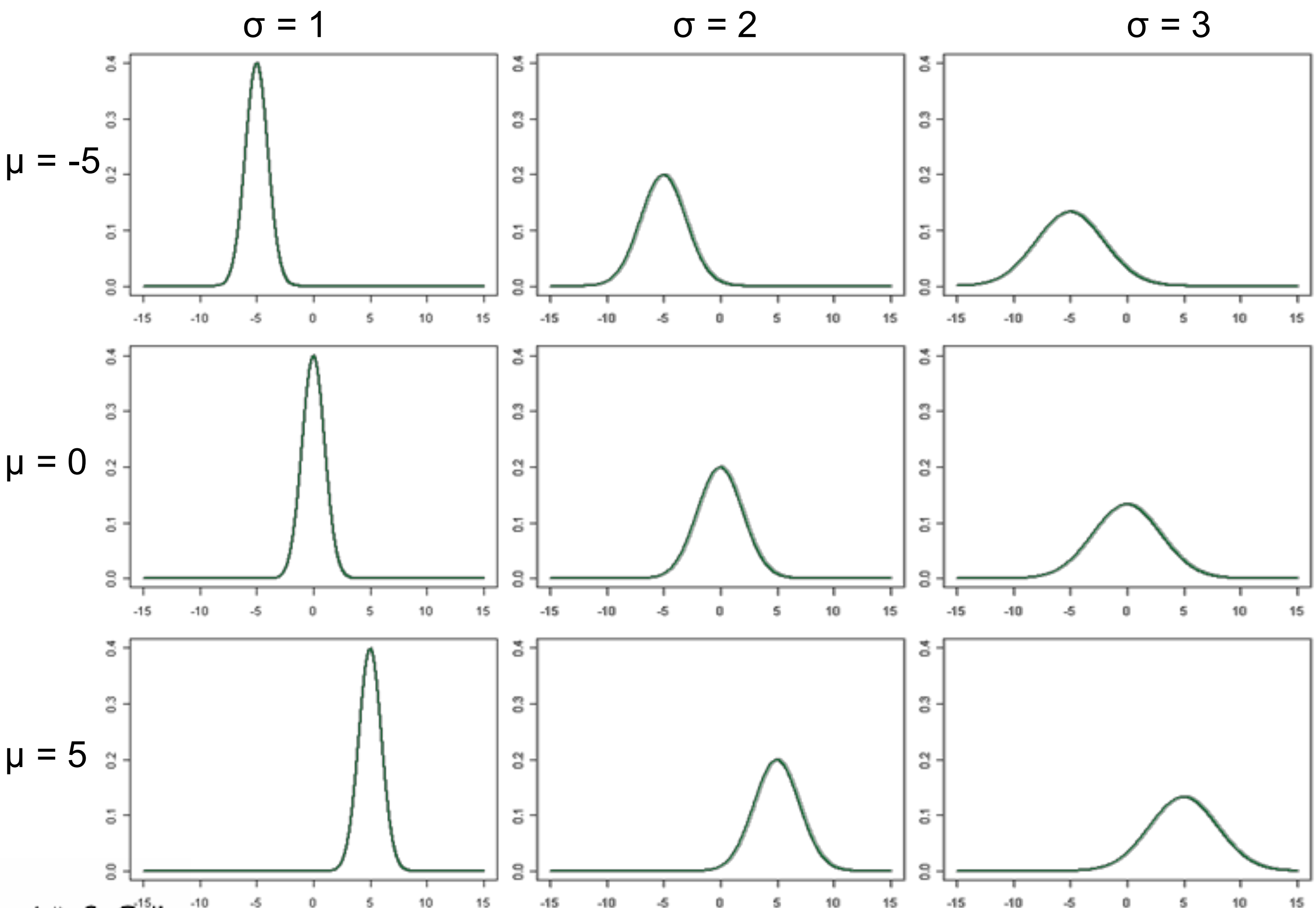
$$\overline{X} := \frac{1}{n} \sum_{i=1}^{n} X_i ,$$

dass

$$\overline{X} \sim \mathcal{N}(\overline{\mu}, \overline{\sigma}^2), \text{ für}$$

$$\overline{\mu} = \frac{1}{n} \sum_{i=1}^{n} \mu_i , \qquad \overline{\sigma}^2 = \frac{1}{n^2} \sum_{i=1}^{n} \sigma_i^2$$

Illustration der Normalverteilung (Dichtefunktion) mit verschiedenen Lage- und Streuungsparametern



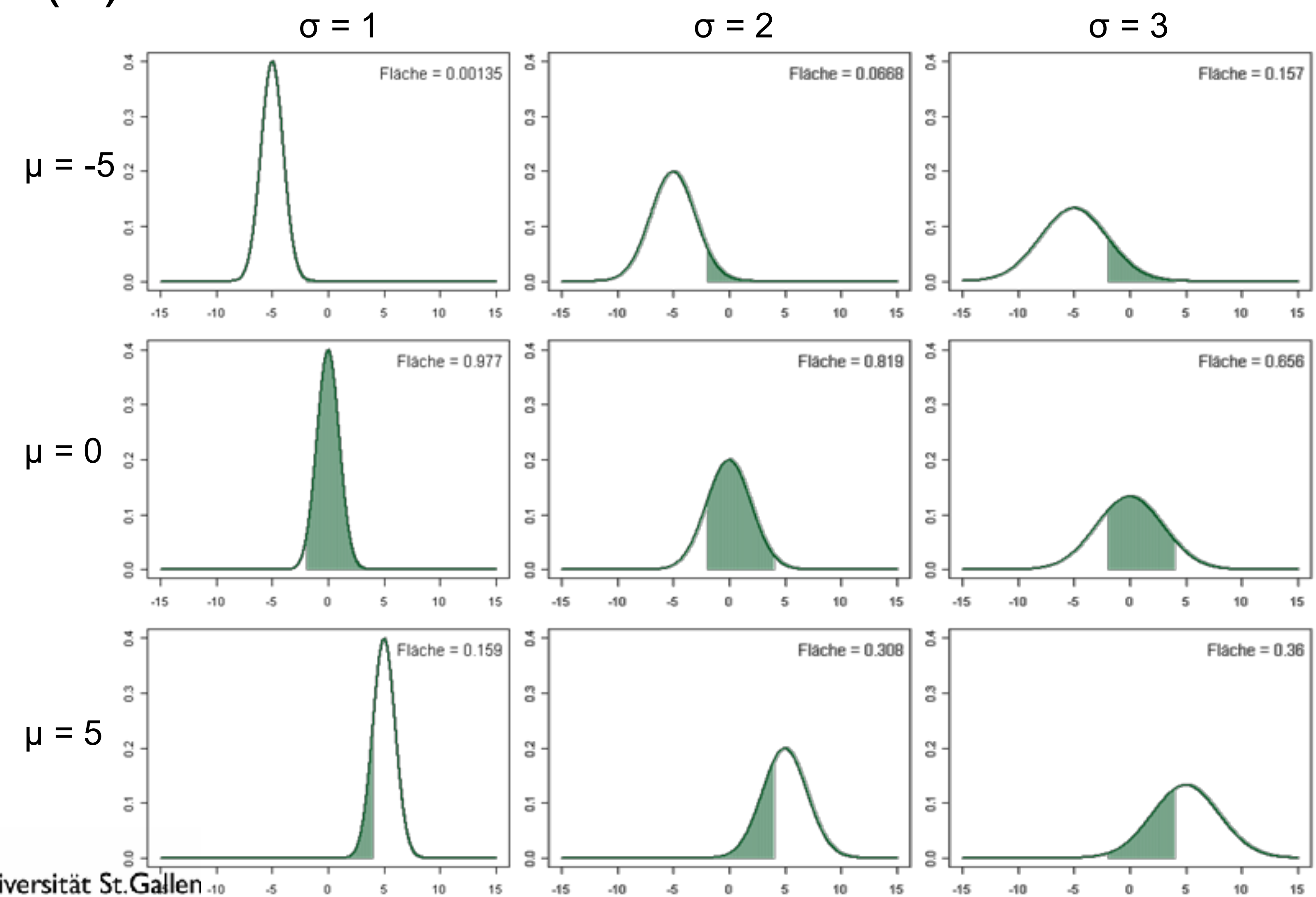
Beispiel 4.7.1:

Die Zufallsvariable X sei normalverteilt mit
E[X] = 5 und V(X) = 9.
Gesucht ist die Wahrscheinlichkeit $P(-2 < X \leq 4)$.

$$
\begin{aligned}
P(-2 < X \leq 4) &= P\left(\frac{-2-\mu}{\sigma} < \frac{X-\mu}{\sigma} \leq \frac{4-\mu}{\sigma}\right) \\
&= P\left(-\frac{7}{3} < Z \leq \frac{-1}{3}\right) \\
&= F_Z\left(-\frac{1}{3}\right) - F_Z\left(-\frac{7}{3}\right) \\
&= 0.3694 - 0.0098 = 0.3596
\end{aligned}
$$

Beispiel 4.7.1 *(Fortsetzung):*

Wahrscheinlichkeit $P(-2 < X \le 4)$ für $E[X] = \mu$ und $V(X) = \sigma^2$.



Beispiel: Asset Allocation

R = Rendite eines Assets im Aktienmarkt

Ann: $R_i \sim \mathcal{N}(\mu_i, \sigma_i)$ für $i = 1, 2, 3$

$\mu = \mathbb{E}[R]$: Erwartete Rendite

$\sigma = \sigma_R := \sqrt{Var(R)}$: Volatilität (Risiko) des Assets.

Assets:	erwartete Rendite	Risiko / Volatilität
A_1	44%	22%
A_2	36%	20%
A_3	10%	4%

Standard normalverteilte Z.V.

$$\mathbb{P}(R_1 < 0) = \mathbb{P}\left(\underbrace{\frac{R_1 - \mu_1}{\sigma_1}}_{Z} \leq \frac{0 - \mu_1}{\sigma_1}\right) = \mathbb{P}\left(Z \leq \frac{0 - \mu_1}{\sigma_1}\right)$$

$$= \mathbb{P}\left(Z - \frac{0 - 44\%}{22\%}\right) = 0.0228$$

$$\mathbb{P}(R_2 < 0) = \ldots = 0.0359$$

$$\mathbb{P}(R_3 < 0) = \ldots = \underline{0.0062}$$

das Verlustrisiko von A_3 ist am kleinsten

IDEE: Gleichmässiges Aufteilen in A_1, A_2, A_3

$$\bar{R} = \frac{1}{3}(R_1 + R_2 + R_3)$$

$$\underline{\mathbb{E}[\bar{R}]} = \frac{1}{3}\left(\underbrace{\mathbb{E}[R_1]}_{44\%} + \underbrace{\mathbb{E}[R_2]}_{36\%} + \underbrace{\mathbb{E}[R_3]}_{10\%}\right)$$

$$= \underline{30\%} = \bar{\mu}$$

$$\mathrm{Var}(\bar{R}) = \left(\frac{1}{3}\right)^2 \left(\mathrm{Var}(R_1) + \mathrm{Var}(R_2) + \mathrm{Var}(R_3)\right)$$

$$= \frac{1}{9}\left(0.22^2 + 0.2^2 + 0.04^2\right)$$

$$= 0.01 = \bar{\sigma}^2$$

$$\mathbb{P}(\bar{R} < 0) = \ldots = 0.0013$$

Wissen $\bar{R} \sim \mathcal{N}(\bar{\mu}, \bar{\sigma}^2)$

Beispiel 4.7.2: *Asset Allocation*

In der Portfoliotheorie wird oft mit der Annahme gearbeitet, dass Renditen von Wertpapieren normal-verteilte Zufallsvariablen sind.

$\rightarrow$ $E[R] = \mu$ erwartete Rendite

$\sigma_R = \sqrt{V(R)}$ Volatilität (Risiko) des Wertpapiers

Ein Anleger möchte einen bestimmten Geldbetrag in Aktien anlegen

	Aktie	Mittelwerte erwartete Rendite	Standardabweichungen Volatilität
risikoreich	A_1	44%	22%
	A_2	36%	20%
risikoarm	A_3	10%	4%

Universität St.Gallen

Beispiel 4.7.2 *(Fortsetzung)*

•) Bei einer Anlage, deren Gesamtbetrag in einer einzelnen Aktie angelegt ist (Vermeidung eines Kapitalverlustes):

$$P\left(R_1 < 0\right) = P\left(\frac{R_1 - \mu_1}{\sigma_1} < \frac{0 - 0.44}{0.22}\right) = P\left(Z < -2\right) = 0.0228$$

$$P\left(R_2 < 0\right) = P\left(\frac{R_2 - \mu_2}{\sigma_2} < \frac{0 - 0.36}{0.2}\right) = P\left(Z < -1.8\right) = 0.0359$$

$$P\left(R_3 < 0\right) = P\left(\frac{R_3 - \mu_3}{\sigma_3} < \frac{0 - 0.1}{0.04}\right) = P\left(Z < -2.5\right) = 0.0062$$

Verlustrisiko am kleinsten (aber auch Rendite-Erwartung)

Universität St.Gallen

Beispiel 4.7.2 *(Fortsetzung)*

•) Bei gleichmässiger Aufteilung des Anlagebetrages und unkorrellierten Renditen:

Portfolio-Rendite: $\bar{R} = \frac{1}{3}\left(R_1 + R_2 + R_3\right) : f_N\left(x; \mu_{\bar{R}}, \sigma_{\bar{R}}^{\,2}\right)$

$$\rightarrow \quad \mu_{\bar{R}} = E\left[\bar{R}\right] = \frac{1}{3}\left(44 + 36 + 10\right) = 30\,[\%]$$

$$\sigma_{\bar{R}}^{\,2} = V\left(\bar{R}\right) = \left(\frac{1}{3}\right)^2 \left(22^2 + 20^2 + 4^2\right) = 100\,\%^2$$

$$P\left(\bar{R} < 0\right) = \ldots = 0.0013 \;<\; P\left(R_3 < 0\right) \quad !$$

→ Die Streuung des Vermögens auf mehrere Anlagearten erweist sich also als überlegene Strategie!

Universität St.Gallen

Kapitel 10

(Multivariate)

5. Mehrdimensionale Zufallsvariablen

Bei vielen Problemstellungen in der Theorie und in der Praxis sind Analysemethoden, die auf Beziehungen, Zusammenhängen oder Abhängigkeiten zwischen verschiedenen Variablen basieren, angemessen, weil sie das Wesentliche erfassen.

Bsp: (Portfoliorendite)

$\mathbb{R} \ni X$: Rendite in SMI

$\mathbb{R} \ni Y$: Rendite in S&P 500

$Z := \begin{pmatrix} X \\ Y \end{pmatrix} \in \mathbb{R}^2$ Multi-variate Z.V.

Beispiel 5.0.1*: Portfoliorendite*

Wir betrachten ein Finanzportfolio basierend auf zwei Indizes (dem S&P 500 und dem FTSE 100). Um die Porftoliorendite zu modellieren könnte man nun die beiden Indexrenditen getrennt analysieren (zwei univariate Zufallsvariablen). Dabei würde jedoch die stochastische Abhängigkeit der beiden Variablen vernachlässigt.

Um diese zu berücksichtigen verwendet man üblicher weise einen multivariaten (in diesem Fall bivariaten) Ansatz.

→ Beziehung zwischen den Variablen ist stochastischer Natur

→ zweidimensionale Zufallsvariable:
→ bivariate Verteilung

→ mehrdimensionale Zufallsvariable:
→ multivariate Verteilung

Universität St.Gallen

Kapitel 10.1

5.1. Gemeinsame Verteilung und Randverteilungen

(X, Y)

Diskrete Zufallsvariablen:

bivariater Fall
$Z = \binom{X}{Y}$

Es seien X und Y die diskreten Komponenten einer 2-dimensionalen Zufallsvariable $Z = (X, Y)$.
Dann heisst die Funktion

$$
\begin{aligned}
f_Z(x,y) &:= \mathbb{P}(\{X = x\} \cap \{Y = y\}) \\
&= \mathbb{P}(X = x, Y = y) \\
&= \mathbb{P}(Z = (x,y))
\end{aligned}
$$

die gemeinsame (Wahrscheinlichkeits)-Massenfunktion von X und Y.

Eigenschaften:

1. $f_Z(x_i, y_j) \geq 0 \quad \forall i \, \forall j$
2. $\sum_{\forall i} \sum_{\forall j} f_Z(x_i, y_j) = 1$

$\Longrightarrow$ 3. $f_Z(x_i, y_j) \leq 1 \quad \forall i \, \forall j$

5.1. Gemeinsame Verteilung und Randverteilungen

Diskrete Zufallsvariablen:

Sind X und Y die diskreten Komponenten einer zweidimensionalen Zufallsvariablen heisst die Funktion

$$f(x,y) = P\,[\{X=x\} \cap \{Y=y\}]$$

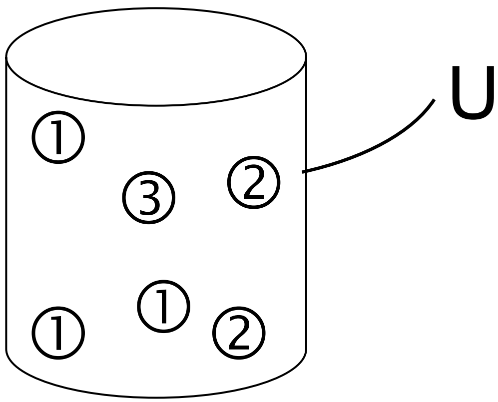
die **gemeinsame Wahrscheinlichkeitsmassen-funktion** von X und Y.

Eigenschaften:

$$\left.\begin{aligned}(1)\qquad & f\left(x_i, y_j\right) \geq 0 \\ (2)\qquad & \sum_{\forall i} \sum_{\forall j} f\left(x_i, y_j\right) = 1\end{aligned}\right\} \Rightarrow \quad (3)\ f\left(x_i, y_j\right) \leq 1, \quad \forall i, j$$

Beispiel 5.1.1:

In einer Urne befinden sich sechs Kugeln. Drei sind mit einer "1" beschriftet, zwei Kugeln mit "2" und eine mit "3".



Nacheinander (ohne Zurücklegen) werden zwei Kugeln gezogen. Die Wahrscheinlichkeitsmassenfunktion von

$Z :=$ (X, Y) = (Aufschrift 1. Kugel, Aufschrift 2. Kugel) ist: [Siehe nächste Seite]

$\mathbb{P}(X=1, Y=1)$

$= \mathbb{P}(X=1) \cdot \mathbb{P}(Y=1 \mid X=1)$

$= 3/6 \cdot 2/5$

Beispiel 5.1.1 (Fortsetzung)

Y / X	y_1=1	y_2=2	y_3=3	f_x
x_1=1	1/5	1/5	1/10	1/2
x_2=2	1/5	1/15	1/15	1/3
x_3=3	1/10	1/15	0	1/6
f_y	1/2	1/3	1/6	1

Zum Beispiel:

$$f(1, 1) = \frac{3}{6} \cdot \frac{2}{5} = \frac{1}{5}$$

$$f(2, 1) = \frac{2}{6} \cdot \frac{3}{5} = \frac{1}{5}$$

$$f(3, 1) = \frac{1}{6} \cdot \frac{3}{5} = \frac{1}{10}$$

Randverteilung

Randverteilung

Beispiel 5.1.2:

Wie oben, aber mit Zurücklegen
(Siehe Beispiel 5.1.1)

X \ Y	**y_1=1**	**y_2=2**	**y_3=3**	**f_x**
x_1=1	1/4	1/6	1/12	1/2
x_2=2	1/6	1/9	1/18	1/3
x_3=3	1/12	1/18	1/36	1/6
f_y	1/2	1/3	1/6	1

Zum Beispiel:

$$f(1,1) = \frac{3}{6} \cdot \frac{3}{6} = \frac{1}{4}$$

~~K~~

~~K~~

~~K~~

Randverteilung

$X = \#$ Köpfe $Y = \#$ der Wechsel

Beispiel 5.1.3 *(Buch Seite 311):*

Eine faire Münze werde viermal hintereinander geworfen.
Sei (X, Y) = (Anzahl der Köpfe, Anzahl der Wechsel)

→ # Möglichkeiten: $2^4=16$

↘

ZZZZ: (0,0); ZZZK: (1,1); ZZKZ: (1,2); ZKZZ: (1,2); KZZZ: (1,1);
ZZKK: (2,1); ZKZK: (2,3); ZKKZ: (2,2); KZKZ: (2,3); KKZZ: (2,1);
KZZK: (2,2); ZKKK: (3,1); KZKK: (3,2); KKZK: (3,2); KKKZ: (3,1);
KKKK: (4,0).

X = # Köpfe

Beispiel 5.1.3 *(Fortsetzung)* Y = # Wechsel

Massenfkt.

Y / X	$y_1 = 0$	$y_2 = 1$	$y_3 = 2$	$y_4 = 3$	f_x
$x_1 = 0$	1/16	0	0	0	1/16
$x_2 = 1$	0	1/8	1/8	0	1/4
$x_3 = 2$	0	1/8	1/8	1/8	3/8
$x_4 = 3$	0	1/8	1/8	0	1/4
$x_5 = 4$	1/16	0	0	0	1/16
f_y	1/8	3/8	3/8	1/8	1

Massenfkt.

KZZZ + ZZZK

$P(X=1, Y=1)$

Merke:

$\sum_j p_{i,j}$

In diesen drei Beispielen:

$$f_x(x_i) = P[X = x_i] = \sum_j f(x_i, y_j) = p_{i,g};$$

$$f_y(y_j) = P[Y = y_j] = \sum_i f(x_i, y_j) = p_{g\,j};$$

heissen **Randverteilung** von X, resp. Y.

$p_{i,j} := f(x_i, y_j)$

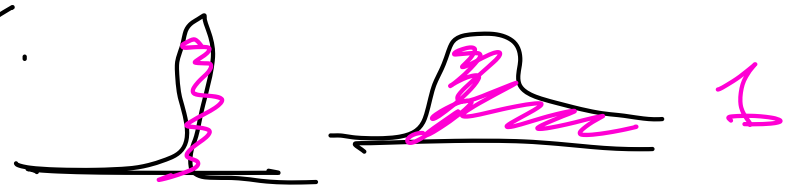
Stetige Zufallsvariablen:

Seien X, Y stetige Zufallsvariablen, dann heisst die Funktion

$f(x,y)$ mit der Eigenschaft $\quad a \leq b,\ c \leq d$

$$\int_a^b \int_c^d f(x,y)\,dy\,dx = \mathbb{P}\left(\{a \leq X \leq b\} \cap \{c \leq Y \leq d\}\right)$$
$$= \mathbb{P}\left(a \leq X \leq b,\ c \leq Y \leq d\right)$$

gemeinsame (Wahrscheinlichkeits-) Dichtefunktion von X und Y.



Eigenschaften:

① $f(x,y) \geq 0 \quad \forall x,y$

② $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\,dy\,dx = 1$

Frage ist es möglich $f(x,y) > 1$ für einige $x, y \in \mathbb{R}$? $\Longrightarrow$ Ja

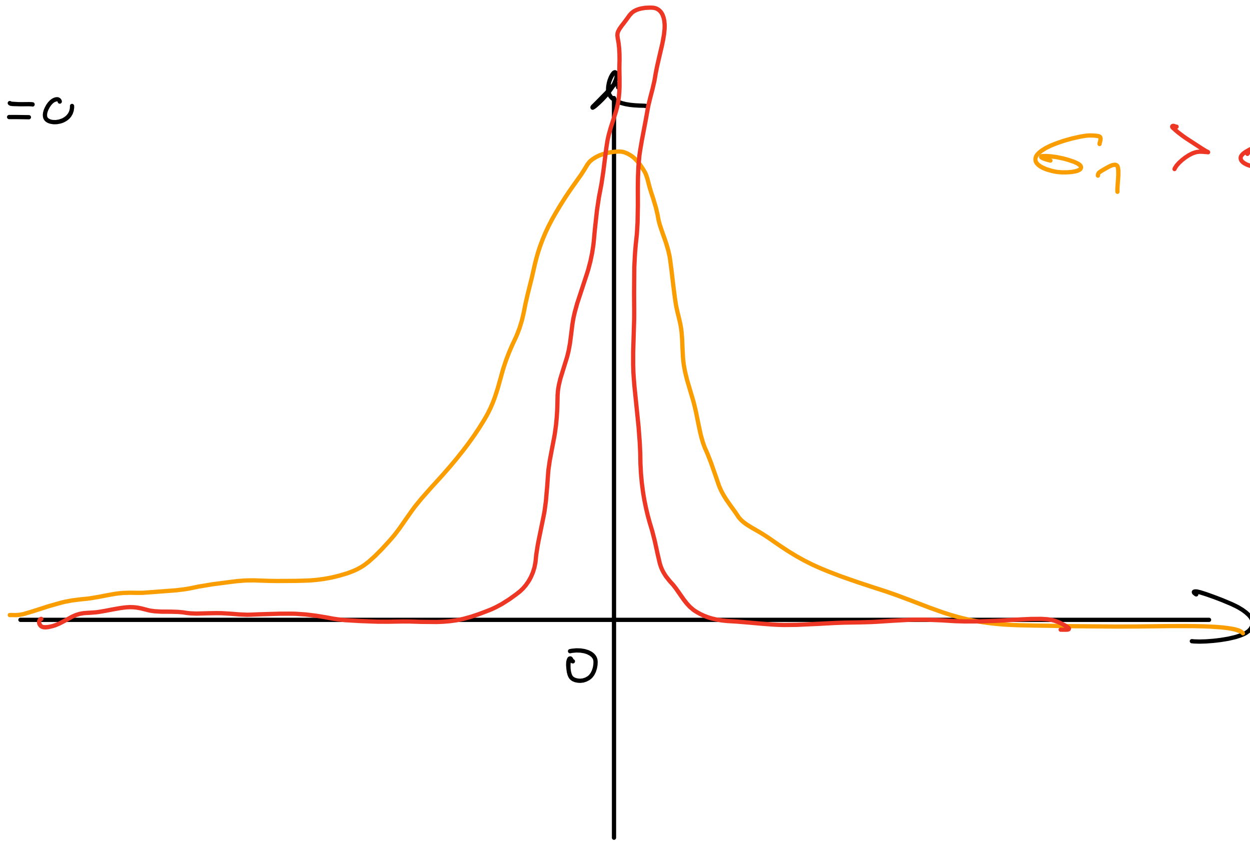
Exkurs:

$$X \sim N(\mu, \sigma^2)$$

Dichtefunktion $\boxed{f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}}$

$\mu = 0$

$\sigma_1 > \sigma_2$



Stetige Zufallsvariablen:

Sind X und Y stetige Zufallsvariablen, dann heisst die Funktion f(x,y) mit

$$\int_a^b \int_c^d f(x,y)\,dy\,dx \quad = \quad P\left[\{a < X \leq b\} \cap \{c < Y \leq d\}\right]$$

für a<b und c<d die **gemeinsame Wahrscheinlichkeitsdichtefunktion** von X und Y.

$\int_{-\infty}^{\infty} F(x)\,dx$

Eigenschaften:

(1) $f(x,y) \geq 0$

(2) $\int_{-\infty}^{\infty} \underbrace{\int_{-\infty}^{\infty} f(x,y)\,dy}_{F(x)}\,dx = 1$

Beispiel 5.1.4:

Die Dichtefunktion der zweidimensionalen Normalverteilung könnte lauten:

$$f(x, y) = \frac{1}{2\pi} \cdot e^{-\frac{x^2+y^2}{2}}, \qquad (x, y) \in \mathbb{R}^2.$$

Beispiel 5.1.5:

Die Dichtefunktion von (X, Y) sei

$$f(x,y) = \begin{cases} \frac{12}{7}\left(x^2 + xy\right), & \text{für } 0 \le x \le 1 \text{ und } 0 \le y \le 1 \\ 0, & \text{sonst} \end{cases}$$

Ges: Finden sie c:

IDEE: $\displaystyle\int_{-\infty}^{\infty} \underbrace{\int_{-\infty}^{\infty} f(x,y)\, dy}_{=:F(x)}\, dx = 1 \longrightarrow c = \ldots = \frac{12}{7}$

$$F(x) = c\int_0^1 (x^2+xy)\, dy = c\left(x^2y + \frac{xy^2}{2}\right)\Big|_0^1 = c\left(x^2 + \frac{x}{2}\right)$$

$$\Rightarrow \int_0^1 F(x)\, dx = c\int_0^1 \left(x^2 + \frac{x}{2}\right) dx = c\left(\frac{x^3}{3} + \frac{x^2}{4}\right)\Big|_0^1$$

$$= c\left(\frac{1}{3} + \frac{1}{4}\right) = \frac{c \cdot 7}{12} \overset{!}{=} 1$$

$$\Rightarrow c = \frac{12}{7}$$

Beispiel 5.1.5:

Die Dichtefunktion von (X, Y) sei

$$f(x,y) = \begin{cases} \dfrac{12}{7}\left(x^2 + xy\right), & \text{für } 0 \le x \le 1 \text{ und } 0 \le y \le 1 \\ 0, & \text{sonst} \end{cases}$$

$\rightarrow$ nicht-negativ ? ✓

$$\frac{12}{7}\int_0^1\int_0^1\left(x^2+xy\right)dydx = \frac{12}{7}\int_0^1 x^2y\Big|_0^1 + \frac{xy^2}{2}\Big|_0^1 \; dx = \frac{12}{7}\int_0^1\left(x^2+\frac{x}{2}\right)dx$$

$$= \frac{12}{7}\left(\frac{x^3}{3}+\frac{x^2}{4}\right)\Big|_0^1 = \frac{12}{7}\left(\frac{1}{3}+\frac{1}{4}\right) = \frac{12}{\cancel{7}}\cdot\frac{\cancel{7}}{12} \cancel{=} 1 \qquad ✓$$

Bei stetigen Zufallsvariablen sind die ***Randverteilungen*** von X und Y durch:

$$f_x(x) = \int_{-\infty}^{\infty} f(x,y)\,dy, \qquad \geq 0$$

$$f_y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx, \qquad \geq 0$$

bestimmt.

$$\int_{\mathbb{R}} f_X(x)\,dx = 1 = \int_{\mathbb{R}} f_Y(y)\,dy$$

Universität St.Gallen

$$f(x,y)=\begin{cases}\frac{12}{7}\left(x^2+xy\right), & \text{für } 0\le x\le 1 \text{ und } 0\le y\le 1\\ 0, & \text{sonst}\end{cases}$$

Gesucht: Randverteilungen

$$\underline{f_X(x)} \overset{\text{Def}}{:=} \int_0^1 f(x,y)\,dy = \frac{12}{7}\int_0^1 (x^2+xy)\,dy$$

$$= \frac{12}{7}\left(x^2 y + x\frac{y^2}{2}\right)\Big|_0^1 \quad \text{für } x\in[0,1]$$

$$= \underline{\frac{12}{7}\left(x^2+\frac{x}{2}\right)}$$

$$f_Y(y) \overset{\text{Def}}{=} \int_0^1 f(x,y)\,dx = \frac{12}{7}\int_0^1 (x^2+xy)\,dx$$

$$= \frac{12}{7}\left(\frac{x^3}{3}+\frac{x^2 y}{2}\right)\Big|_{x=0}^{x=1}$$

$$= \underline{\frac{12}{7}\left(\frac{1}{3}+\frac{y}{2}\right)} \quad \text{für } y\in[0,1]$$

Wofür sind Randverteilungen nützlich?

$(X, Y) \longrightarrow$ gemeinsame Dichte $f(x,y)$

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

Dichte der Randverteilung

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y) \, dy$$

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} y \, f_Y(y) \, dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y) \, dx$$

$$\mathrm{Var}(X) = \mathbb{E}\Big[\big(X - \underbrace{\mathbb{E}[X]}_{\mu_X}\big)^2\Big]$$

$$= \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) \, dx$$

$$\mathrm{Var}(X) := \mathbb{E}\Big[\big(X - \mathbb{E}[X]\big)^2\Big]$$

$$= \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Beispiel 5.1.5 *(Fortsetzung):*

$$f_x(x) = \frac{12}{7}\int_0^1 \left(x^2 + xy\right)dy = \frac{12}{7}\cdot\left[x^2 y + x\frac{y^2}{2}\right]_0^1$$

$$= \frac{12}{7}\cdot\left(x^2 + \frac{x}{2}\right), \qquad x \in [0,1].$$

$$f_y(y) = \frac{12}{7}\int_0^1 \left(x^2 + xy\right)dx = \frac{12}{7}\cdot\left[\frac{x^3}{3} + \frac{x^2 y}{2}\right]_0^1$$

$$= \frac{12}{7}\cdot\left(\frac{1}{3} + \frac{y}{2}\right), \qquad y \in [0,1].$$

Bemerkung:

Für Zufallsvariablen mit einer gemeinsamen Verteilung kann man Erwartungswert und Varianz der einzelnen Komponenten **mittels der Randverteilung** berechnen:

$$\rightarrow \quad \mu_x = E[X] = \sum_i x_i f_x(x_i);$$

$$\sigma_x^2 = V(X) = \sum_i (x_i - \mu_x)^2 f_x(x_i) \quad (\text{diskret})$$

$$\rightarrow \quad \mu_x = E[X] = \int_{-\infty}^{\infty} x\, f_x(x)\, dx;$$

$$\sigma_x^2 = V(X) = \int_{-\infty}^{\infty} (x - \mu_x)^2 f_x(x)\, dx \quad (\text{stetig})$$

Definition

Die gemeinsame Verteilungsfunktion

$$F(x,y) = P[X \leq x, Y \leq y]$$

gibt an, mit welcher Wahrscheinlichkeit X-Werte kleiner oder gleich x und gleichzeitig Y-Werte kleiner oder gleich y sind.

Berechnung:

$$F(x,y) = \sum_{x_i \le x} \sum_{y_j \le y} f(x_i, y_j), \qquad ((X,Y)\, diskret)$$

$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v) \cdot dv \cdot du, \qquad ((X,Y)\, stetig)$$

Universität St.Gallen

Kapitel 10.2

Bedingte Verteilung

(1) Diskreter Fall: Seien X, Y diskrete Z.V. mit gemeinsamer Massenfunktion $f_{X,Y}(x,y)$. Dann ist die bedingte Massenfunktion von X gegeben $Y = y_j$ definiert als

$$f_{X|Y=y_j}(x) := \frac{f_{X,Y}(x, y_j)}{f_Y(y_j)}, \quad (*)$$

$$\mathbb{P}(X = x \mid Y = y_j) \overset{\text{Bedingte}}{\underset{\text{W'keit}}{=}} \frac{\mathbb{P}(X = x, Y = y_j)}{\mathbb{P}(Y = y_j)}$$

(*) nur definiert für y_j mit $f_Y(y_j) > 0$

(2) Stetiger Fall: Seien X, Y stetige Z.V. mit gemeinsamer Dichte $f_{X,Y}(x,y)$. Dann ist die bedingte Dichtefunktion von X gegeben $Y = y$, definiert als

$$f_{X|Y=y}(x) := \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

(für $f_Y(y) > 0$)

$$Z = (X, Y) \Rightarrow \mathbb{E}[Z] = (\mathbb{E}[X], \mathbb{E}[Y])$$

$$\iint (x,y)\, f_{X,Y}(x,y)\, dxy$$

5.2. Bedingte Verteilungen und stochastische Unabhängigkeit

Definition

a) Seien X und Y diskrete Zufallsvariablen mit gemeinsamer Massenfunktion p_{ij}. Die **bedingte Massenfunktion von X, gegeben dass $Y=y_j$**, ist definiert durch

$$f_{X|Y=y_j}(x_i) = \frac{f_{X,Y}(x_i, y_j)}{f_Y(y_j)} = \frac{p_{i,j}}{p_{\cdot,j}}, \quad \textit{für } p_{\cdot,j} > 0 \textit{ und } 0 \textit{ sonst}.$$

b) Seien X und Y stetige Zufallsvariablen mit gemeinsamer Dichtefunktion $f(x,y)$. Die ***bedingte Dichtefunktion von X, gegeben dass $Y = y$***, ist definiert durch

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}, \qquad \textit{für } f_y(y) > 0 \textit{ und } 0 \textit{ sonst}.$$

Beispiel 5.2.1:

Beim dreimaligen Münzwurf mit (X,Y)=(Kopfanzahl beim ersten Wurf, Gesamtanzahl der Köpfe) suchen wir die bedingte Verteilung von X gegeben dass Y=1.

Y / X	$y_1=0$	$y_2=1$	$y_3=2$	$y_4=3$	f_x
$x_1=0$	1/8	1/4	1/8	0	1/2
$x_2=1$	0	1/8	1/4	1/8	1/2
f_y	1/8	3/8	3/8	1/8	1

$$f_{X|Y=1}(x) \longrightarrow \frac{f_{X|Y=1}(0)}{f_{X|Y=1}(1)}$$

Wkeit der Kopfzahl
beim ersten Wurf
gegeben dass total
1 Kopf gesehen wurde

$$f_{X|Y=1}(0) = \frac{f(0,1)}{f_Y(1)} = \frac{1/4}{3/8} = \frac{2}{3}$$

$$f_{X|Y=1}(1) = \frac{f(1,1)}{f_Y(1)} = \frac{1/8}{3/8} = \frac{1}{3}$$

Beispiel 5.2.1 *(Fortsetzung):*

$f_{X|Y=1}(x):\quad X = x_1 = 0: \quad \dfrac{p_{1,2}}{p_{\times,2}} = \dfrac{1/4}{3/8} = \dfrac{2}{3}$

$\qquad\qquad\quad X = x_2 = 1: \quad \dfrac{p_{2,2}}{p_{\times,2}} = \dfrac{1/8}{3/8} = \dfrac{1}{3}$

should sum up to one

Andere Beispiele:

$f_{Y|X=0}(y):\quad Y = 0: \quad \dfrac{1}{4}$

$Y = 1: \quad \dfrac{1}{2}$

$Y = 2: \quad \dfrac{1}{4}$

$Y = 3: \quad 0 \qquad \ldots$

Beispiel 5.2.2:

gemeinsamer

(X, Y) sei ein stetiger Zufallsvektor mit Dichtefunktion

$$f(x,y) = \begin{cases} \lambda^2 \cdot e^{-\lambda x} & , \quad 0 \le y \le x \\ 0 & , \quad \text{sonst.} \end{cases} \quad , \lambda > 0$$

→ Randverteilung? bedingte Dichte?

$$\underline{f_X(x)} \overset{\text{Def.}}{=} \int_0^x f(x,y)\,dy = \lambda^2 \int_0^x e^{-\lambda x}\,dy$$

$$= \lambda^2 e^{-\lambda x}\, y \Big|_0^x = \underline{\lambda^2 x e^{-\lambda x}}$$

$$\underline{f_Y(y)} \overset{\text{Def}}{=} \int_y^\infty f(x,y)\,dx = \lambda^2 \int_y^\infty e^{-\lambda x}\,dx$$

$$= \lambda^2 \left(-\frac{1}{\lambda}\right) e^{-\lambda x} \Big|_y^\infty$$

$$= \underline{\lambda e^{-\lambda y}}$$

$$\Rightarrow \boxed{Y \sim \mathrm{Exp}(\lambda)}$$

falls C stimmt

$$\mathbb{1}_{\{C\}} = \begin{cases} 1, \\ 0, \text{sonst} \end{cases}$$

Seien X und Y stetige Zufallsvariablen mit gemeinsamer Dichtefunktion $f(x,y)$. Die ***bedingte Dichtefunktion von X, gegeben dass Y = y***, ist definiert durch

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}, \qquad \text{für } f_y(y) > 0 \text{ und } 0 \text{ sonst.}$$

$$f(x,y) = \begin{cases} \lambda^2 \cdot e^{-\lambda x} & , \ 0 \le y \le x \\ 0 & , \ \text{sonst.} \end{cases}$$

1) Bedingte Dichte von $X \mid Y=y$

$$\underline{f_{X|Y=y}(x)} \overset{\text{Def}}{=} \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

$f_Y(y) \longrightarrow$ Randverteilung

$$= \frac{\lambda^{\cancel{2}} e^{-\lambda x} \mathbb{1}\{0 \le y \le x\}}{\cancel{\lambda} e^{-\lambda y} \mathbb{1}\{y \ge 0\}}$$

$$= \underline{\lambda e^{-\lambda(x-y)}, \quad \text{für } x \ge y \ge 0}$$

2) Bedingte Dichte von $Y \mid X = x$

$$\underline{f_{Y|X=x}(y)} \overset{\text{Def}}{=} \frac{f_{X,Y}(x,y)}{f_X(x)}$$

$$= \frac{\lambda^2 e^{-\lambda x} \mathbb{1}\{0 \le y \le x\}}{\lambda^2 x e^{-\lambda x} \mathbb{1}\{0 \le x\}}$$

$$= \underline{\frac{1}{x}} \quad \text{für } 0 \le y \le x$$

Beispiel 5.2.2:

(X,Y) sei ein stetiger Zufallsvektor mit Dichtefunktion

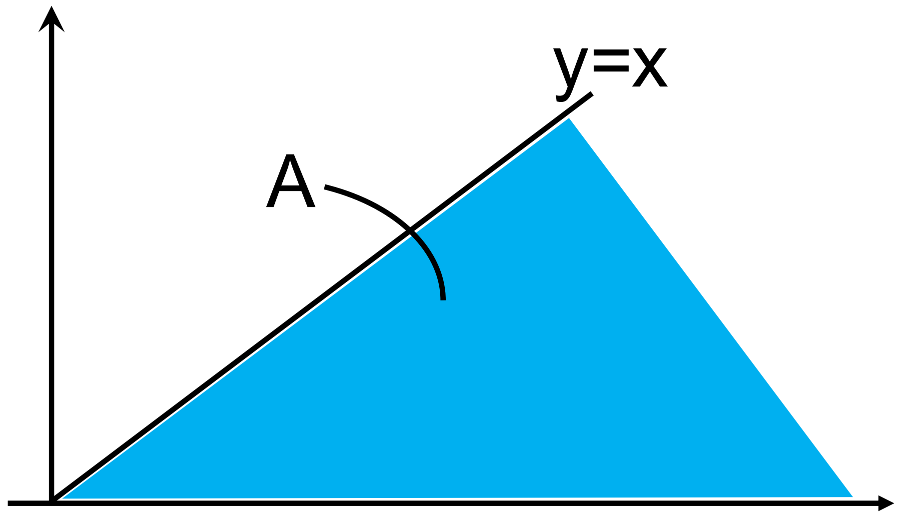
$$f(x,y)=\begin{cases}\lambda^2\cdot e^{-\lambda x} & , \quad 0\le y\le x\\ 0 & , \quad \text{sonst.}\end{cases}$$

$\rightarrow$ Randverteilung:

$$f_x(x)=\int_0^x \lambda^2\cdot e^{-\lambda x}\,\mathrm{d}y=\lambda^2\cdot e^{-\lambda x}\cdot x, \quad x\ge 0$$

$$f_y(y)=\int_y^{+\infty}\lambda^2\cdot e^{-\lambda x}\,\mathrm{d}x=\lambda^2\cdot\left[-\frac{1}{\lambda}e^{-\lambda x}\right]_y^{+\infty}=\lambda\cdot e^{-\lambda y}, \quad y\ge 0$$

$$\left(\rightarrow y:\ \mathrm{Exp}(\lambda)\right).$$



$$A = \{(x, y) \mid x \in [0, \infty), y \in [0, x]\}$$

oder

$$A = \{(x, y) \mid y \in [0, \infty), x \in [y, \infty)\}$$

Beispiel 5.2.2 *(Fortsetzung):*

→ Bedingte Dichte von *X* gegeben *Y*:

$$f_{X|Y=y}(x) = \frac{\lambda^2 \cdot e^{-\lambda x} \cdot \ 1_{\{0 \le y \le x\}}}{\lambda \cdot e^{-\lambda y} \cdot \ 1_{\{y \ge 0\}}} = \lambda \cdot e^{-\lambda(x\text{-}y)}, \qquad x \ge y \ge 0.$$

→ Bedingte Dichte von *Y* gegeben *X*:

$$f_{Y|X=x}(y) = \frac{\cancel{\lambda^2 \cdot e^{-\lambda x}} \cdot \ 1_{\{0 \le y \le x\}}}{\cancel{\lambda^2 \cdot e^{-\lambda x}} \mathrm{x} \cdot \ 1_{\{x \ge 0\}}} = \frac{1}{x}, \qquad 0 \le y \le x.$$

Def: (Stochastische Unabhängigkeit)

Zwei Zufallsvariablen X, Y heissen (stochastisch) unabhängig, wenn

$$f_{X,Y}(x,y) = f_X(x)\, f_Y(y) \quad \forall x, y$$

Definition: Stochastische Unabhängigkeit

Die Zufallsvariablen X und Y heissen **stochastisch unabhängig** (oder kurz: unabhängig) wenn die gemeinsame Massen- bzw. Dichtefunktion gerade gleich dem Produkt der beiden Randverteilungen ist. Das heisst:

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y), \quad \text{für } \forall (x,y)$$

Universität St.Gallen

Teil I: Wahrscheinlichkeitsrechnung

1. Grundlagen der Wahrscheinlichkeit
   - 1.1. Ereignisse, Ereignisraum und Ereignismenge
   - 1.2. Das Rechnen mit Ereignissen
   - 1.3. Der Wahrscheinlichkeitsbegriff
   - 1.4. Axiomatik der Wahrscheinlichkeitstheorie
   - 1.5. Wichtige Regeln der Wahrscheinlichkeitsrechnung
   - 1.6. Wahrscheinlichkeitsräume
   - 1.7. Bedingte Wahrscheinlichkeit und stochastische Unabhängigkeit
   - 1.8. Totale Wahrscheinlichkeit
   - 1.9. Das Bayes-Theorem

2. Elementare Kombinatorik
   - 2.1. Fakultät und Binomialkoeffizient
   - 2.2. Das Fundamentalprinzip der Kombinatorik
   - 2.3. Permutationen
   - 2.4. Kombinationen
   - 2.5. Variationen

3. Zufallsvariablen
   - 3.1. Die Verteilungsfunktion
   - 3.2. Diskrete Zufallsvariablen
   - 3.3. Stetige Zufallsvariablen
   - 3.4. Erwartungswerte von Zufallsvariablen
   - 3.5. Varianz
   - 3.6. Standardisieren

4. Stochastische Modelle und spezielle Verteilungen
   - 4.1. Gleichförmige Verteilung (diskret)
   - 4.2. Bernoulli-Verteilung (diskret)
   - 4.3. Binomialverteilung (diskret)
   - 4.4. Poisson-Verteilung (diskret)
   - 4.5. Rechteckverteilung (stetig)
   - 4.6. Exponentialverteilung (stetig)
   - 4.7. Normalverteilung (stetig)

5. **Mehrdimensionale Zufallsvariablen**
   - 5.1. Gemeinsame Verteilung und Randverteilungen
   - 5.2. Bedingte Verteilungen und stochastische Unabhängigkeit
   - **5.3. Kovarianz und Korrelationskoeffizient**
   - 5.4. Summe von zwei oder mehreren Zufallsvariablen

6. Der zentrale Grenzwertsatz

Teil II: Statistik

7. Beschreibende/Deskriptive Statistik
   - 7.1. Häufigkeitsverteilung, Histogramm und Verteilungsfunktion
   - 7.2. Messzahlen zur Beschreibung statistischer Verteilungen
   - 7.3. Boxplot
   - 7.4. Quantile-Quantile Plot
   - 7.5. Streudiagramm

8. Schätzung unbekannter Parameter: Punktschätzung
   - 8.1. Intuitiv heuristische Ansätze für Schätzfunktionen
   - 8.2. Eigenschaften von Punktschätzungen
   - 8.3. Methoden zur Konstruktion von Schätzfunktionen

9. Intervallschätzungen-Konfidenzintervalle
   - 9.1. Konzept des Konfidenzintervalls
   - 9.2. Ableitung von Konfidenzintervallen (bei grossen Stichproben)
   - 9.3 Zusammenhang mit Hypothesentests

10. Hypothesentests
   - 10.1 Arten von Hypothesen
   - 10.2 Kritischer Bereich und Teststatistik
   - 10.3 Gütefunktion und Arten von Fehlern
   - 10.4 Der p-Wert

Teil III: Aufgaben

$X \in \mathbb{R}, Y \in \mathbb{R}, g: \mathbb{R}^2 \to \mathbb{R}$

$f_{X,Y}(x,y)$

Kapitel10.3

Eindimensionaler Fall: $g : \mathbb{R} \to \mathbb{R}$

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

$$\mathbb{E}[g(X)] = \sum_{\forall i} g(x_i) f(x_i)$$

5.3. Kovarianz und Korrelationskoeffizient

Definition

Seien X und Y zwei Zufallsvariablen mit der gemeinsamen Massen- bzw. Dichtefunktion $f(x,y)$. Der **Erwartungswert der Funktion** $g(X,Y)$ ist definiert als

$$E\left[g(X,Y)\right] = \sum_{\forall i} \sum_{\forall j} g(x_i, y_j)\, f(x_i, y_j), \quad \text{falls } (X,Y) \text{ diskret;}$$

$$E\left[g(X,Y)\right] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y)\, f(x,y)\, dy\, dx, \quad \text{falls } (X,Y) \text{ stetig.}$$

Erinnerung: $\mathrm{Var}(X) = \mathbb{E}[(X-\mu_X)(X-\mu_X)]$

Definition: Seien X, Y zwei Zufallsvariablen mit Erwartungswerten $\mathbb{E}[X] = \mu_X$, $\mathbb{E}[Y] = \mu_Y$. Dann heisst

$$\mathrm{Cov}(X,Y) := \mathbb{E}[(X-\mu_X)\cdot(Y-\mu_Y)]$$

heisst Kovarianz zwischen X, Y

Randverteilung

$$\mu_X = \mathbb{E}[X] = \int_{-\infty}^{\infty} x\, f_X(x)\, dx$$

Multiplikationssatz für Erwartungswerte

$$\mathbb{E}[X\cdot Y] = \mathbb{E}[X]\cdot\mathbb{E}[Y] + \mathrm{cov}(X,Y)$$

Bew: Ann: X, Y stetig

$$\mathrm{Cov}(X,Y) \overset{\mathrm{Def}}{=} \iint \underbrace{(x-\mu_X)(y-\mu_Y)}_{xy - x\mu_Y - y\cdot\mu_X + \mu_X\mu_Y} f_{X,Y}(x,y)\, dx\, dy$$

$$= \iint x y f_{X,Y}(x,y)\,dx\,dy \qquad \mathbb{E}[XY]$$

$$x f_X(x)$$

$$- \mu_Y \iint x f_{X,Y}(x,y)\,dy\,dx$$

$$y f_Y(y)$$

$$- \mu_X \iint y f_{X,Y}(x,y)\,dx\,dy$$

$$+ \mu_X \mu_Y \iint f_{X,Y}(x,y)\,dx\,dy = 1$$

$$= \mathbb{E}[XY] - \mu_Y \int x f_X(x)\,dx \qquad \mathbb{E}[X] = \mu_X$$

$$- \mu_X \int y f_Y(y)\,dy \qquad \mathbb{E}[Y] = \mu_Y$$

$$+ \mu_X \cdot \mu_Y$$

$$= \mathbb{E}[XY] - \mu_X \mu_Y \cancel{- \mu_X \mu_Y} \cancel{+ \mu_X \mu_Y}$$

Eine Masszahl für den stochastischen Zusammenhang zwischen *X* und *Y* ist die **Kovarianz**.

Definition

Seien *X* und *Y* zwei Zufallsvariablen mit den Erwartungswerten μ_x und μ_y.

Die Grösse

$$Cov(X,Y) = \mathrm{E}\left[(X - \mu_X) \cdot (Y - \mu_Y)\right]$$

heisst **Kovarianz** zwischen X und Y.

Universität St.Gallen

Multiplikationssatz für Erwartungswerte:

$$E[XY] = E[X] \cdot E[Y] + \mathrm{Cov}(X,Y)$$

Bew: $\mathrm{Cov}(X,Y) = \iint \underbrace{(x-\mu_X)(y-\mu_Y)}_{xy - x\mu_Y - y\mu_X + \mu_X\mu_Y} f_{X,Y}(x,y)\,dx\,dy$

$$= \underbrace{\iint xy\, f_{X,Y}(x,y)\,dx\,dy}_{\mathbb{E}[XY]}$$

$$- \mu_Y \iint x\, f_{X,Y}(x,y)\,dx\,dy$$

$$- \mu_X \iint y\, f_{X,Y}(x,y)\,dx\,dy$$

$$+ \mu_X\mu_Y \underbrace{\iint f_{X,Y}(x,y)\,dx\,dy}_{1}$$

$$= \mathbb{E}[XY] + \mu_X\mu_Y - \mu_Y \int x \underbrace{\int f_{X,Y}(x,y)\,dy}_{f_X(x)}\,dx$$

$$- \mu_X \int y \underbrace{\int f_{X,Y}(x,y)\,dx}_{f_Y(y)}\,dy$$

$$= \mathbb{E}[XY] + \mu_X\mu_Y - \mu_Y\mu_X - \mu_X\mu_Y = \mathbb{E}[XY] - \mu_X\mu_Y$$

☒

$$\boxed{A \Rightarrow B \iff \neg B \Rightarrow \neg A}$$

Multiplikationssatz für Erwartungswerte:

$$E[XY] = E[X] \cdot E[Y] + \text{Cov}(X,Y)$$

$\rightarrow$ Sind X und Y unabhängig, ist

$$E[XY] = E[X] \cdot E[Y]$$

Um festzustellen, ob zwei Zufallsvariablen unabhängig oder abhängig sind, könnte man also die Kovarianz ausrechnen und nachschauen, ob sie Null ist oder nicht.

A B

$X,\ Y$ sind unabhängig $\rightarrow$ $Cov(X,Y) = 0$

oder $\underbrace{Cov(X,Y) \neq 0}_{\neg B}$ $\rightarrow$ X,Y sind nicht unabhängig

Rechenregeln mit Kovarianzen

Seien X, Y, U, V Zufallsvariablen

Seien $a, b, c, d \in \mathbb{R}$ Parameter (reelle Zahlen)

Dann gilt:

$$\boxed{\begin{aligned} &\mathrm{Cov}(a \cdot U + b \cdot V, cX + dY) \\ &= a \cdot c \, \mathrm{Cov}(U, X) + a \cdot d \, \mathrm{Cov}(U, Y) \\ &\quad + b \cdot c \, \mathrm{Cov}(V, X) + bd \, \mathrm{Cov}(Y, Y) \end{aligned}}$$

Bsp: $\mathrm{Cov}(2 \cdot X, 5 \cdot Y)$

$$= 10 \, \mathrm{Cov}(X, Y$$

Erinnerung: Skalarprodukt!

$$\langle a \cdot u + bv, cx + dy \rangle = ac \langle u, x \rangle + ad \langle u, y \rangle$$
$$+ bc \langle v, x \rangle + bd \langle v, y \rangle$$

Lineare Algebra

$v \in \mathbb{R}^n, \quad u \in \mathbb{R}^n$

$\langle v, u \rangle = v^T u \quad = 0$

$\Longrightarrow v, u$ sind orthogonal

$$Cov(X,Y) = \mathrm{E}\left[(X-\mu_X)\cdot(Y-\mu_Y)\right]$$

$$\underbrace{Cov(\overset{\tilde{X}}{aX}, \overset{\tilde{Y}}{bY})}_{} = ab\, Cov(X,Y)$$

$$\mathbb{E}\Big[\underbrace{(aX - a\mu_X)\cdot(bY - b\mu_Y)}_{}\Big]$$

$$abXY - abX\mu_Y - abY\mu_X$$

$$+ ab\mu_X\mu_Y$$

! $Cov(X,Y) \in \mathbb{R}$

Rechenregeln für die Kovarianz

Die Kovarianz ist ***bilinear***, das heisst, sie ist linear in beiden Argumenten. Seien zu diesem Zweck U, V, X, Y geeignete Zufallsvariablen und a, b, c, d reelle Zahlen. Dann gilt

$$\mathrm{Cov}(a\cdot U+b\cdot V,\, c\cdot X+d\cdot Y) =$$
$$a\cdot c\cdot \mathrm{Cov}(U,X) + a\cdot d\cdot \mathrm{Cov}(U,Y) + b\cdot c\cdot \mathrm{Cov}(V,X)+ b\cdot d\cdot \mathrm{Cov}(V,Y)$$

Bemerkung: Die Kovarianz erfüllt also dieselben Rechenregeln wie ein Skalarprodukt. Deshalb meint man mit X und Y sind ***orthogonal***, dass ihre Kovarianz verschwindet, d.h., Cov(X,Y) = 0.

Zufallsvariablen mit positiver (negativer) Kovarianz heissen **positiv (negativ)** korreliert. Verschwindet die Kovarianz, so sind die Zufallsvariablen **unkorreliert**.

Definition

Der Quotient aus der Kovarianz und $\sigma_x\sigma_y$ von X und Y

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y}$$

heisst **Korrelationskoeffizient** zwischen X und Y.

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

$\sigma_Y = \sqrt{Var(Y)}$

Teil I: Wahrscheinlichkeitsrechnung

1. Grundlagen der Wahrscheinlichkeit
   - 1.1. Ereignisse, Ereignisraum und Ereignismenge
   - 1.2. Das Rechnen mit Ereignissen
   - 1.3. Der Wahrscheinlichkeitsbegriff
   - 1.4. Axiomatik der Wahrscheinlichkeitstheorie
   - 1.5. Wichtige Regeln der Wahrscheinlichkeitsrechnung
   - 1.6. Wahrscheinlichkeitsräume
   - 1.7. Bedingte Wahrscheinlichkeit und stochastische Unabhängigkeit
   - 1.8. Totale Wahrscheinlichkeit
   - 1.9. Das Bayes-Theorem

2. Elementare Kombinatorik
   - 2.1. Fakultät und Binomialkoeffizient
   - 2.2. Das Fundamentalprinzip der Kombinatorik
   - 2.3. Permutationen
   - 2.4. Kombinationen
   - 2.5. Variationen

3. Zufallsvariablen
   - 3.1. Die Verteilungsfunktion
   - 3.2. Diskrete Zufallsvariablen
   - 3.3. Stetige Zufallsvariablen
   - 3.4. Erwartungswerte von Zufallsvariablen
   - 3.5. Varianz
   - 3.6. Standardisieren

4. Stochastische Modelle und spezielle Verteilungen
   - 4.1. Gleichförmige Verteilung (diskret)
   - 4.2. Bernoulli-Verteilung (diskret)
   - 4.3. Binomialverteilung (diskret)
   - 4.4. Poisson-Verteilung (diskret)
   - 4.5. Rechteckverteilung (stetig)
   - 4.6. Exponentialverteilung (stetig)
   - 4.7. Normalverteilung (stetig)

5. Mehrdimensionale Zufallsvariablen
   - 5.1. Gemeinsame Verteilung und Randverteilungen
   - 5.2. Bedingte Verteilungen und stochastische Unabhängigkeit
   - 5.3. Kovarianz und Korrelationskoeffizient
   - 5.4. Summe von zwei oder mehreren Zufallsvariablen

6. Der zentrale Grenzwertsatz

Teil II: Statistik

7. Beschreibende/Deskriptive Statistik
   - 7.1. Häufigkeitsverteilung, Histogramm und Verteilungsfunktion
   - 7.2. Messzahlen zur Beschreibung statistischer Verteilungen
   - 7.3. Boxplot
   - 7.4. Quantile-Quantile Plot
   - 7.5. Streudiagramm

8. Schätzung unbekannter Parameter: Punktschätzung
   - 8.1. Intuitiv heuristische Ansätze für Schätzfunktionen
   - 8.2. Eigenschaften von Punktschätzungen
   - 8.3. Methoden zur Konstruktion von Schätzfunktionen

9. Intervallschätzungen-Konfidenzintervalle
   - 9.1. Konzept des Konfidenzintervalls
   - 9.2. Ableitung von Konfidenzintervallen (bei grossen Stichproben)
   - 9.3 Zusammenhang mit Hypothesentests

10. Hypothesentests
   - 10.1 Arten von Hypothesen
   - 10.2 Kritischer Bereich und Teststatistik
   - 10.3 Gütefunktion und Arten von Fehlern
   - 10.4 Der p-Wert

Teil III: Aufgaben

Universität St.Gallen

Kapitel 10.4

5.4. Summe von zwei oder mehreren Zufallsvariablen

Erwartungswert?

$$E\left[X+Y\right] = \sum_{\forall i}\sum_{\forall j}\left[(x_i+y_j)\cdot f(x_i,y_j)\right]$$

$$= \sum_{\forall i}\sum_{\forall j}\left[x_i \cdot f(x_i,y_j)\right] + \sum_{\forall i}\sum_{\forall j}\left[y_j \cdot f(x_i,y_j)\right]$$

$$= E\left[X\right] + E\left[Y\right]$$

Varianz?

$$\underline{V(X+Y)} = E\left[\left((X+Y)-(\mu_x+\mu_y)\right)^2\right] = E\left[\left((X-\mu_x)+(Y-\mu_y)\right)^2\right]$$

$$= E\left[(X-\mu_x)^2 + (Y-\mu_y)^2 + 2\cdot(X-\mu_x)(Y-\mu_y)\right]$$

$$= E\left[(X-\mu_x)^2\right] + E\left[(Y-\mu_y)^2\right] + 2\cdot E\left[(X-\mu_x)(Y-\mu_y)\right]$$

$$\underline{= V(X) + V(Y) + 2\cdot Cov(X,Y)}$$

Arithmetisches Mittel von IID Zufallsvariablen:

$$E\left(\overline{X}_n\right) = E\left[\frac{1}{n}\left(X_1 + \ldots + X_n\right)\right] = \frac{1}{n}\cdot\sum_{i=1}^{n} E\left[X_i\right] = \frac{1}{n}\cdot n\cdot\mu = \mu$$

$$V\left(\overline{X}_n\right) = V\left(\frac{1}{n}\left(X_1 + \ldots + X_n\right)\right) = \frac{1}{n^2}\cdot\sum_{i=1}^{n} V\left[X_i\right] = \frac{1}{n^2}\cdot n\cdot\sigma^2 = \frac{\sigma^2}{n}$$

$$\Rightarrow \quad \sigma_{\overline{X}_n} = \frac{\sigma}{\sqrt{n}}$$

(um den Faktor $\sqrt{n}$ kleiner, $\sqrt{n}$ - gesetz!)

Beispiel 5.4.1:

Betrachten Sie folgendes Spiel: Zur Teilnahme muss eine Gebühr von 1 Euro bezahlt werden. Es wird dreimal eine ideale Münze geworfen, für jeden Kopf wird “1 Euro“ ausgezahlt.

a) Beschreiben Sie die Zufallsvariable X: “Gewinn pro Spiel" über Bernoulli-verteilte Zufallsvariablen. Zeigen Sie, dass $E[X] = \frac{1}{2}$ und $V(X) = \frac{3}{4}$ gilt.

Beispiel 5.4.1 *(Fortsetzung)*

a) Beschreiben Sie die Zufallsvariable *X*: “Gewinn pro Spiel" über Bernoulli-verteilte Zufallsvariablen.
Zeigen Sie, dass E[*X*] = ½ und V(*X*) = ¾ gilt.

$$X = \sum_{i=1}^{3} Y_i - 1, \quad Y_i = \begin{cases} y_{i,1}=1, & \text{"Kopf"}, \quad p = \dfrac{1}{2} \qquad Y_i: \quad \text{i.i.d.} \\ y_{i,2}=0, & \text{"Zahl"}. \end{cases}$$

$$\rightarrow \quad E[X] = E\left[\sum_{i=1}^{3} Y_i - 1\right] = \sum_{i=1}^{3} E[Y_i] - 1 = \sum_{i=1}^{3} p - 1 = 3 \cdot \frac{1}{2} - 1 = \frac{1}{2};$$

$$\rightarrow \quad V(X) = V\left(\sum_{i=1}^{3} Y_i - 1\right) = \sum_{i=1}^{3} V(Y_i) = \sum_{i=1}^{3} p \cdot (1\text{-}p) = \frac{1}{4} \cdot 3 = \frac{3}{4}.$$

Beispiel 5.4.1 *(Fortsetzung)*

b) Anton spielt dieses Spiel drei Runden hinter-einander. U sei sein Gewinn nach drei Spielen. Drücken Sie die Zufallsvariable U unter Zuhilfenahme von X_i, i = 1, 2, 3 aus. Berechnen Sie E[U] und V(U).

$$U = X_1 + X_2 + X_3, \qquad \mathrm{X}_i \, : \text{ i.i.d.}$$

$$\rightarrow \quad \mathrm{E}[U] = 3 \cdot \mathrm{E}[X] = \frac{3}{2};$$

$$\rightarrow \quad \mathrm{V}(U) = 3 \cdot \mathrm{V}(X) = 3 \cdot \frac{3}{4} = \frac{9}{4}.$$

Teil I: Wahrscheinlichkeitsrechnung

1. Grundlagen der Wahrscheinlichkeit
   1.1. Ereignisse, Ereignisraum und Ereignismenge
   1.2. Das Rechnen mit Ereignissen
   1.3. Der Wahrscheinlichkeitsbegriff
   1.4. Axiomatik der Wahrscheinlichkeitstheorie
   1.5. Wichtige Regeln der Wahrscheinlichkeitsrechnung
   1.6. Wahrscheinlichkeitsräume
   1.7. Bedingte Wahrscheinlichkeit und stochastische Unabhängigkeit
   1.8. Totale Wahrscheinlichkeit
   1.9. Das Bayes-Theorem

2. Elementare Kombinatorik
   2.1. Fakultät und Binomialkoeffizient
   2.2. Das Fundamentalprinzip der Kombinatorik
   2.3. Permutationen
   2.4. Kombinationen
   2.5. Variationen

3. Zufallsvariablen
   3.1. Die Verteilungsfunktion
   3.2. Diskrete Zufallsvariablen
   3.3. Stetige Zufallsvariablen
   3.4. Erwartungswerte von Zufallsvariablen
   3.5. Varianz
   3.6. Standardisieren

4. Stochastische Modelle und spezielle Verteilungen
   4.1. Gleichförmige Verteilung (diskret)
   4.2. Bernoulli-Verteilung (diskret)
   4.3. Binomialverteilung (diskret)
   4.4. Poisson-Verteilung (diskret)
   4.5. Rechteckverteilung (stetig)
   4.6. Exponentialverteilung (stetig)
   4.7. Normalverteilung (stetig)

5. Mehrdimensionale Zufallsvariablen
   5.1. Gemeinsame Verteilung und Randverteilungen
   5.2. Bedingte Verteilungen und stochastische Unabhängigkeit
   5.3. Kovarianz und Korrelationskoeffizient
   5.4. Summe von zwei oder mehreren Zufallsvariablen

6. Der zentrale Grenzwertsatz

Teil II: Statistik

7. Beschreibende/Deskriptive Statistik
   7.1. Häufigkeitsverteilung, Histogramm und Verteilungsfunktion
   7.2. Messzahlen zur Beschreibung statistischer Verteilungen
   7.3. Boxplot
   7.4. Quantile-Quantile Plot
   7.5. Streudiagramm

8. Schätzung unbekannter Parameter: Punktschätzung
   8.1. Intuitiv heuristische Ansätze für Schätzfunktionen
   8.2. Eigenschaften von Punktschätzungen
   8.3. Methoden zur Konstruktion von Schätzfunktionen

9. Intervallschätzungen-Konfidenzintervalle
   9.1. Konzept des Konfidenzintervalls
   9.2. Ableitung von Konfidenzintervallen (bei grossen Stichproben)
   9.3 Zusammenhang mit Hypothesentests

10. Hypothesentests
   10.1 Arten von Hypothesen
   10.2 Kritischer Bereich und Teststatistik
   10.3 Gütefunktion und Arten von Fehlern
   10.4 Der p-Wert

Teil III: Aufgaben

Universität St.Gallen

Kapitel 12, insb. 12.4

6. Der zentrale Grenzwertsatz

Denken wir uns eine n-dimensionale Zufalls-variable und nehmen an, dass die einzelnen Komponenten $X_1, \ldots, X_n$, **unabhängig** und **identisch** (i.i.d) verteilt sind mit

$$E\left[X_i\right] = \mu \quad \text{und} \quad V\left(X_i\right) = \sigma^2.$$

Wie entsteht nun eine solche Folge von Zufallsvariablen in der Praxis und welche Bedeutung hat sie? → Zwei verschiedene Fälle:

1) *Stichprobe mit Zurücklegen*

2) *Versuchsreihe*

Beispiel 6.1:

Der Gewinn bei einem Glücksspiel wird durch die Zufallsvariable X beschrieben. Bei wiederholtem Spielen erhalten wir also Zufallsvariablen $X_1,\ldots, X_n$.

Der **Gesamtgewinn nach n Runden** ist $S_n = \sum_{i=1}^{n} X_i$.

Der **durchschnittliche Gewinn pro Runde** ist $\bar{X}_n = \frac{1}{n} S_n$.

Frage: Wie gross ist nach vielen Runden die Wahrscheinlichkeit für einen Gesamtgewinn zwischen a und b, d.h. $P[a \leq S_n \leq b] = ?$

→ Der **zentrale Grenzwertsatz** gibt dafür eine Approximation für grosse n ($n \to \infty$).

Theorem: Zentraler Grenzwertsatz (ZGS)

Seien $X_1, X_2, \ldots, X_n$ i.i.d. Zufallsvariablen mit $\mu = E[X_i]$ und $\sigma^2 = V(X_i)$. Sei S_n die Summe und $\overline{X_n} = \frac{S_n}{n}$ das arithmetische Mittel dieser Zufallsvariablen. Dann strebt die Verteilungsfunktion F_n der standardisierten Grösse

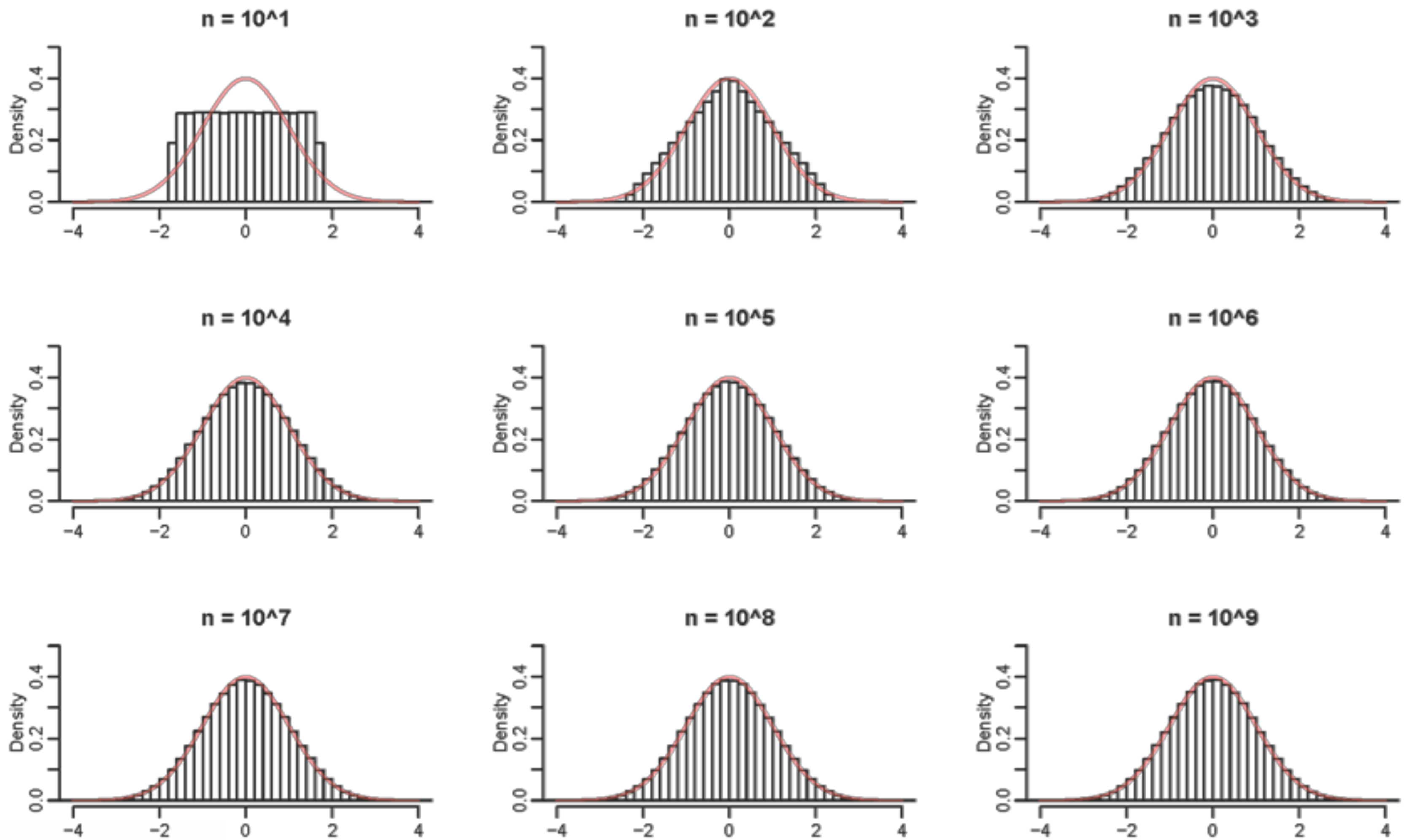
$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\overline{X}_n - \mu}{\sigma / \sqrt{n}}$$

mit wachsendem n $(n \to \infty)$ gegen die Standard-normalverteilung: $F_n(Z_n) \to F_Z(Z)$

Spezialfall: Alle X_i i.i.d. Bernoulli-verteilt.

Illustration des zentralen Grenzwertsatzes

Seien $X_1, X_2, \ldots, X_n$ sind i.i.-gleichverteilt auf [0,1].



Theorem: Grenzwertsatz von De Moivre und Laplace

Sei S_n eine binomialverteilte Zufallsvariable mit den Parametern *n* und *p*. Dann strebt ihre Verteilungsfunktion $F_{Bi}\left(s_n; n, p\right) \rightarrow F_N\left(s_n; np, np\left(1-p\right)\right)$ mit wachsendem *n* gegen die Normalverteilung mit den entsprechenden Momenten. Die Verteilungsfunktion *F* der standardisierten Zufallsvariablen

$$Z_n = \frac{S_n - np}{\sqrt{np\left(1-p\right)}} \equiv \frac{\overline{X}_n - p}{\sqrt{\frac{p\left(1-p\right)}{n}}}$$

konvergiert für $n \rightarrow \infty$ gegen die Standardnormal-verteilung $F_n\left(Z_n\right) \rightarrow F_Z\left(Z\right)$.

Universität St.Gallen

Beispiel 6.2 *(Buch Seite 412):*

Eine Theorie behauptet, die Entwicklung von Aktien- und Wechselkursen auf informations-effizienten Märkten folge einem sogenannten **Random Walk**

$K_t = K_{t-1} + \varepsilon_t$, wobei die Kursänderungen von heute auf morgen $\varepsilon_t = K_t - K_{t-1}$, $E[\varepsilon_t] = 0$

$$V(\varepsilon_t) = \sigma^2$$

den Erwartungswert Null und die gleiche Varianz hätten.

Universität St.Gallen

Beispiel 6.2 *(Fortsetzung):*

Die monatliche Kursänderung wäre dann eine Summe $\varepsilon_t + \varepsilon_{t+1} + \ldots + \varepsilon_{t+n}$, (Anzahl 25 der Han-delstage im Monat).

$\xrightarrow{\text{ZGS}}$ Monatliche Kursänderung asymptotisch normalverteilt mit Erwartungswert Null, aber mit der n-fachen Varianz.

→ Naive Prognose: Der Kurs bleibt, wie er ist.

Beispiel 6.3:

Seien $X_1,\dots, X_{12}$ i.i.d. gleichverteilt auf [-½, ½], und $S_{12} = \sum_{i=1}^{12} X_i$.

Wegen $\mathrm{E}[X_i] = 0 = \mu$, $\mathrm{V}(X_i) = \frac{(b-a)^2}{12} = \frac{1}{12} = \sigma^2$, sagt uns der ZGS, dass

$$S_{12} \overset{\text{approx}}{:} N(n\mu, n\sigma^2) = N(0, 1).$$

Bemerkung: Mit n=12 finden wir eine erstaunlich gute Übereinstimmung der beiden Verteilungen.

Beispiel 6.4:

In einem Experiment wird eine Münze 100 Mal geworfen, und man erhält 60 Mal Kopf. Ist die Münze fair?

$$\text{Sei } X_i = \begin{cases} 1, & \text{Kopf bei Wurf } i \\ 0, & \text{sonst} \end{cases}, \quad i=1,\ldots,100.$$

X_i ist Bernoulli-verteilt mit p=½.
(Annahme: faire Münze)

Beispiel 6.4 *(Fortsetzung):*

S_{100} ist binomialverteilt mit p=½, n=100.

$$P\left[S_{100} \geq 60\right] \overset{\text{Standardisierung}}{\underset{\downarrow}{=}} P\left[Z_{100} = \frac{S_{100}-50}{\sqrt{25}} \geq \frac{60-50}{\sqrt{25}}\right]$$

$$; \quad 1 - F_Z(2) = 0.0228.$$

Also: Für eine faire Münze ist die Wahrscheinlichkeit für das Ereignis $\{S_{100} \geq 60\}$ sehr klein; dass wir nun dieses Ereignis beobachtet haben stellt die Fairness der betrachteten Münze in Frage.

Notiz: Für die Approximation einer diskreten Verteilung mit dem ZGS ist eine Stetigkeitskorrektur nötig.

Teil II: Statistik

Kapitel 1

Repräsentative Stichprobe:

Wünschenswert wäre es, eine Stichprobe auszuwählen, die **repräsentativ** für die Grundgesamtheit ist, also eine Struktur bezüglich der interessierenden Merkmale aufweist, die der Grundgesamtheit möglichst ähnlich ist.

Achtung: Es gibt ***keine strikt statistische*** Definition von ***repräsentativ***.

Eine sehr wichtige Einteilung der Typen von statistischen Variablen ist mit welcher **Skala** sie gemessen werden können. Das Niveau der Messbarkeit bestimmt dabei die Möglichkeiten und Grenzen der statistischen Auswertungen.

1. Nominal messbare Variablen:

Eine Variable ist **nominal skaliert**, wenn lediglich die Gleichheit oder Andersartigkeit verschiedener Ausprägungen festgestellt werden kann.

2. Ordinal messbare Variablen:

Eine Variable ist **ordinal skaliert**, wenn die möglichen Ausprägungen unterscheidbar und zusätzlich in eine natürliche oder sinnvoll festzulegende Rangordnung gebracht werden können.

3. Kardinal messbare Variablen:

Eine Variable ist **kardinal skaliert**, wenn die verschiedenen Ausprägungen nicht nur eine Rangfolge ausdrücken, sondern ausserdem der quantitative Unterschied zwischen ihnen bestimmt ist. Die Ausprägungen müssen Zahlen sein.

7. Beschreibende / deskriptive Statistik

Ziel:

Übersicht über einige Methoden zur **graphischen** Aufbereitung von Daten; das ist oft nützlich, um einen ersten Eindruck und erste Ideen zu bekommen.

Kapitel 1.5

7.1. Häufigkeitsverteilung, Histogramm und Verteilungsfunktion

Beispiel 7.1.1: *Radioaktiver Zerfall von Americium-241*

Das radioaktive Element Am-241 emittiert beim Zerfall α – Teilchen. Wir interessieren uns für die Anzahl der Emissionen innerhalb eines Zeitintervalls bestimmter Länge; z.B. von 10 Sekunden.

Man beobachtet den Zerfallsprozess eine Weile und möchte mit Hilfe der erhobenen Daten ein passendes Modell finden.

Zur Messung der Daten wird die Beobachtungs-periode in 1207 Intervalle von 10 Sekunden unterteilt. In jedem Intervall zählt man die Anzahl der Emissionen und erhält so Daten

$$X_1, \ldots, X_n; \ n = 1207,$$

Die Gesamtzahl der Emissionen ist 10'129.

Beispiel 7.1.1 (Fortsetzung): *Radioaktiver Zerfall von Americium-241)*

In einem ersten Schritt werden nun Klassen gebildet:

für $y_j = 3, 4, \ldots, 16,$ zählt man die Anzahl t_j der Intervalle mit genau y_j Emissionen, also die Anzahl der x_j mit $x_j = y_j$.

Zusätzlich bildet man eine Klasse für 0 - 2 Emissionen und eine Klasse für 17 und mehr Emissionen.

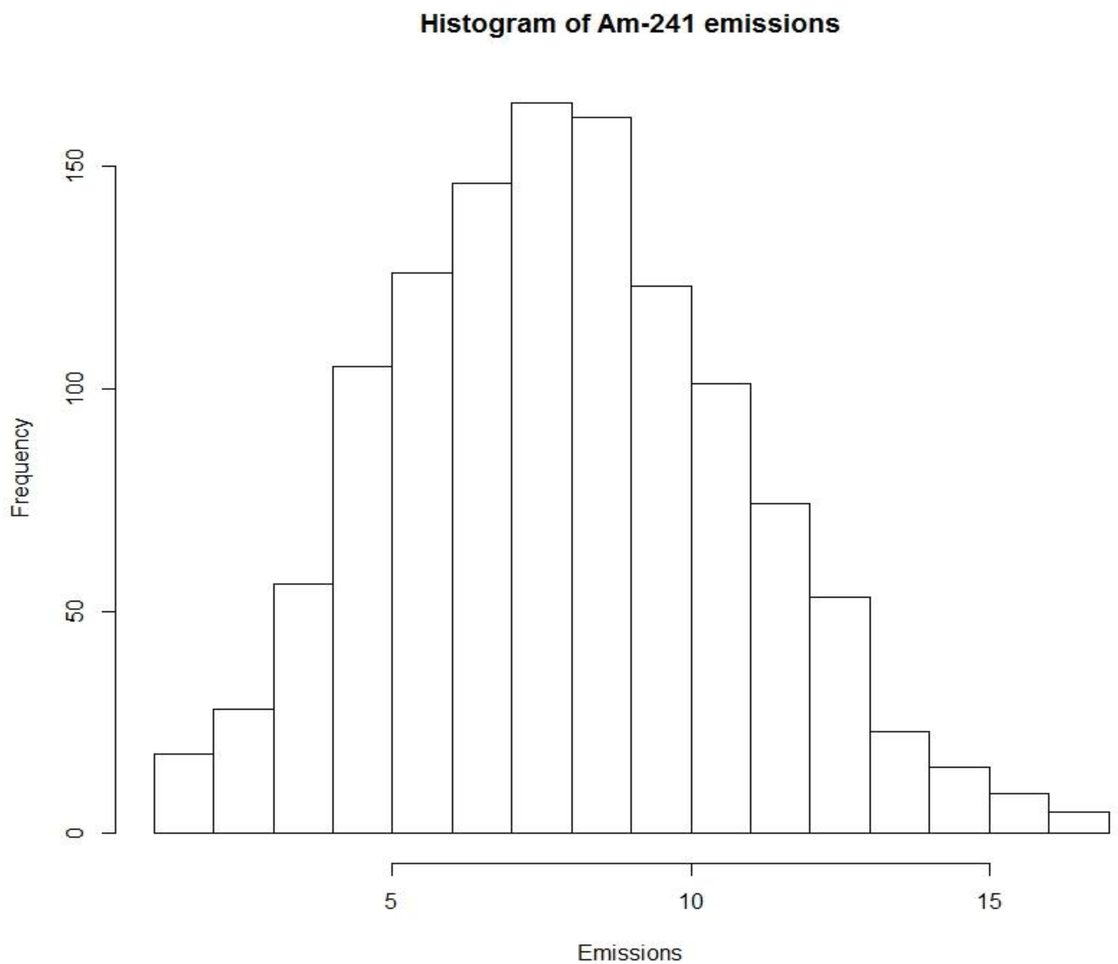
Daraus ergibt sich die folgende Tabelle:

Class interval (emissions)	0-2	3	4	5	6	7	8	9
Numbers	18	28	56	105	126	146	164	161

Class interval (emissions)	10	11	12	13	14	15	16	>17
Numbers	123	101	74	53	23	15	9	5

Beispiel 7.1.1 *(Fortsetzung)**:* *Radioaktiver Zerfall von Americium-241*

Aus der Tabelle erhält man das folgende Histogramm:



***Beispiel 7.1.2**: Bevölkerungspyramide und Histogramme: (Siehe Buch S. 34)*

Bemerkung:

Histogramme können auch für nominal messbare Variablen gebildet werden.

Ein weiteres Hilfsmittel ist die **empirische Verteilungsfunktion** F_n .Sie ist definiert durch

$$F_n(y) = \frac{(Anzahl\ der\ Daten\ \mathrm{x}_i\ mit\ \mathrm{x}_i \leq y)}{n} = \frac{1}{n} \sum_{j,\ \mathrm{y}_j \leq y} f_j\ .$$

Universität St.Gallen

F_n liefert eine Schätzung für die Verteilungsfunktion F, die den Mechanismus hinter den Daten modelliert.

Weil offenbar F_n ausserhalb der Werte $y_1, \ldots, y_m$ unserer Daten konstant ist, plottet man für die graphische Darstellung die Punkte

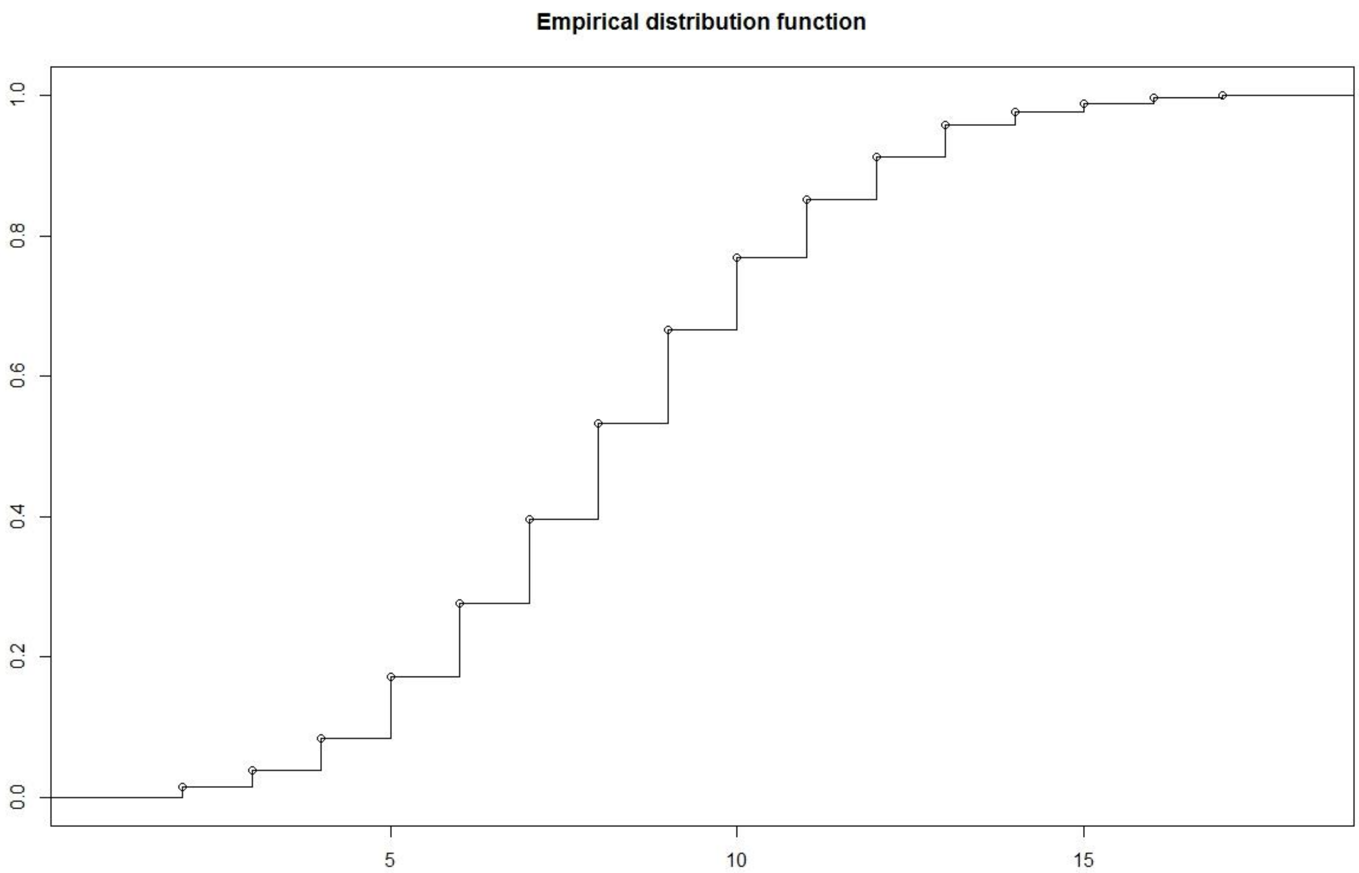
$$(y_j,\ F_n(y_j)) \quad \textit{für} \quad j = 1, \ldots, m.$$

Manchmal verbindet man diese Punkte auch noch mit einer Treppenfunktion.

Beispiel 7.1.1 *(Fortsetzung)**:** *Radioaktiver Zerfall von Americium-241*

→ Tabelle und Bild

y_j	0	1	2	3	4	5	6	7	……
$F_n(y_j)$	?	?	0.0149	0.0381	0.0845	0.1715	0.2759	0.3969	…….



Teil I: Wahrscheinlichkeitsrechnung

1. Grundlagen der Wahrscheinlichkeit
   1.1. Ereignisse, Ereignisraum und Ereignismenge
   1.2. Das Rechnen mit Ereignissen
   1.3. Der Wahrscheinlichkeitsbegriff
   1.4. Axiomatik der Wahrscheinlichkeitstheorie
   1.5. Wichtige Regeln der Wahrscheinlichkeitsrechnung
   1.6. Wahrscheinlichkeitsräume
   1.7. Bedingte Wahrscheinlichkeit und stochastische Unabhängigkeit
   1.8. Totale Wahrscheinlichkeit
   1.9. Das Bayes-Theorem

2. Elementare Kombinatorik
   2.1. Fakultät und Binomialkoeffizient
   2.2. Das Fundamentalprinzip der Kombinatorik
   2.3. Permutationen
   2.4. Kombinationen
   2.5. Variationen

3. Zufallsvariablen
   3.1. Die Verteilungsfunktion
   3.2. Diskrete Zufallsvariablen
   3.3. Stetige Zufallsvariablen
   3.4. Erwartungswerte von Zufallsvariablen
   3.5. Varianz
   3.6. Standardisieren

4. Stochastische Modelle und spezielle Verteilungen
   4.1. Gleichförmige Verteilung (diskret)
   4.2. Bernoulli-Verteilung (diskret)
   4.3. Binomialverteilung (diskret)
   4.4. Poisson-Verteilung (diskret)
   4.5. Rechteckverteilung (stetig)
   4.6. Exponentialverteilung (stetig)
   4.7. Normalverteilung (stetig)

5. Mehrdimensionale Zufallsvariablen
   5.1. Gemeinsame Verteilung und Randverteilungen
   5.2. Bedingte Verteilungen und stochastische Unabhängigkeit
   5.3. Kovarianz und Korrelationskoeffizient
   5.4. Summe von zwei oder mehreren Zufallsvariablen

6. Der zentrale Grenzwertsatz

Teil II: Statistik

7. Beschreibende/Deskriptive Statistik
   7.1. Häufigkeitsverteilung, Histogramm und Verteilungsfunktion
   7.2. Messzahlen zur Beschreibung statistischer Verteilungen
   7.3. Boxplot
   7.4. Quantile-Quantile Plot
   7.5. Streudiagramm

8. Schätzung unbekannter Parameter: Punktschätzung
   8.1. Intuitiv heuristische Ansätze für Schätzfunktionen
   8.2. Eigenschaften von Punktschätzungen
   8.3. Methoden zur Konstruktion von Schätzfunktionen

9. Intervallschätzungen-Konfidenzintervalle
   9.1. Konzept des Konfidenzintervalls
   9.2. Ableitung von Konfidenzintervallen (bei grossen Stichproben)
   9.3 Zusammenhang mit Hypothesentests

10. Hypothesentests
   10.1 Arten von Hypothesen
   10.2 Kritischer Bereich und Teststatistik
   10.3 Gütefunktion und Arten von Fehlern
   10.4 Der p-Wert

Teil III: Aufgaben

7.2. Messzahlen zur Beschreibung statistischer Verteilungen

Definition

a$_1$) ***Arithmetisches Mittel*** (oder Mittelwert) als Lagemass: ($\rightarrow$ kardinal skaliert)

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Definition

a$_2$) ***Median*** als Lagemass: (→ ordinal skaliert)

$$x_{Med} = \begin{cases} x_{\left(\frac{n+1}{2}\right)}, & \text{falls n ungerade} \\ \frac{1}{2}\left(x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}\right), & \text{falls n gerade} \end{cases}$$

a$_3$) ***Modus*** als Lagemass: (→ nominal skaliert)

$$x_M = x_i, \text{ mit } h_i \geq h_j, \text{ für alle } j \neq i.$$

Beispiel 7.2.1: (Siehe Buch S. 46, Bsp. 2)

Die statistische Reihe

4, 7, 7, 7, 12, $\boxed{12,}$ $\boxed{13,}$ 16, 19, 23, 23, 97 hat

Mittelwert $\bar{x} = 20;$

Modus $x_M = 7;$

Median $x_{Med} = \frac{12+13}{2} = 12.5.$

Definition

b) Aus den obgenannten Ordnungsstatistiken $x_{(1)}$, ..., $x_{(n)}$ (den geordneten Daten) erhält man das ***empirische*** α***-Quantil*** für $\alpha \in (0,1)$ wie folgt:

Man bestimmt zuerst $K = \lfloor \alpha n \rfloor + 1$, wobei $\lfloor \cdot \rfloor$ den ganzzahligen Teil bezeichnet und erhält dann als ***empirisches*** α ***-Quantil***

$$\begin{cases} x_{(K)}, & \text{für } \alpha \cdot n \text{ nicht ganzzahlig} \\ \frac{1}{2}\left(x_{(K)} + x_{(K-1)}\right), & \text{für } \alpha \cdot n \text{ ganzzahlig.} \end{cases}$$

Damit liegt etwa der Anteil α der Daten unterhalb des empirischen α -Quantils.

Beispiel 7.2.2:

Sei $\alpha = 75\%$. Für n = 100 ist $\alpha \cdot n = 75$ ganzzahlig, also K = 76.

Damit gerade 75% der Daten links von z liegen müssen wir den Wert z zwischen $x_{(75)}$ und $x_{(76)}$ wählen. Das empirische 75%-Quantil (Q_3)

$\frac{1}{2}\left(x_{(75)} + x_{(76)}\right)$ erfüllt das.

$\rightarrow$ Für n = 101: $\alpha \cdot$ n = 75.75 $\Rightarrow$ K=76,

$\alpha \cdot$ n nicht ganzzahlig $\Rightarrow$ $Q_3 = x_{(76)}$.

Beispiel 7.2.3: (Siehe Buch Seite 61, Bsp. 14)

Die Lageparameter wie Modus, Median und arithmetisches Mittel geben jeweils nur eine **zentrale Tendenz** einer Verteilung an. Nun soll aber auch noch das Ausmass der **Streuung** oder **Dispersion** der Werte einer statistischen Reihe in einer Masszahl ausgedrückt werden.

Das Ausmass der **Streuung** kann helfen, zwei Verteilungen mit derselben zentralen Tendenz zu unterscheiden.

Definition

c_1) **Spannweite** als Streuungsmass: Sie ist die Differenz zwischen der grössten und der kleinsten Merkmalsausprägung: Spannweite = $x_{max} - x_{min}$.

c_2) **Mittlerer Quartilsabstand** als Streuungsmass:

$$MQA = \frac{(Q_3 - Q_2) + (Q_2 - Q_1)}{2} = \frac{IQA}{2},$$

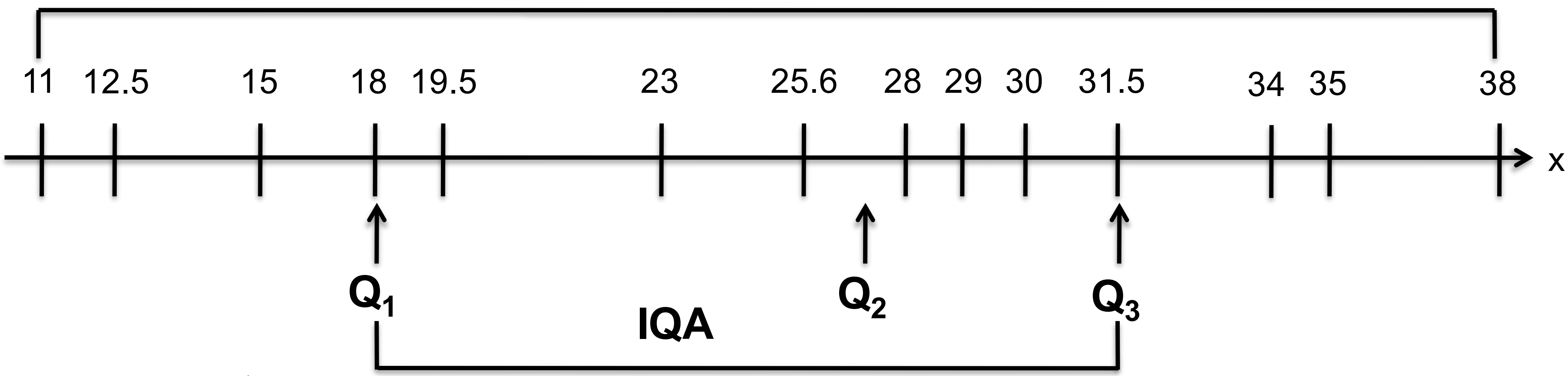
wobei man $IQA = Q_3 - Q_1$ **Interquartilsabstand** nennt.

Universität St.Gallen

Beispiel 7.2.4: (Siehe Buch Seite 53, Bsp. 9)

n = 14 Daten

Spannweite = 38-11=27



$$Q_2 = \frac{1}{2}\left(x_{(7)} + x_{(8)}\right) = 26.8$$

$$Q_1 = x_{(4)} = 18$$

$$\Rightarrow \ \text{IQA} = 13.5 \ \Rightarrow \ \text{MQA} = 6.75.$$

$$Q_3 = x_{(11)} = 31.5$$

Universität St.Gallen

Definition

c_3) ***Empirische Varianz und Standardabweichung als Streuungsmass***:

Die mittlere quadratische Abweichung vom arithmetischen Mittel

$$S_x^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

heisst **empirische Varianz.**

Die positive Wurzel aus der empirischen Varianz

$$S_x = +\sqrt{S_x^2}$$

heisst empirische **Standardabweichung**.

Beispiel 7.2.5: (Siehe Buch Seite 54, Bsp. 11)

Bei der statistischen Reihe

3, 5, 9, 9, 6, 6, 3, 7, 7, 6, 7, 6, 5, 7, 6, 9, 6, 5, 3, 5

rechnen wir mit folgender Arbeitstabelle:

j	x_j	n_j	h_j	$h_j x_j$	$x_j - \bar{x}$	$\left(x_j - \bar{x}\right)^2$	$h_j\left(x_j - \bar{x}\right)^2$
1	3	3	0.15	0.45	-3	9	1.35
2	5	4	0.20	1	-1	1	0.20
3	6	6	0.30	1.8	0	0	0
4	7	4	0.20	1.4	1	1	0.20
5	9	3	0.15	1.35	3	9	1.35
		$n=20$	$\sum=1$	$\bar{x}=6$	$\sum=0$ ✓		$S_x^2=3.1$

Definition

Die bisherigen Masse der Streuung sind alle Masse der *absoluten Streuung* und ihre Werte hängen von der Einheit ab, in der die Variable gemessen wird.
Um das Mass der Streuung zweier Variablen in verschiedenen Einheiten zu vergleichen, müssen wir ein Mass der *relativen Streuung* definieren, wie den **Variationskoeffizienten** (sofern $\bar{x} \neq 0$):

$$VK_x = \frac{S_x}{|\bar{x}|}$$

Universität St.Gallen

Beispiel 7.2.6: Aktienkurse (250 Handelstage):
(Siehe Buch Seite 58, Bsp. 13)

Daimler Chrysler-Aktie: $\overline{x} = 50.59 \text{ Euro} \quad , \quad S_x = 36.18 \text{ Euro}$

Porsche AG-Aktie: $\overline{y} = 396.10 \text{ Euro} \; , \; S_y = 182.96 \text{ Euro}$

$$\Rightarrow VK_x = \frac{36.18}{50.59} = 0.72$$

$$VK_y = \frac{182.96}{396.10} = 0.46$$

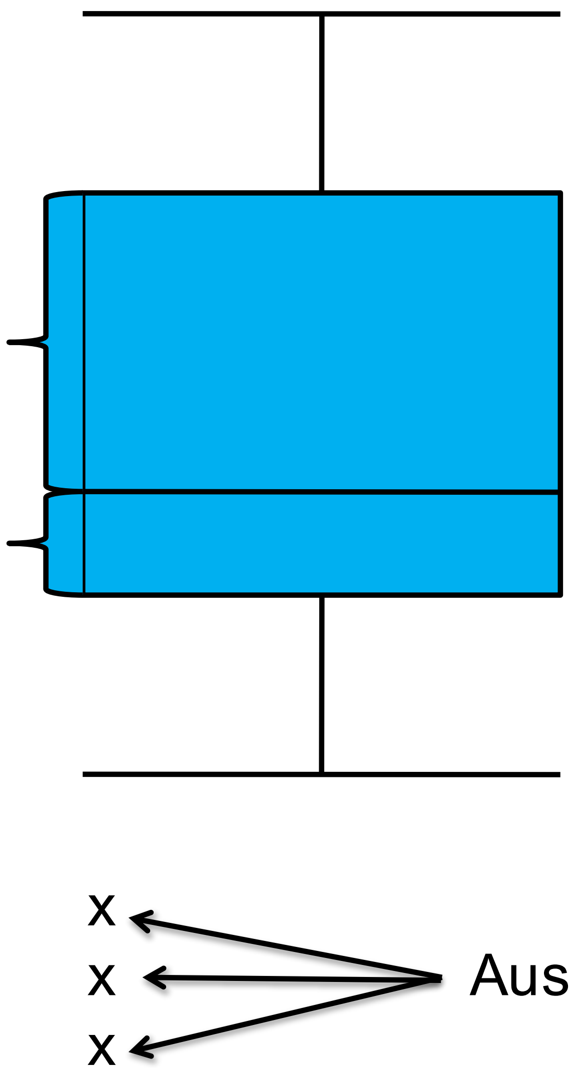
Somit streute die Daimler Chrysler-Aktie trotz geringerer Standardabweichung relativ stärker.

→ VK oft benützt als Mass für die Volatilität eines Aktienkurses!

Kapitel 2

Universität St.Gallen

7.3. Boxplot *Bildlich:* alles skaliert!



25% der Daten liegen hier

25% der Daten liegen hier

Ausreisser

d = grösster Datenwert x_i

mit $x_i - Q_3 < 1.5 \cdot IQA$

Q_3 = 75%-Quantil

Q_2 = Median

Q_1 = 25%-Quantil

$IQA = Q_3 - Q_1$

c = kleinster Datenwert x_i

mit $Q_1 - x_i < 1.5 \cdot IQA$

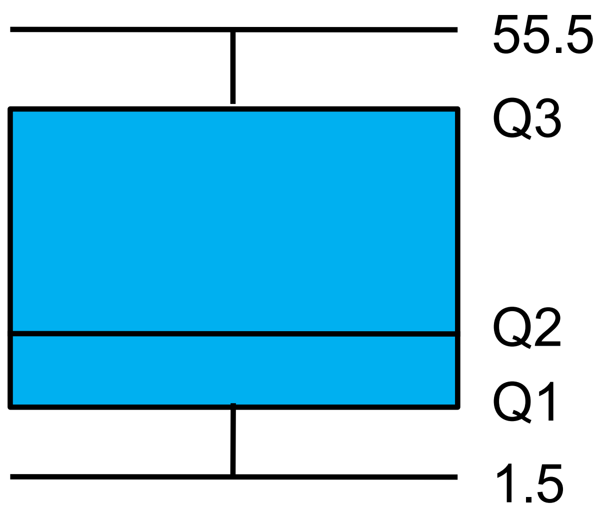
Beispiel 7.3.1: *Lebensdauer von 16 Geräten (Angaben in Monaten)*

1.5; 3.5; 6.5; 11.50; 12.50; 14; 17; 17; 19; 20; 23.5; 32.5; 34.5; 39; 55.5; 119

$$Q_2 = \frac{\left(X_{(8)} + X_{(9)}\right)}{2} = 18\ ; \qquad Q_1 = \frac{\left(X_{(4)} + X_{(5)}\right)}{2} = 12\ ; \qquad Q_3 = \frac{\left(X_{(13)} + X_{(12)}\right)}{2} = 33.5$$

$$\Rightarrow\ \mathrm{IQA} = 21.5 \qquad Q_3 + 1.5 \cdot \mathrm{IQA} = 65.75; \qquad Q_1 - 1.5 \cdot \mathrm{IQA} < 0$$

$$1.5 \cdot \mathrm{IQA} = 32.25$$



- 119 Ausreisser
- relativ kleine Streuung
- nicht symmetrisch!
  (mehr konzentriert zwischen Q1/Q2)

Universität St.Gallen

Kapitel 2

7.4. Quantile-Quantile-Plot (QQ-Plot)

Manchmal hat man eine Idee, von welcher Verteilung die vorliegenden Daten stammen könnten. Ein **QQ-Plot** ist eine graphische Methode mit der man beurteilen kann, wie gut die Daten zur vermuteten Verteilung passen.

Für praktische Beispiele wird man den Computer benutzen.

Doch was für eine theoretische Überlegung steckt dahinter?

Annahme:

Die Daten $x_1, \ldots, x_n$ stammen von Zufallsvariablen $X_1, \ldots, X_n$, die alle die Verteilungsfunktion F haben.

Aus der Analogie zwischen empirischen und theoretischen Quantilen erwarten wir dann

$$x_{(\lfloor \alpha n \rfloor + 1)} \approx F^{-1}(\alpha)$$

denn zirka der Anteil α der Daten liegt links von $x_{(\lfloor \alpha n \rfloor + 1)}$, und mit Wahrscheinlichkeit α liegen die Werte X_i links von

$$F^{-1}(\alpha)\Big(P\left[X_i \le F^{-1}(\alpha)\right] = F\left(F^{-1}(\alpha)\right) \cong \alpha\Big).$$

Ist $\mathrm{K}=\lfloor \alpha \mathrm{n} \rfloor + 1$, so ist $\dfrac{\mathrm{K}-1}{\mathrm{n}} \approx \alpha$, und wir erwarten also

$$\mathrm{x}_{(\mathrm{K})} \approx \mathrm{F}^{-1}(\alpha) \approx \mathrm{F}^{-1}\left(\frac{\mathrm{K}-1}{\mathrm{n}}\right) \approx \mathrm{F}^{-1}\left(\frac{\mathrm{K}-\frac{1}{2}}{\mathrm{n}}\right).$$

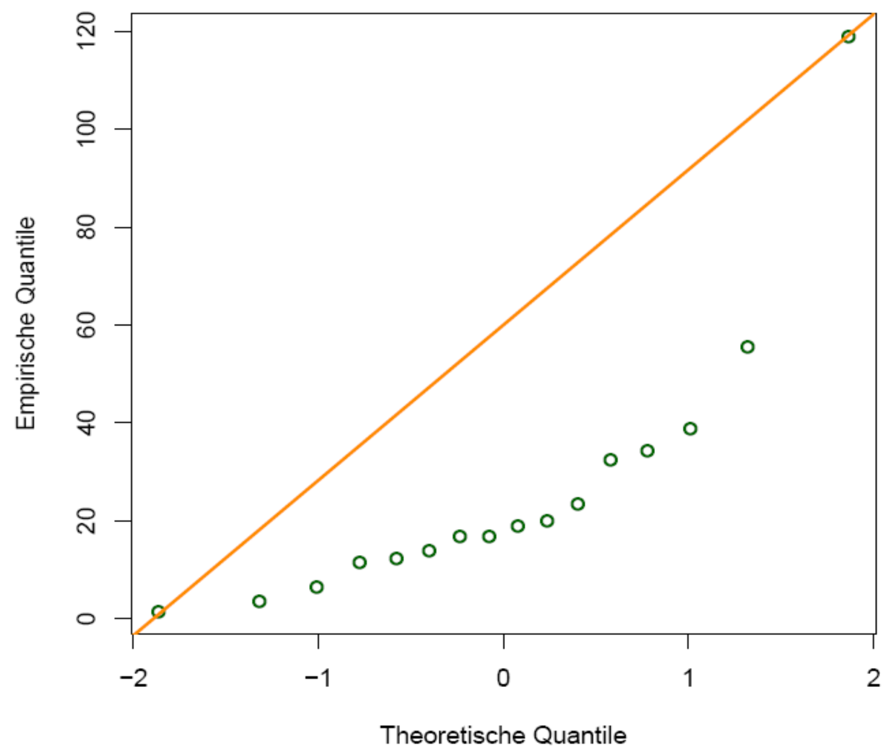
Also sollten die Punkte $\left(\mathrm{F}^{-1}\left(\dfrac{\mathrm{K}-\frac{1}{2}}{\mathrm{n}}\right), \mathrm{x}_{(\mathrm{K})}\right)$, $\mathrm{K}=1,\ldots,\mathrm{n}$,

etwa auf der Winkelhalbierenden im ersten Quadranten liegen.

Wir zeichnen also die empirischen Quantile $x_{(K)}$ gegen die theoretische Quantile $F^{-1}\left(\frac{K-\frac{1}{2}}{n}\right)$ und untersuchen, wie stark diese Punktwolkle von einer Geraden abweicht.

Bespiel 7.3.1 *(Fortsetzung)*:

QQ - Plot der Daten von 7.3.1 (theoretische Verteilung: Normalverteilung)



Universität St.Gallen

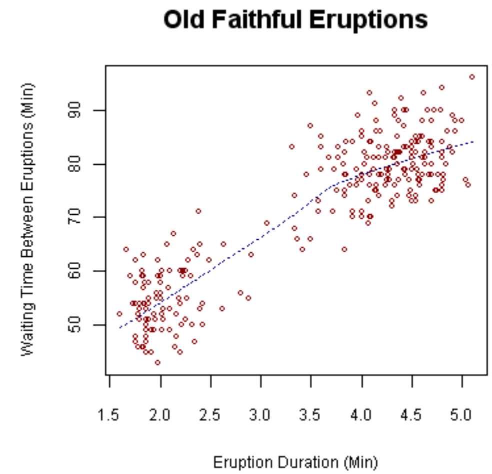
(Kapitel 3.1)

7.5. Streudiagramm

Für Intervall- und Verhältnisdaten sind Abstände zwischen den Merkmalsausprägungen definiert.
Eindimensional: Die einzelnen Beobachtungen lassen sich als Punkte auf der x-Achse darstellen.
Zweidimensional: Die Daten lassen sich als Sammlung von Punkten darstellen, jeder mit dem Wert einer Variable als x-Koordinate und dem Wert der anderen als y-Koordinate.

Der übliche Zweck dieser Abbildungsmethode ist es, eine Idee für eine zweckmässige Klassenbildung in einem Histogramm zu bekommen oder die Beziehung zwischen zwei Variablen zu veranschaulichen.

Beispiel 7.5.1: *Streudiagramm*



Zweidimensional:
Wartezeit zwischen Ausbrüchen und Dauer der Ausbrüche für den Old Faithful Geysir im Yellowstone National Park.

Streudiagramme sind einfach zu konstruieren und zu interpretieren (sofern die Stichprobe nicht zu gross ist): Sie bringen den Messbereich, Extremwerte sowie allfällig vorhandene Konzentrationen augenfällig zum Ausdruck.

Kapitel 13

8. Schätzung unbekannter Parameter: Punktschätzung

Universität St.Gallen

Ein zentraler Problemkreis der Statistik ist die Schätzung unbekannter Parameter von Grundgesamtheiten. Von einer Zufallsvariablen sei das Verteilungsgesetz bekannt, es enthalte jedoch unbekannte Parameter.

Diese sollen aufgrund einer Stichprobe $X_1, \ldots, X_n$ möglichst gut geschätzt werden.

Bei der Schätzung unbekannter Parameter unterscheidet man grundsätzlich zwei Methoden:

1) ***Punktschätzung:***
   Man erhält einen einzigen Wert aus der Stichprobe, welcher für die Schätzung herangezogen wird.

2) ***Intervallschätzungen:***
   Diese lassen Schlüsse über einen Bereich zu, welcher mit grosser Wahrscheinlichkeit den unbekannten Parameter enthält.

Ausgangspunkt für beide Ansätze bilden sogenannte ***Schätzfunktionen***

$$\hat{\theta}_n\left(X_1,\ldots,X_n\right),$$

welche angeben, in welcher Art und Weise die Stichprobenvariablen im Hinblick auf optimale Schätzungen zu verarbeiten sind.

Kapitel 13

8.1. Intuitiv heuristische Ansätze für Schätzfunktionen

In der Realität steht man oft vor dem Problem, dass man von einer Zufallsvariablen zwar das Verteilungsgesetz kennt, letzteres hingegen unbekannte Parameter enthält. Bezeichnet $f_X(x)$ die Wahrscheinlichkeitsfunktion (Dichte) einer Zufallsvariablen, so gilt für ihren Mittelwert:

$$\mu = E[X] = \sum_j x_j \; f_x\left(x_j\right).$$

μ soll über eine Stichprobe $X_1,\ldots, X_n$ aus einer Grundgesamtheit mit der Wahrscheinlichkeits-funktion f_x geschätzt werden.

Idee: Benutze das arithmetische Mittel

$$\hat{\mu} = t(X_1, \ldots, X_n) = \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i .$$

Eine solche Schätzung heisst **Punktschätzung**, weil ein punktueller Wert als Schätzwert genannt wird (und nicht etwa ein Intervall).

Beispiel 8.1.1: Aus der Grundgesamtheit der Studenten in einer Vorlesung wurde eine Stichprobe vom Umfang 10 gezogen. Die Körpergrösse *X* in cm wurde festgestellt und in folgender Tabelle erfasst:

i	1	2	3	4	5	6	7	8	9	10
x_i	176	180	181	168	177	186	184	173	182	177

Die Schätzung für die Körpergrösse der Studenten im Hörsaal lautet einfach:

$$\hat{\mu} = \bar{x}_n = 178.4 \text{ cm.}$$

Frage: *Handelt es sich um eine gute Schätzformel?*

Um eine solche Frage zu beantworten, muss man sogenannte Eigenschaften von Punktschätzungen einführen. Dies werden wir im nächsten Abschnitt machen.

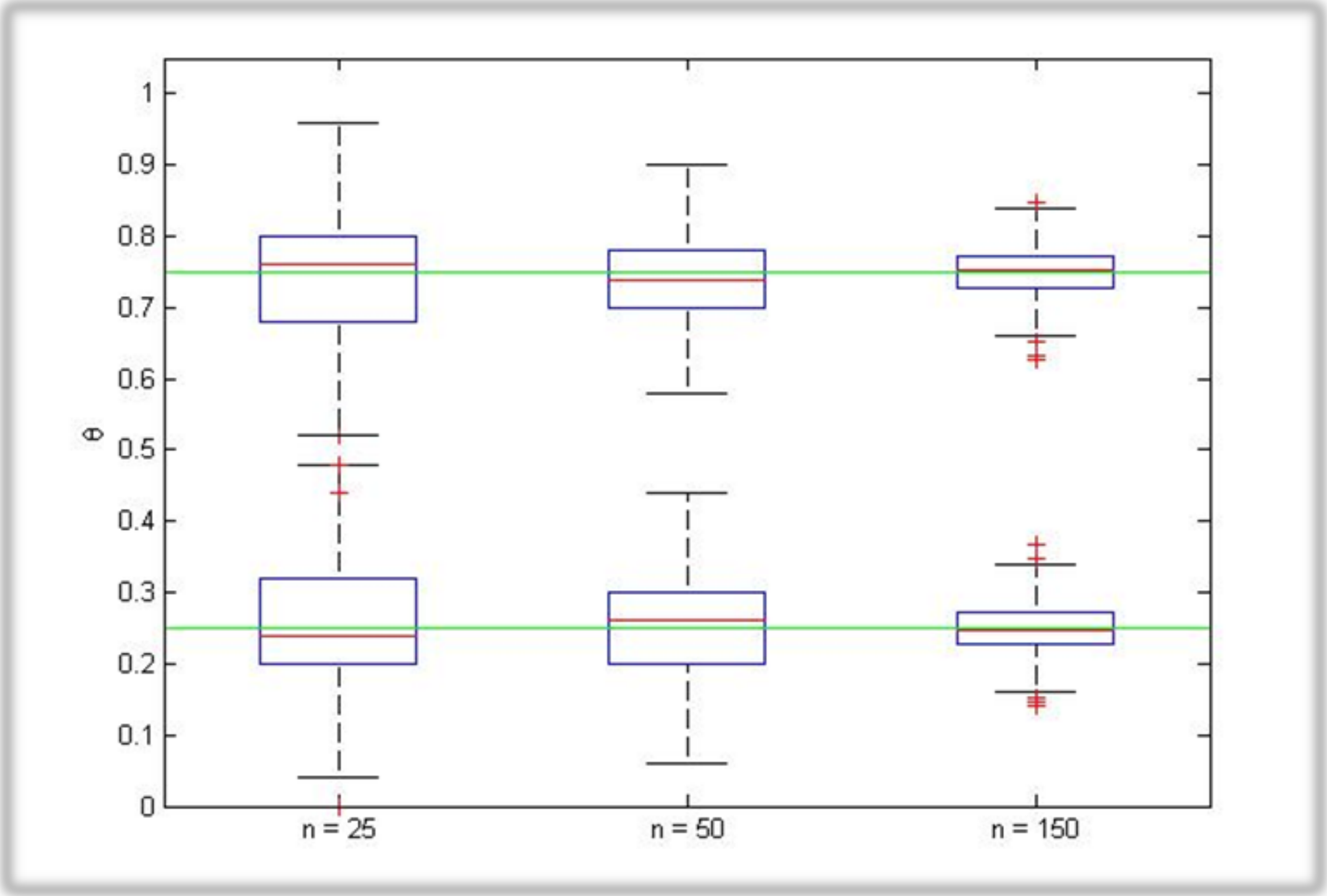
Würde man nach demselben Ansatz Schätzfunktionen für die Varianz σ^2 einer Zufallsvariablen oder für den Erfolgsanteil p in einem Binomialexperiment suchen, so wären folgende Schätzfunktionen angezeigt:

$$\hat{\sigma}^2 = t(X_1, \ldots, X_n) = S_x^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

respektive

$$\hat{p} = t(X_1, \ldots, X_n) = \frac{1}{n}\sum_{i=1}^{n}X_i .$$

Illustration: Binomialexperiment



Universität St.Gallen

Beispiel 8.1.2:

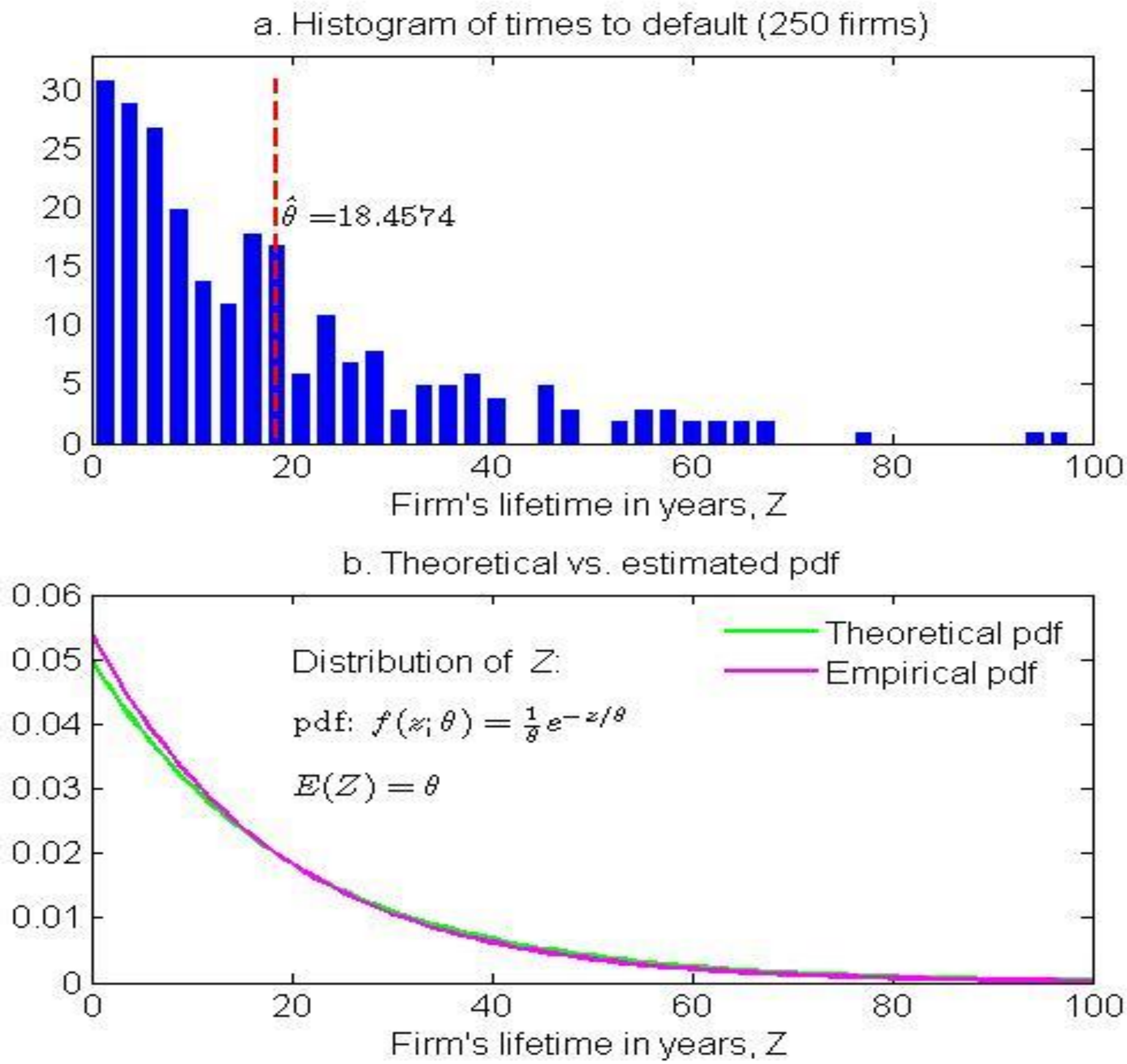
Mit den Beobachtungswerten aus *Beispiel 8.1.1* bekommen wir als Varianz der Stichprobe

$$S_x^2 = 25.84$$

und damit die Punktschätzung

$$\hat{\sigma}^2 = S_x^2 = 25.84.$$

Beispiel 8.1.3: Firms' lifetime



Schätzfunktionen $t(X_1,\ldots, X_n)$ sind ebenfalls Zufallsvariablen und unterliegen somit auch einem Verteilungsgesetz.

Notation:

- Symbol "^" bedeutet Schätzfunktion ("Dach" oder "Hat" auf Englisch)
- $T = t(X_1,\ldots, X_n)$ ist eine Zufallsvariable, während die auf den Realisationen x_i $(i = 1,\ldots, n)$ basierenden Werte Realisationen dieser Zufallsvariablen darstellen.
- $E[T] = \mu_T$ ist der Erwartungswert der Schätzfunktion T.
- $\mathrm{V(T)} = \mathrm{E}\left[\left(\mathrm{T} - \mu_\mathrm{T}\right)^2\right] = \sigma_\mathrm{T}{}^2$ ist die Varianz der Schätzfunktion T.

Für ein und denselben Parameter stehen oft mehrere Schätzfunktionen zur Verfügung.

Beispiel 8.1.4:

Sei X poissonverteilt mit Parameter λ, so gilt:

$$P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda} \quad \text{mit} \quad E[X] = V(X) = \lambda.$$

Soll nun

$$\hat{\lambda} = t(X_1, \ldots, X_n) = \overline{X},$$

$$\hat{\lambda} = t(X_1, \ldots, X_n) = S_x^{\,2}$$

oder andere Möglichkeiten gewählt werden?

Es stellt sich somit ein Bewertungsproblem für Schätzfunktionen. Ihre Qualität wird an wünschenswerten Eigenschaften gemessen.

Kapitel 13.2

8.2. Eigenschaften von Punktschätzungen

Definition

A) Erwartungstreue Schätzfunktionen

Eine Schätzfunktion $T = t(X_1, \ldots, X_n) = \hat{\theta}$ heisst erwartungstreu für θ falls: $E[T] = \mu_T$ existiert und $E[T] = \mu_T = \theta$.

Es erscheint vernünftig zu fragen, ob eine Schätzformel im Mittel den gesuchten Wert trifft.

Die Differenz $\text{bias} = \theta - E[T]$ heisst Verzerrung (Bias).

Beispiel 8.2.1:

i) $T = t(X_1,\dots, X_n) = \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ ist eine erwartungs-

treue Schätzfunktion für

$$\mu = E[X] \ \left(\text{d.h. } \mu = E[X_i],\ i = 1,\dots, n\right).$$

$$E[T] = E\left[\overline{X}_n\right] = E\left[{}^1\!/_n \sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n} \cdot n \cdot \mu = \mu.$$

ii) $T = t(X_1,\dots, X_n) = S^2 = {}^1\!/_n \sum_{i=1}^{n} (X_i - \overline{X})^2$ ist keine

erwartungstreue Schätzfunktion für $\sigma^2 = V(X)$

und ist um den Faktor $\frac{(n-1)}{n}$ verzerrt.

Beispiel 8.2.1 *(Fortsetzung)*:

$$E[T]=E\left[S^2\right]=\frac{1}{n}E\left[\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2\right]=\frac{1}{n}E\left[\sum_{i=1}^{n}\left((X_i-\mu)-\left(\overline{X}-\mu\right)\right)^2\right]$$

$$=\frac{1}{n}E\left[\sum_{i=1}^{n}(X_i-\mu)^2-n\left(\overline{X}-\mu\right)^2\right]$$

$$=\frac{1}{n}\left[\sum_{i=1}^{n}E\left[(X_i-\mu)^2\right]-n\,E\left[\left(\overline{X}-\mu\right)^2\right]\right]$$

$$=\frac{1}{n}\cdot\left[n\cdot V(X)-n\,V\left(\overline{X}\right)\right]$$

$$=\frac{1}{n}\left[n\cdot\sigma^2-n\cdot\frac{\sigma^2}{n}\right]=\frac{1}{n}\cdot\sigma^2\cdot(n-1)=\frac{n-1}{n}\cdot\sigma^2$$

=> Eine erwartungstreue Schätzfunktion für σ^2 ist somit

$$T=t\left(X_1,\ldots,X_n\right)=\frac{n}{n-1}S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2.$$

Beispiel 8.2.1 *(Fortsetzung)****:***

iii) Sind $Z_1,\ldots, Z_n$ iid Bernoulli Variablen mit

$$f_z(z) = p^z(1\text{-}p)^{1\text{-}z},\ z \in \{0,1\},$$

so gilt für die Schätzfunktion

$$Z = t(Z_1,\ldots, Z_n) = \overline{Z} = 1/n \sum_{i=1}^{n} Z_i;$$

$$E[Z] = E\left[1/n \sum_{i=1}^{n} Z_i\right] = p.$$

Der Erfolgsanteil in der Stichprobe schätzt die Erfolgswahrscheinlichkeit p eines Binomial-experimentes erwartungstreu.

Frage: Welche Schätzfunktion für V(Z)=p(1-p)?

Idee: $\overline{Z}(1-\overline{Z})$

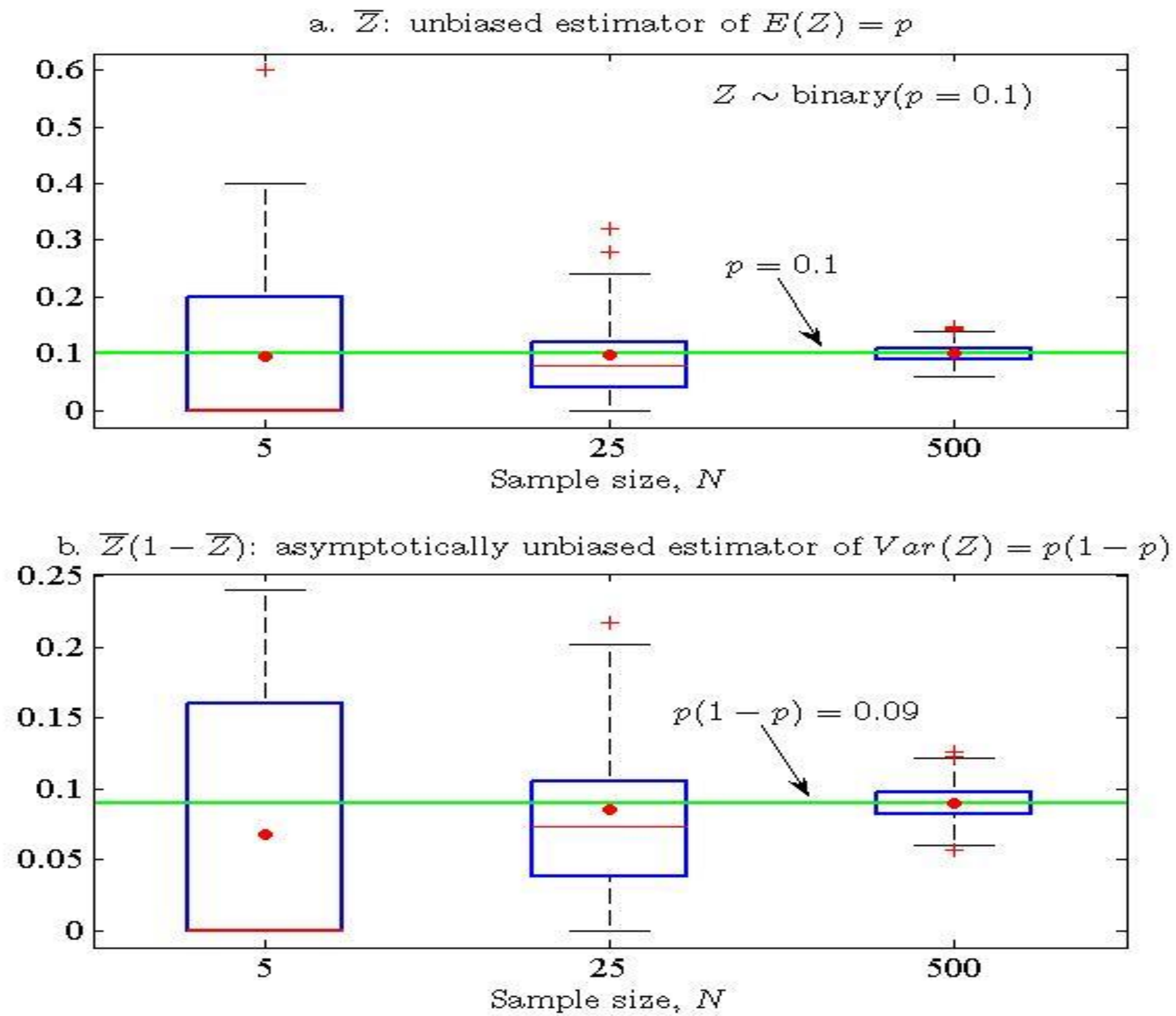
Ist diese Schätzfunktion erwartungstreu?

$$E\left[\overline{Z}(1-\overline{Z})\right] = E\left[\overline{Z}\right] - E\left[\overline{Z}^2\right] = p - \frac{n^2p^2 + np(1-p)}{n^2}$$

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

Aber: $E\left[\overline{Z}(1-\overline{Z})\right] \to p(1-p)$ für $n \to \infty$

Graphische Illustration:



Definition

B) Asymptotisch erwartungstreue Schätz-funktionen

Wenn die Verzerrung mit zunehmendem Stich-probenumfang geringer wird und für $n \to \infty$ verschwindet, dann heisst der Schätzer asymptotisch erwartungstreu:

$$\lim_{n \to \infty} \mathrm{E}\left[\mathrm{T}\right] = \theta.$$

Definition

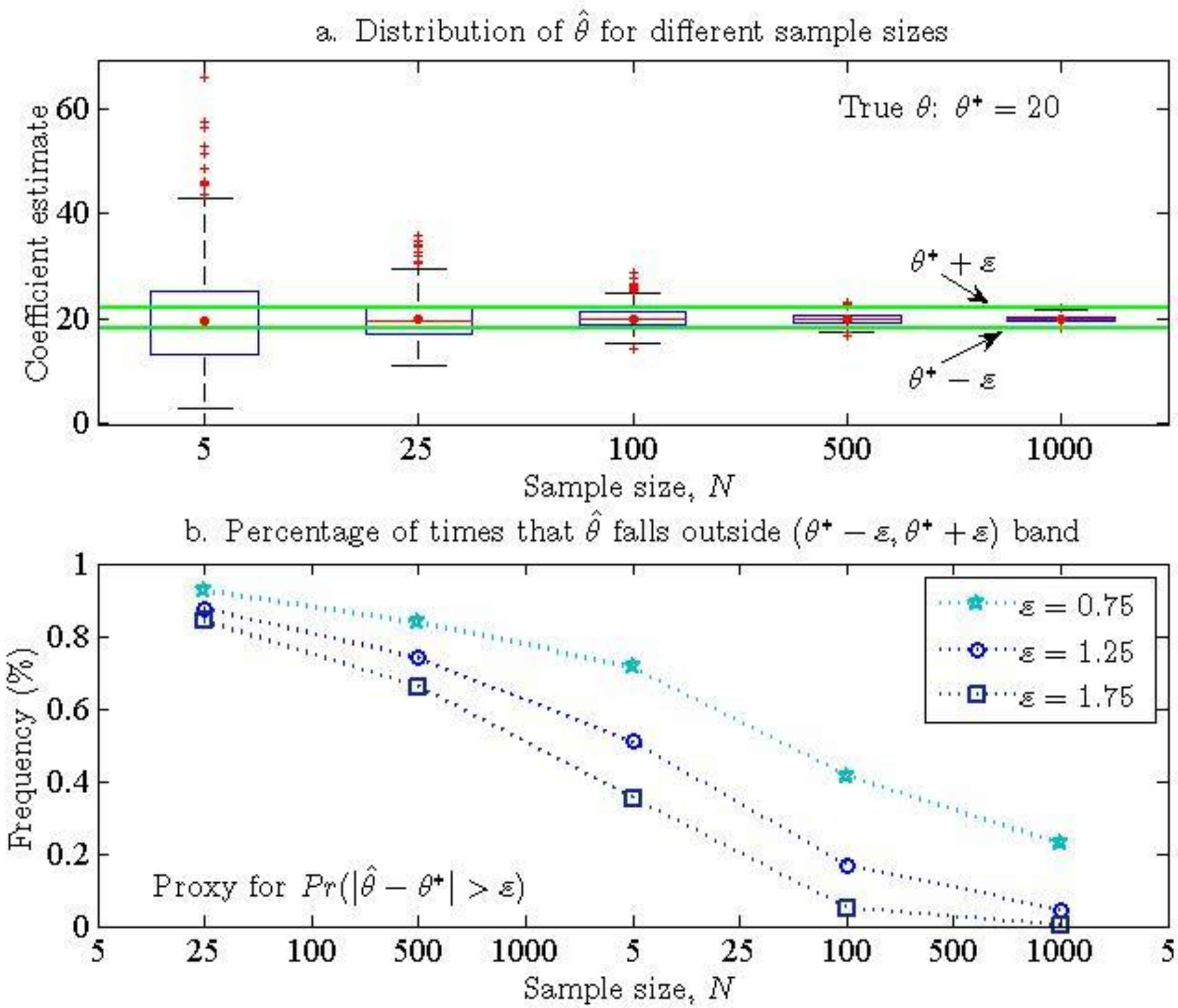
C) Konsistente Schätzfunktionen

Eine Schätzfunktion $\hat{\theta}_n = t\left(X_1, \ldots, X_n\right)$ für den unbekannten Parameter θ heisst konsistent, wenn die Folge $\left\{\hat{\theta}_n\right\}_n$, $n \to \infty$ stochastisch gegen θ konvergiert:

$$P\left(\left|\hat{\theta}_n - \theta\right| > \varepsilon\right) \to 0 \quad \text{für} \quad n \to \infty, \qquad \forall \varepsilon > 0.$$

Notation: $\hat{\theta}_n \xrightarrow{P} \theta \quad \text{für} \quad n \to \infty.$

Graphische Illustration der Konsistenzeigenschaft:



Universität St.Gallen

Theorem (Konsistente Schätzer):

Ein Schätzer ist konsistent, wenn er:

1. erwartungstreu ist (mindestens asymptotisch); und
2. wenn ausserdem seine Varianz für $n \rightarrow \infty$. gegen Null geht.

Beispiel 8.2.2:

$X_1, \dots, X_n$ sei eine Stichprobe aus einer Gesamtheit mit Erwartungswert μ und Standardabweichung σ.

Sei $$T_n = t(X_1, \dots, X_n) = \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Wir wissen: $\mathbf{T_n}$ ist erwartungstreu für μ.

Nun: $$V(T_n) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{\sigma^2}{n} \underset{n\to\infty}{\to} 0.$$

$\underset{\text{Regel}}{\Rightarrow}$ T_n ist konsistent für μ, oder $T_n \xrightarrow{p} \mu$.

Definition

D) Effiziente Schätzfunktion

Bezeichnen T sowie $U_1, U_2, \ldots, U_K$ erwartungstreue Schätzfunktionen für den unbekannten Parameter θ mit

$$E[T] = E[U_1] = \ldots = E[U_K] = \theta$$

so heisst T effizient, falls

$$V(T) \leq V(U_i), \quad i=1,\ldots,K.$$

Existieren also mehrere erwartungstreue Schätz-funktionen, so wählt man jene mit der kleinsten Varianz. (→ Sie liefert Schätzwerte, welche im Mittel am wenigsten vom wahren θ abweichen.)

Graphische Illustration der Effizienzeigenschaft (I)

Seien $X_1, \ldots, X_n$ Zufallsvariablen mit Erwartung μ und Varianz σ^2 und n eine gerade Zahl. Dann sind

$$\bar{X}_a = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$\bar{X}_b = \frac{2}{n}\sum_{i=1}^{n/2} X_{2i}$$

zwei erwartungstreue Schätzer für den Mittelwert, da

$$E(\bar{X}_a) = E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} \cdot n \cdot E(X_i) = \mu$$

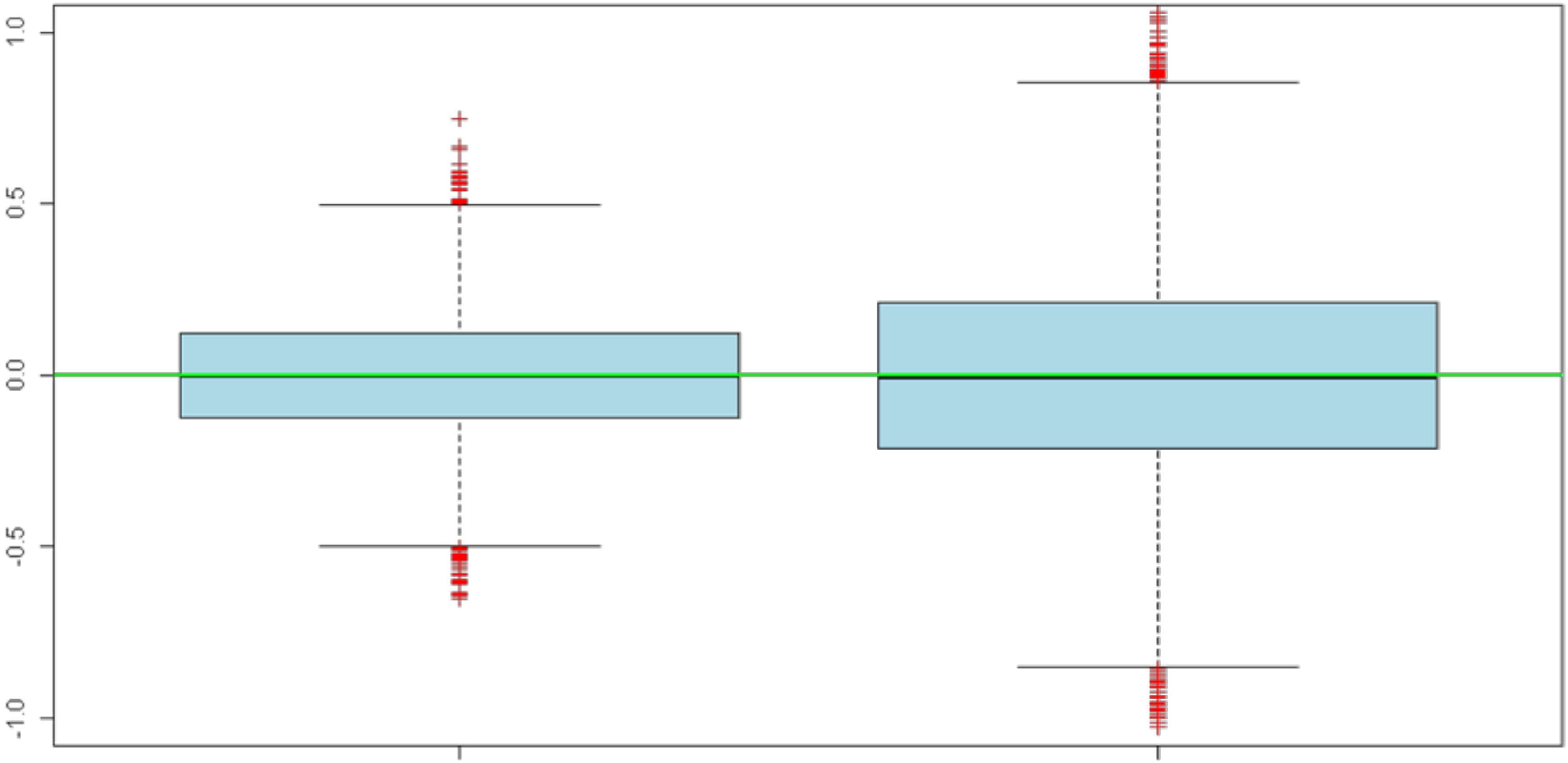
$$E(\bar{X}_b) = E(\frac{2}{n}\sum_{i=1}^{n/2} X_{2i}) = \frac{2}{n}\sum_{i=1}^{n/2} E(X_{2i}) = \frac{2}{n} \cdot n/2 \cdot E(X_{2i}) = \mu$$

Varianz?

Graphische Illustration der Effizienzeigenschaft (I, Fortsetzung)

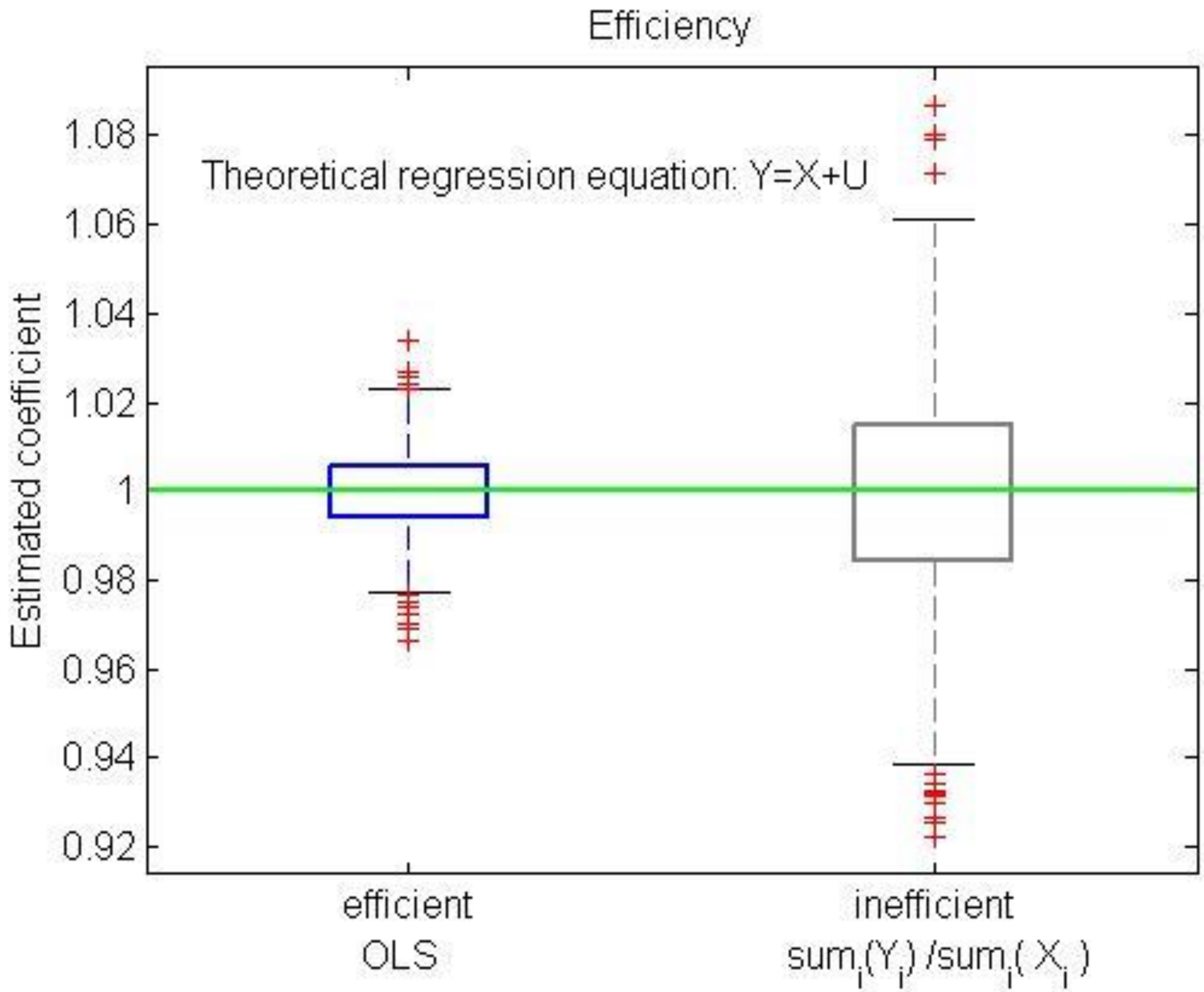
Es gilt, dass

$$Var(\bar{X}_a) = \frac{1}{n}\sigma^2 \text{ und } Var(\bar{X}_b) = \frac{2}{n}\sigma^2$$



$\bar{X}_a$ $\bar{X}_b$

Graphische Illustration der Effizienzeigenschaft (II)



Definition

E) Mittlerer quadratischer Fehler (MSE)

Der MSE ergänzt die bisherigen Kriterien zur Beurteilung der Güte von Schätzfunktionen.

Bezeichnet $T = t\left(X_1,\ldots, X_n\right)$ die Schätzfunktion für den unbekannten Parameter θ, so heisst

$$E\left[\left(T-\theta\right)^2\right] = MSE\left(\theta\right)$$

mittlerer quadratischer Fehler der Schätzfunktion.

$$\text{Nun: } \underline{\mathrm{MSE}(\theta)} = E\left[(T-\theta)^2\right] = E\left[\left(T-\mu_T-(\theta-\mu_T)\right)^2\right]$$

$$= E\left[(T-\mu_T)^2\right] - 2(\theta-\mu_T)\cdot \underbrace{E\left[T-\mu_T\right]}_{=0} + (\theta-\mu_T)^2$$

$$= E\left[(T-\mu_T)^2\right] + (\theta-\mu_T)^2$$

$$= \underline{V(T) + \mathrm{bias}^2}$$

Universität St.Gallen

Kapitel 13

8.3. Methoden zur Konstruktion von Schätzfunktionen

Bisher beschränkten sich die Überlegungen auf die Diskussion der Qualität von Schätzfunktionen.

Frage: Nach welchen Methoden und Prinzipien können solche Schätzfunktionen konstruiert werden?

Es gibt eine reiche Palette von möglichen Ansätzen:

- Momentenmethode
- Minimumquadratmethode
- Maximum-Likelihood-Methode

A) Momentenmethode

Regel: Man schätze die Momente der Verteilung der Grundgesamtheit mit den entsprechenden Momenten der Stichprobe.

Grundgesamtheit	Stichprobe
$E[X] = \mu$	$\frac{1}{n}\sum_{j=1}^{n} X_j$
$E\left[X^2\right] = \mu^2 + \sigma^2$	$\frac{1}{n}\sum_{j=1}^{n} X_j^2$
.	.
.	.
.	.
$E\left[X^m\right]$	$\frac{1}{n}\sum_{j=1}^{n} X_j^m$

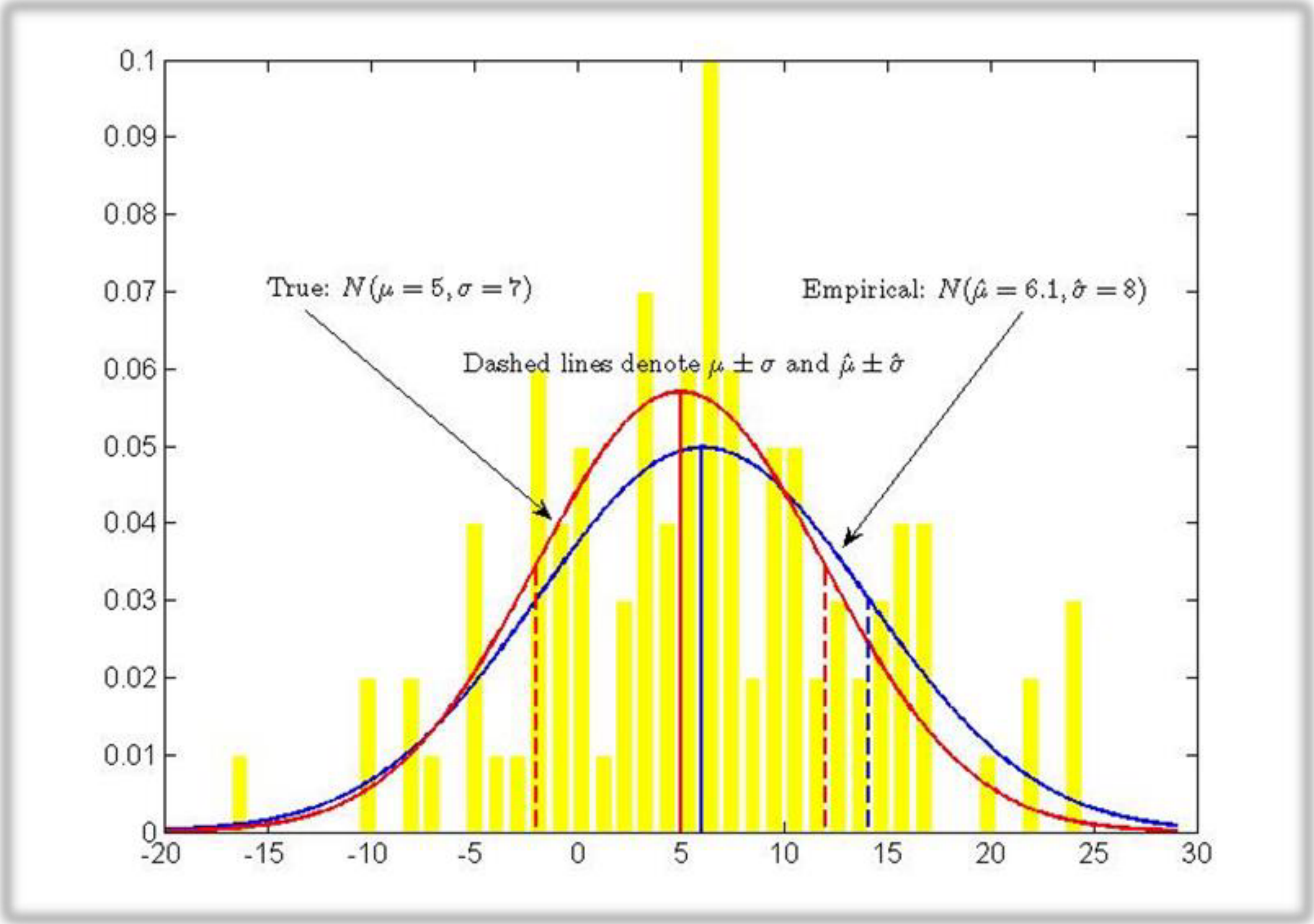
Daraus lassen sich die Schätzformeln

$$\hat{\mu} = \overline{X}_n$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 = S^2$$

herleiten.

Illustration: Normalverteilung



B) Minimumquadratmethode

Regel: Auf den unbekannten Mittelwert μ angewandt, schlägt das Prinzip vor, diejenige Grösse $\hat{\mu}_{KQ}$ als Schätzfunktion zu wählen, von der die Summe der quadrierten Abstände zu den Stichprobenwerten X_j minimal ist: $\hat{\mu}_{KQ} = \underset{\mu}{\operatorname{argmin}} \sum_{i=1}^{n} (X_i - \mu)^2.$

Lösung: $$\hat{\mu}_{KQ} = \frac{1}{n} \sum_{i=1}^{n} X_i = \overline{X}_n.$$

Universität St.Gallen

C) Die Maximum-Likelihood-Methode
("Methode der maximalen Mutmasslichkeit")

Einführungsbeispiel:

Wir betrachten eine Urne mit 2 Sorten Kugeln (S, W) im Mischungsverhältnis 1:3.

Der Anteil p der Erfolgskugeln nimmt dann entweder den Wert 0.25 oder 0.75 an.

Der Entscheid für p soll aufgrund einer konkreten Stichprobe X_1, X_2, X_3 mit Zurücklegen vom Umfang n=3 gefällt werden.

$$X \ : \ \ f_{Bi}(x\,;n,p); \ \ P(X=x) = \binom{3}{x} p^x (1\text{-}p)^{3\text{-}x} \ , \ \ x = 0,1,2,3.$$

Universität St.Gallen

Problem:

Aufgrund einer konkreten Beobachtung von X schätze man die Erfolgswahrscheinlichkeit p.

p	X	0	1	2	3
0.25	$f_{Bin}(x\,;3\,,0.25)$	$27/64$	$27/64$	$9/64$	$1/64$
0.75	$f_{Bin}(x\,;3\,,0.75)$	$1/64$	$9/64$	$27/64$	$27/64$

$f_{Bi}(x\,;3\,,p)$ hängt im Falle einer konkreten Realisation von X nur noch von p ab.

Man bezeichnet diese Funktion als **Likelihoodfunktion** $L(p\,;x)$.

Annahme: Die Stichprobe zeigt einen Erfolg (d.h. X=1).

Aufgrund dieser Beobachtung: $\hat{p} = 0.25.$

Maximum Likelihood Prinzip:

$X_1, \ldots, X_n$ sei eine Zufallsstichprobe aus einer Grundgesamtheit mit bekanntem Verteilungsgesetz f_X und zu schätzendem Parameter θ.

Dann wird die Funktion:

$$L(\theta; x_1, \ldots, x_n) = f_{(x_1, \ldots, x_n)}(\theta; x_1, \ldots, x_n)$$

$$= \prod_{i=1}^{n} f_x(\theta; x_i)$$

als ***Likelihoodfunktion*** bezeichnet.

Nun:

$$\hat{\theta}_{ML} = \underset{\theta}{\operatorname{argmax}} \quad L\left(\theta\,; x_1,\ldots, x_n\right)$$

$$= \underset{\theta}{\operatorname{argmax}} \quad \log L\left(\theta\,; x_1,\ldots, x_n\right).$$

ist ein ***Maximumlikelihood-Schätzer*** für θ .

Beispiel 8.3.1: ML-Schätzer für p einer Bernoulli-verteilten Grundgesamtheit:

$$L(p\,;x_1,\ldots,x_n) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \quad , \; x_i \in \{0,1\}.$$

$$\log L(p\,;x_1,\ldots,x_n) = \sum_{i=1}^{n}\left(x_i \log p + (1-x_i)\log(1-p)\right)$$

$$\frac{\partial \log}{\partial p} L(p\,;x_1,\ldots,x_n) = \frac{\sum_{i=1}^{n} x_i}{\hat{p}} - \frac{\sum_{i=1}^{n}(1-x_i)}{1-\hat{p}} \overset{!}{=} 0$$

$$\Leftrightarrow \quad (1-\hat{p})\cdot\sum_{i=1}^{n} x_i = \hat{p}\cdot\left(n - \sum_{i=1}^{n} x_i\right)$$

$$\Rightarrow \quad \hat{p}_{ML} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}_n.$$

Beispiel 8.3.2: $X_1,\ldots, X_n$ sei eine Zufallsstichprobe aus einer stetig gleichverteilten Grundgesamtheit:

$$f_R(X)=\begin{cases} \dfrac{1}{\theta} & , \quad 0 \le x \le \theta \\ 0 & , \quad \text{sonst} \end{cases}$$

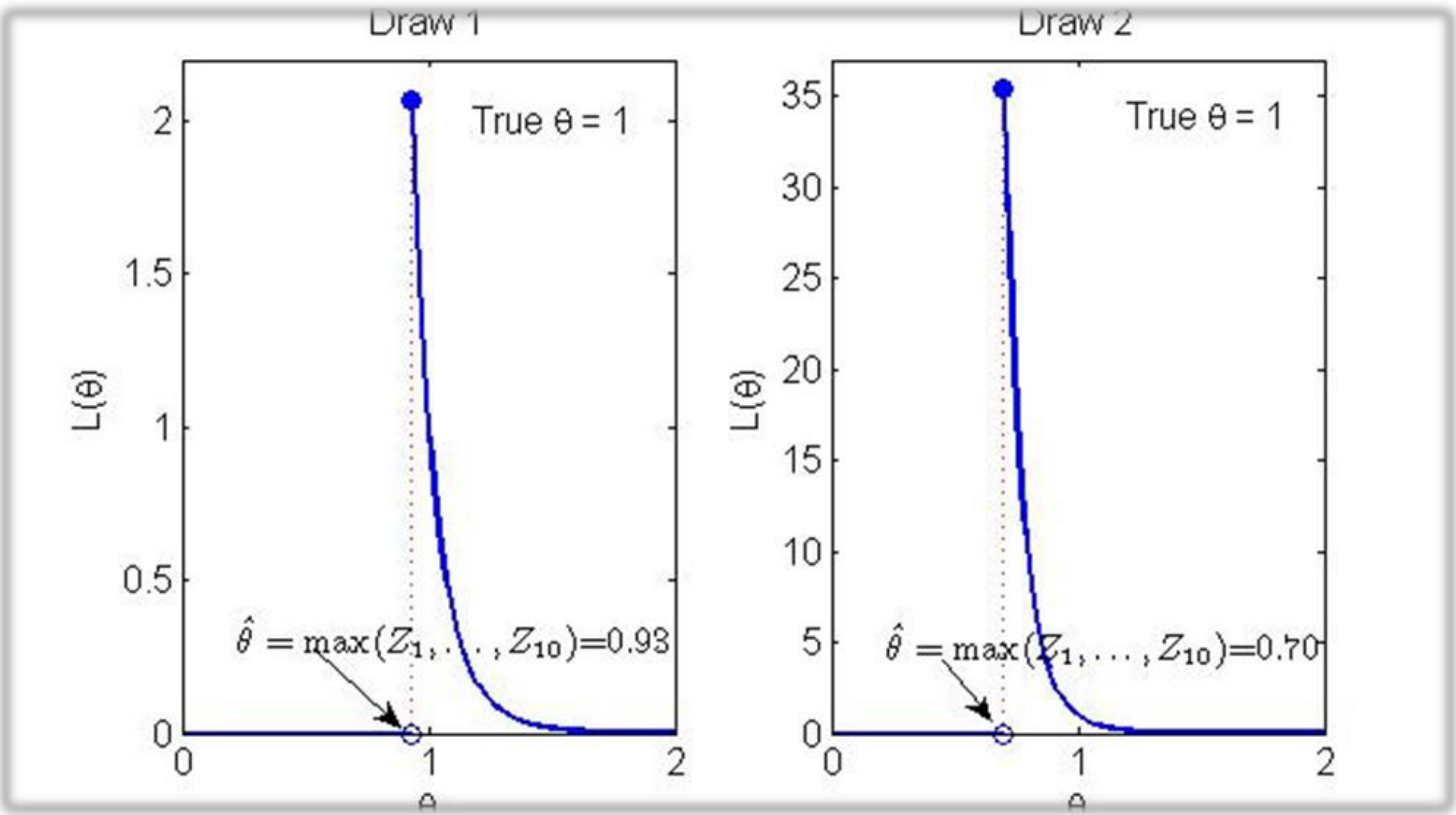
$$L(\theta\,;\, x_1,\ldots, x_n) = \left(\frac{1}{\theta}\right)^n, \qquad \text{falls alle } x_i \in [0,\theta].$$

L ist streng monoton fallend in θ, aber: alle $x_i \in [0,\theta]$,

d.h. $\theta \ge x_i$, $i=1,\ldots, n$

$$\Rightarrow \; \hat{\theta}_{ML} = \max(X_1,\ldots, X_n).$$

Graphische Illustration von Beispiel 8.3.2:



Eigenschaften von Likelihoodschätzfunktionen

1. Maximumlikelihood-Schätzfunktionen genügen dem Invarianzprinzip:

   Ist $\hat{\theta}$ eine ML-Schätzung für θ und g(.) eine eineindeutige Abbildung von θ, so ist $g(\hat{\theta})$ eine ML-Schätzung für $g(\theta)$.

2. ML-Schätzfunktionen sind konsistent und asymptotisch erwartungstreu.

9. Intervallschätzungen-Konfidenzintervalle

9.1. Konzept des Konfidenzintervalls

Während bei der Punktschätzung aus einem Stichprobenergebnis nur ein punktueller Schätzwert ermittelt wird, informiert eine **Intervallschätzung** zusätzlich über den **Stichprobenfehler**.
Sie stellt einen Zusammenhang zwischen der Punktschätzung und dem Parameterwert der Grundgesamtheit her.

Definition

Symmetrisches (1-α)-Konfidenzintervall:

$$\mathrm{KONF}_{1-\alpha}(\theta) = \left[\hat{\theta}_n - f_n \; ; \; \hat{\theta}_n + f_n\right],$$

wobei der Stichprobenfehler f_n so bestimmt wird, dass das Konfidenzintervall den unbekannten Parameter θ mit einer vorgegebenen Wahrscheinlichkeit $(1-\alpha)$ überdeckt:

$$P\left[\theta \in \mathrm{KONF}_{1-\alpha}(\theta)\right] = 1-\alpha.$$

Die Wahrscheinlichkeit (1-α) heisst ***Überdeckungs-wahrscheinlichkeit*** oder ***Konfidenzniveau***.

Üblicherweise werden Konfidenzniveaus von 0.95 oder 0.99 gewählt.

Interpretation:

Das Konfidenzintervall ist als ein (aus dem Stichprobenergebnis resultierendes) Zufallsintervall für einen unbekannten, aber numerisch fixierten Parameter zu interpretieren.

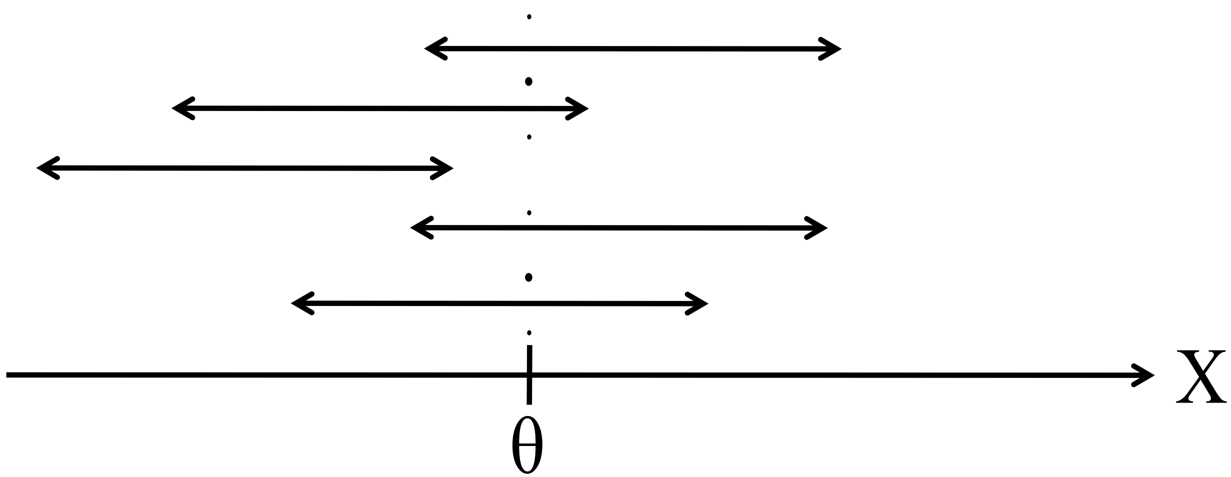


Abbildung: Aus verschiedenen Stichproben berechnete Konfidenzintervalle für θ.

9.2. Ableitung von Konfidenzintervallen (bei grossen Stichproben)

Exemplarisch betrachten wir die Ableitung des (1-α)-Konfidenzintervalls für das arithmetische Mittel μ der Grundgesamtheit.

$$\text{ZGS:} \qquad P\left[\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \le q\right] \underset{n\to\infty}{\longrightarrow} F_Z(q).$$

Daraus folgt, dass (für n gross genug):

$$P\left[-q_{1-\frac{\alpha}{2}} \le \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \le q_{1-\frac{\alpha}{2}}\right] \cong 1\text{-}\alpha,$$

wobei $q_{1-\frac{\alpha}{2}}$ das $\left(1-\frac{\alpha}{2}\right)$–Quantil der Standardnormalverteilung bezeichnet.

Universität St.Gallen

Äquivalent können wir umformen und erhalten für das $(1\text{-}\alpha)$ -Konfidenzintervall für μ:

$$\mathrm{KONF}_{1\text{-}\alpha}(\mu) = \left[\overline{X}_n - q_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \;\; ; \;\; \overline{X}_n + q_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

Wenn auch σ unbekannt ist und $n \geq 50$, dann ist

$$\mathrm{KONF}_{1\text{-}\alpha}(\mu) = \left[\overline{X}_n - q_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} \;\; ; \;\; \overline{X}_n + q_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}}\right]$$

das (1- α) - Konfidenzintervall für μ, wobei

$$S_n^{\;2} = {}^{1}\!/_{n} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2 \;.$$

Universität St.Gallen

Beispiel 9.2.1: (Buch, Seiten 452-453)

Für die Berechnung der Daten für den örtlichen Mietspiegel befragt die Gemeindeverwaltung 50 Haushalte nach der Kaltmiete pro m^2 .

Resultat: $\overline{X}_{50} = 8.30$ € und $S_{50} = 2.07$ € .

Wie gross ist das Konfidenzintervall zum Niveau 0.9 für die durchschnittliche Kaltmiete μ?

$$\mathrm{KONF}_{90\%}(\mu) =$$

$$\left[8.30 - 1.645 \cdot \frac{2.07}{\sqrt{50}} \; ; 8.30 + 1.645 \cdot \frac{2.07}{\sqrt{50}}\right] = [7.82 \; ; \; 8.78]$$

9.3. Zusammenhang mit Hypothesentests

Betrachte erneut das statistische Problem bezüglich eines Parameters θ dessen wahrer Wert unbekannt ist, von welchem allerdings bekannt ist das er in einem Parameterraum Θ liegt.

Angenommen Θ kann in zwei disjunkte Teilmengen Θ_0 und Θ_1 geteilt werden, und wir möchten überprüfen ob θ in Θ_0 oder in Θ_1 liegt.

Ein Problem dieser Art wird als **Hypothesentestproblem** bezeichnet. Die beobachteten Werte liefern Informationen über θ um eine Entscheidung zu treffen.

$$H_0:\ \theta \epsilon\ \Theta_0$$
$$H_1:\ \theta \epsilon\ \Theta_1$$

werden als **Ziel (Null)** und ***A*lternativ Hypothese** bezeichnet.

Wird bei der Durchführung eines Tests entschieden dass θ in Θ_1 liegt, so spricht man davon dass die **Nullhypothese verworfen** wird. Entscheidet man das θ in Θ_0 leigt, so sagt man das H_0 nicht verworfen wird.

Eine Möglichkeit, um eine solche Entscheidung zu treffen ist ein Konfidenzintervall zu konstruieren.

In diesem Fall sind wir an Hypothesen folgender Art interessiert:

$$H_0\colon \theta = \theta_0$$
$$H_1\colon \theta \neq \theta_0.$$

Beispiel: Mittelwert einer Verteilung (in grossen Stichproben)

Idee: Verwirf H_0: μ= μ_0 falls der Abstand zwischen dem arithmetischen Mittel und μ_0 gross genug ist:

$$|\overline{X}_n - \mu_0| \geq c,$$

mit *c* bestimmt wird durch das **Signifikanz Niveau** α des Tests: $c = q_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$.

Falls $\overline{X}_n = \overline{x}_n$ beobachtet wird, dann ist die Menge der μ_0 so dass die H_0 verworfen wird gleich der Menge der μ_0 so dass

$$|\overline{x}_n - \mu_0| < q_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

falls σ bekannt ist (ansonsten siehe Folie 367).

Diese Ungleichung kann einfach in die Formel für das Konfidenzintervall auf Folie 367 umgeschrieben werden ($\mu = \mu_0$):

$$\mathrm{KONF}_{1-\alpha}(\mu) = \left[\overline{X}_n - q_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \;\; ; \;\; \overline{X}_n + q_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

10. Hypothesentests

10.1. Arten von Hypothesen

Betrachte eine Zufallsstichprobe $X_1, \ldots, X_n$ einer gegebenen Verteilung f_X, abhängig von einem Parameter θ welcher im Parameterraum Θ liegt. Es soll die folgende Hypothese getestet werden

$$H_0 \colon \ \theta \in \Theta_0$$
$$H_1 \colon \ \theta \in \Theta_1$$

mit $\Theta_0 \cup \Theta_1 = \Theta$.

Wenn $\Theta_i, i = 1{,}2$, nur einen einzigen Wert von θ enthält, so ist H_i eine **einfache Hypothese**. Andernfalls ist H_i eine **zusammengesetzte Hypothese**.

Darüber hinaus haben **einseitige** Hypothesen für einen eindimensionalen Parameter θ die Form

$$H_i\colon\ \theta \geq \theta_0 \text{ or } H_i\colon\ \theta > \theta_0$$

wohingegen **zweiseitige** Hypothesen die Form

$$H_i\colon\ \theta \neq \theta_0$$

annehmen.

Beispiel: Mittelwert einer Normalverteilung mit bekannter Varianz:

$$H_0\colon \mu = \mu_0 \text{ (einfach)}$$

$$H_1\colon \mu \neq \mu_0 \text{(zweiseitig, zusammengesetzt)}$$

Oder wenn eine bestimmte Richtung getestet wird:

$$H_1\colon \mu > \mu_0 \text{ oder } H_1\colon \mu < \mu_0 \text{ (einseitig, zusammengesetzt)}$$

10.2. Kritischer Bereich und Teststatistik

Betrachte erneut folgende Hypothese:

$$H_0: \ \theta \in \Theta_0$$
$$H_1: \ \theta \in \Theta_1$$

mit $\Theta_0 \cup \Theta_1 = \Theta$.

Man nehme an, dass wir, bevor wir entscheiden welche Hypothese wir wählen, eine Zufallsstichprobe $X_1,\ldots, X_n$ einer Verteilung f_X, welche von einem unbekannten Parameter θ abhängt, beobachten können.

Sei S der Stichprobenraum des n-dimensionalen Zufallsvektors $\mathbf{X}=(X_1,\ldots, X_n)$. S ist die Menge aller möglichen Werte der Zufallsstichprobe.

In einem solchen Hypothesentestproblem können wir ein Testverfahren basierend auf den beobachteten Ereignissen spezifizieren, welches S in zwei Teilmengen unterteilt:
S_1 enthält die Werte von **X** sodass H_0 verworfen wird, während S_o die Werte enthält bei denen H_0 nicht verworfen wird.

Die Menge S_1 wird als **kritischer Bereich** des Tests bezeichnet.

In den meisten Hypothesentestproblemen wird der kritische Bereich basierend auf einer **Teststatistik** definiert:

$$T = f(X_1, \ldots, Xn).$$

Beispiel: (fortgesetzt) Mittelwert einer Normalverteilung mit bekannter Varianz:

Für eine zweiseitige Alternativhypothese wird H_0 verworfen, wenn der Stichprobenmittelwert $\overline{X}_n$ weit von μ_0 entfernt ist. Dies führt zur Verwendung folgender Teststatistik:

$$T = |\overline{X}_n - \mu_0|$$

und H_0 wird verworfen falls $T \geq c$.

Universität St.Gallen

10.3. Gütefunktion und Arten von Fehlern

Aufgrund von Unsicherheit besteht bei jedem Testverfahren die Möglichkeit von Fehlern.

Ein **Fehler erster Art** ist die Fehlentscheidung eine wahre Nullhypothese zu verwerfen. Die Wahrscheinlichkeit eines Fehlers erster Art nennt man **Signifikanzniveau** des Tests.

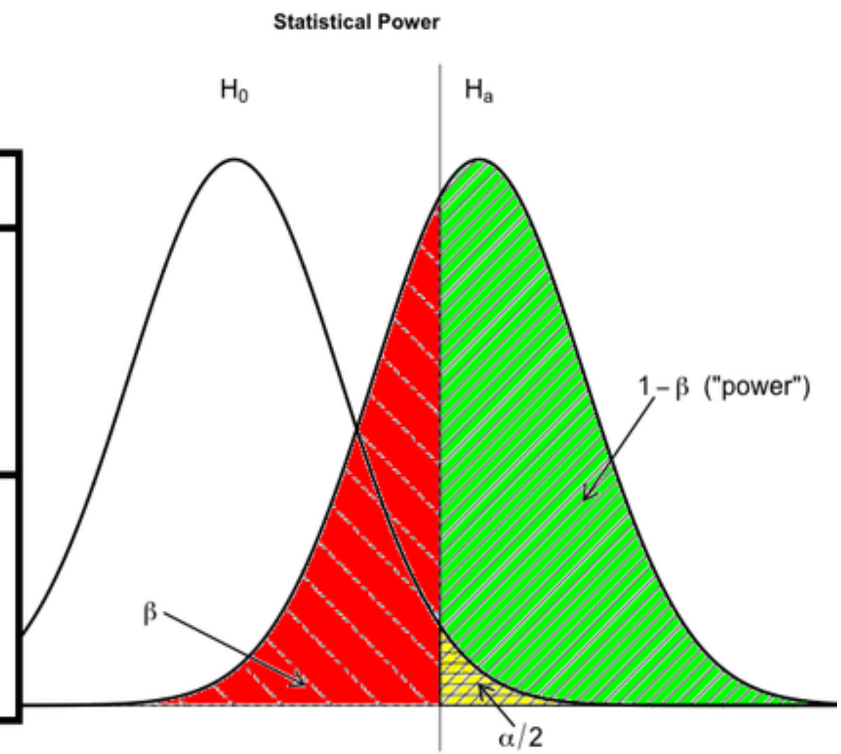
Ein **Fehler zweiter Art** ist die Fehlentscheidung eine falsche Nullhypothese nicht zu verwerfen. Für einen gegebenen Wert unter der Alternativhypothese bezeichnet man eins minus der Wahrscheinlichkeit für einen Fehles zweiter Art als **Güte** des Tests.

Beispiel: (fortgesetzt) Mittelwert einer Normalverteilung mit bekannter Varianz :

$$H_0\text{: } \mu = \mu_0$$
$$H_1\text{: } \mu \neq \mu_0$$

Decision

Null Hypothesis (H_0)	Accept H_0	Reject H_0
True	**Correct** "Confidence Level" Probability = 1-α	**Type I Error** "False Positive" Probability = α
False	**Type II Error** "False Negative" Probability = β	**Correct** "Statistical Power" Probability = 1-β



Der kritische Bereich eines Tests wird grundsätzlich festgelegt, indem das Signifikanzniveau des Tests α fixiert wird.

Betrachte die einfache Nullhypothese $H_0\colon\ \theta = \theta_0$. α kann in Bezug auf die Teststatistik T und den Ablehnbereich R berechnet werden als:

$$P(T \epsilon\, R|\theta_0) = P(\boldsymbol{X} \epsilon\, S_1|\theta_0) = \alpha.$$

Übliche Werte für α sind 10%, 5%, und 1%.

Beispiel: (fortgesetzt) Mittelwert einer Normalverteilung mit bekannter Varianz :

$$H_0\colon \mu = \mu_0 \;;\; H_1\colon \mu \neq \mu_0$$

Basierend auf der Teststatistik $T = |\overline{X}_n - \mu_0|$ mit Ablehnbereich $R = [c, \infty)$ folgt

$$\alpha = P(T \epsilon\, R|\mu_0) = P(|\overline{X}_n - \mu_0| \geq c|\mu_0)$$
$$= 1 - P(-c < \overline{X}_n - \mu_0 < c|\mu_0)$$
$$= 1 - P\left(\frac{\overline{X}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}} < \frac{c}{\frac{\sigma}{\sqrt{n}}}\middle|\mu_0\right) + P\left(\frac{\overline{X}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}} < \frac{-c}{\frac{\sigma}{\sqrt{n}}}\middle|\mu_0\right)$$
$$= 1 - \Phi\left(\sqrt{n}\frac{c}{\sigma}\right) + \Phi\left(\sqrt{n}\frac{-c}{\sigma}\right) = 2\Phi\left(\sqrt{n}\frac{-c}{\sigma}\right).$$

Es folgt

$$\Phi\left(\sqrt{n}\frac{-c}{\sigma}\right) = \frac{\alpha}{2}$$

und

$$-c = q_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

Oder äquivalent

$$c = q_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

mit q_α dem α-Quantil der Standardnormalverteilung.
Die Nullhypothese mit Signifikanzniveau α wird verworfen wenn der Abstand zwischen Stichprobenmittelwert und μ_0 grösser als c ist.

Für ein Testverfahren δ wird die Funktion

$$\pi(\theta|\delta) = P(\boldsymbol{X} \epsilon\, S_1|\theta), \text{ for } \theta\ \epsilon\ \Theta,$$

als **Gütefunktion** bezeichnet. Wird das Testproblem in Bezug auf die Teststatistik T und den Ablehnbereich R formuliert, so lautet die Gütefunktion

$$\pi(\theta|\delta) = P(T \epsilon\, R|\theta), \text{ for } \theta\ \epsilon\ \Theta.$$

Die Gütefunktion gibt für jeden möglichen Wert des Parameters θ die Wahrscheinlichkeit an, dass die Nullhypothese abgelehnt wird.

Beispiel: (fortgesetzt) Mittelwert einer Normalverteilung mit bekannter Varianz:

$$H_0\colon \mu = \mu_0 \;;\; H_1\colon \mu \neq \mu_0$$

Basierend auf der Teststatistik $T = |\overline{X}_n - \mu_0|$ mit Ablehnbereich $R = [c, \infty)$ es folgt für alle $\mu \epsilon \mathbb{R}$

$$\pi(\mu|\delta) = P(T \epsilon R|\mu) = P(|\overline{X}_n - \mu_0| \geq c|\mu)$$

$$= 1 - P(-c < \overline{X}_n - \mu_0 < c|\mu)$$

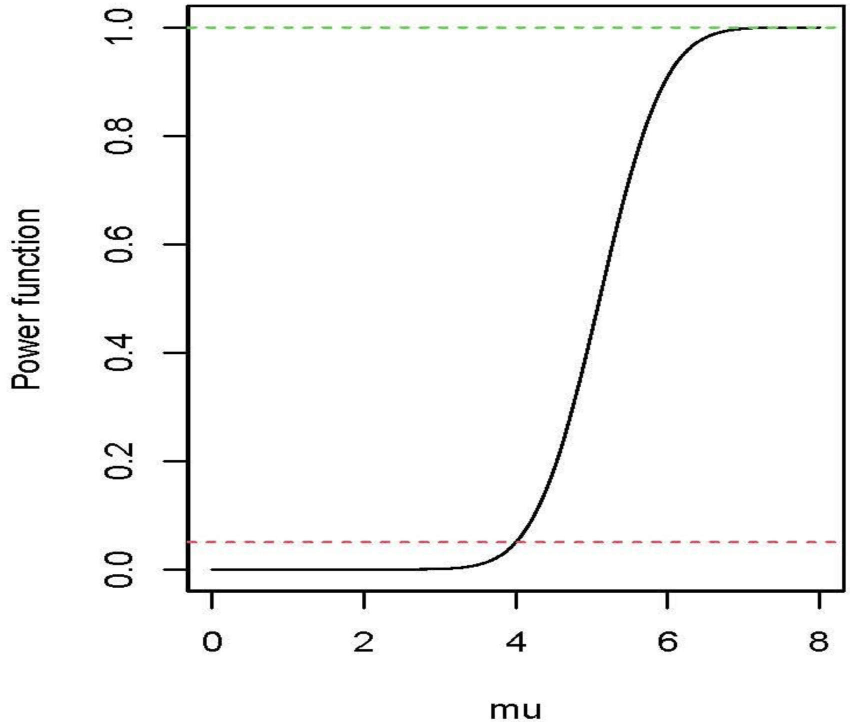
$$= 1 - P\left(\frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} < \frac{c + \mu_0 - \mu}{\frac{\sigma}{\sqrt{n}}} \middle| \mu\right) + P\left(\frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} < \frac{-c + \mu_0 - \mu}{\frac{\sigma}{\sqrt{n}}} \middle| \mu\right)$$

$$= 1 - \Phi\left(\frac{c + \mu_0 - \mu}{\frac{\sigma}{\sqrt{n}}}\right) + \Phi\left(\frac{-c + \mu_0 - \mu}{\frac{\sigma}{\sqrt{n}}}\right).$$

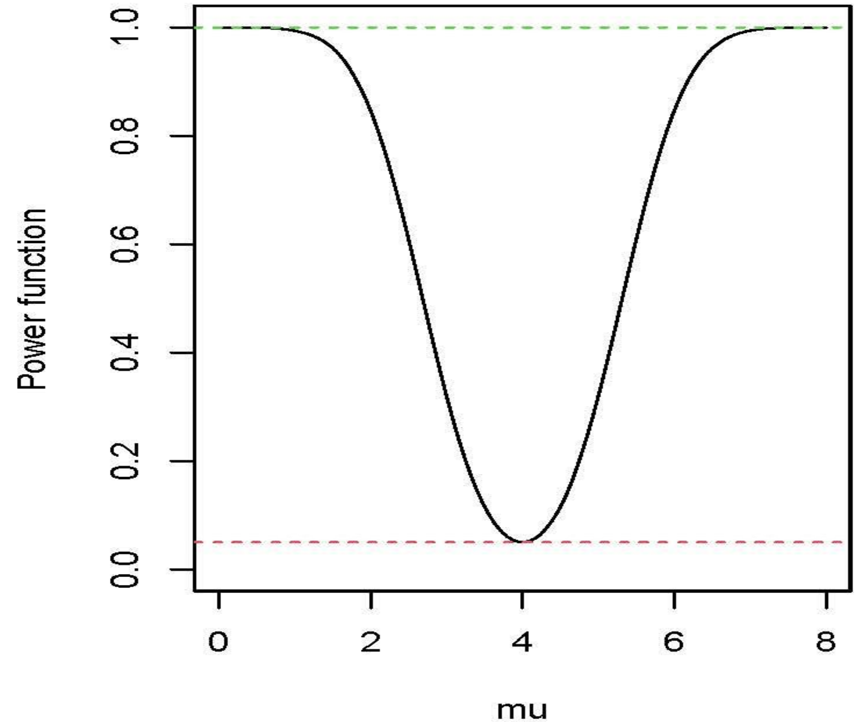
Veranschaulichendes Beispiel: $\mu_0 = 4, n = 15, \sigma = 3$
Signifikanzniveau des Tests: $\alpha = 0.05$ (gestrichelte rote Linie)

H_0: $\mu = \mu_0$

H_1: $\mu > \mu_0$



H_1: $\mu \neq \mu_0$



10.4. Der p-Wert

Das Ergebnis eines Tests hängt vom gewählten Niveau α ab. Das bedeutet, dass bei unterschiedlichen Signifikanzniveaus die Schlussfolgerung des Tests unterschiedlich ausfallen könnte.

Um dieses Problem zu umgehen, wird die Idee des p-Werts eingeführt.

Der **p-Wert** ist das kleinste Niveau α für welches die Nullhypothese mit Niveau α mit den beobachteten Daten abgelehnt würde.

Beispiel: (fortgesetzt) Mittelwert einer Normalverteilung mit bekannter Varianz:

$$H_0\colon \mu = \mu_0 \; ; \; H_1\colon \mu \neq \mu_0$$

Wir beobachten in der Stichprobe den Wert 2.78 für die standardisierte Statistik

$$Z = \frac{\sqrt{n}\,T}{\sigma} = \sqrt{n}\,\frac{|\overline{X}_n - \mu_0|}{\sigma}.$$

Der p-Wert kann berechnet werden, indem das kleinste α gefunden wird, welches die Bedingung

$$2.78 \geq q_{1-\frac{\alpha}{2}}$$

erfüllt. Das ist $\alpha = 0.0054$.

Teil I: Wahrscheinlichkeitsrechnung

1. Grundlagen der Wahrscheinlichkeit
   - 1.1. Ereignisse, Ereignisraum und Ereignismenge
   - 1.2. Das Rechnen mit Ereignissen
   - 1.3. Der Wahrscheinlichkeitsbegriff
   - 1.4. Axiomatik der Wahrscheinlichkeitstheorie
   - 1.5. Wichtige Regeln der Wahrscheinlichkeitsrechnung
   - 1.6. Wahrscheinlichkeitsräume
   - 1.7. Bedingte Wahrscheinlichkeit und stochastische Unabhängigkeit
   - 1.8. Totale Wahrscheinlichkeit
   - 1.9. Das Bayes-Theorem

2. Elementare Kombinatorik
   - 2.1. Fakultät und Binomialkoeffizient
   - 2.2. Das Fundamentalprinzip der Kombinatorik
   - 2.3. Permutationen
   - 2.4. Kombinationen
   - 2.5. Variationen

3. Zufallsvariablen
   - 3.1. Die Verteilungsfunktion
   - 3.2. Diskrete Zufallsvariablen
   - 3.3. Stetige Zufallsvariablen
   - 3.4. Erwartungswerte von Zufallsvariablen
   - 3.5. Varianz
   - 3.6. Standardisieren

4. Stochastische Modelle und spezielle Verteilungen
   - 4.1. Gleichförmige Verteilung (diskret)
   - 4.2. Bernoulli-Verteilung (diskret)
   - 4.3. Binomialverteilung (diskret)
   - 4.4. Poisson-Verteilung (diskret)
   - 4.5. Rechteckverteilung (stetig)
   - 4.6. Exponentialverteilung (stetig)
   - 4.7. Normalverteilung (stetig)

5. Mehrdimensionale Zufallsvariablen
   - 5.1. Gemeinsame Verteilung und Randverteilungen
   - 5.2. Bedingte Verteilungen und stochastische Unabhängigkeit
   - 5.3. Kovarianz und Korrelationskoeffizient
   - 5.4. Summe von zwei oder mehreren Zufallsvariablen

6. Der zentrale Grenzwertsatz

Teil II: Statistik

7. Beschreibende/Deskriptive Statistik
   - 7.1. Häufigkeitsverteilung, Histogramm und Verteilungsfunktion
   - 7.2. Messzahlen zur Beschreibung statistischer Verteilungen
   - 7.3. Boxplot
   - 7.4. Quantile-Quantile Plot
   - 7.5. Streudiagramm

8. Schätzung unbekannter Parameter: Punktschätzung
   - 8.1. Intuitiv heuristische Ansätze für Schätzfunktionen
   - 8.2. Eigenschaften von Punktschätzungen
   - 8.3. Methoden zur Konstruktion von Schätzfunktionen

9. Intervallschätzungen-Konfidenzintervalle
   - 9.1. Konzept des Konfidenzintervalls
   - 9.2. Ableitung von Konfidenzintervallen (bei grossen Stichproben)
   - 9.3 Zusammenhang mit Hypothesentests

10. Hypothesentests
   - 10.1 Arten von Hypothesen
   - 10.2 Kritischer Bereich und Teststatistik
   - 10.3 Gütefunktion und Arten von Fehlern
   - 10.4 Der p-Wert

Teil III: Aufgaben

Universität St.Gallen

Teil III:

Aufgaben

Aufgaben 1.1.1

1) Eine Person wird nach ihrem Geburtstag befragt:

2) K Personen werden nach ihrem Geburtstag befragt:

3) Position eines Zeigers im Einheitskreis:

4) Ist $S = \{0 \text{ Sechsen}, 2 \text{ Sechsen}\}$ ein Ereignisraum für den Doppelwürfel?

Aufgaben 1.2.1

1) Wurf eines regelmässigen Würfels:

$$\left.\begin{array}{l}\text{A: "Augenzahl kleiner als 4"}\\ \text{B: "Augenzahl ungerade"}\end{array}\right\} \rightarrow \begin{array}{l} A \cup B = ?\\ A \cap B = ?\end{array}$$

2)
$$\left.\begin{array}{l} A = \{(x, y) | ax + by + c = 0\}\\ B = \{(x, y) | ax + by + d = 0\}\end{array}\right. \rightarrow A ⍰ B = ?$$

3) Beweise die *De Morgan'sche Gesetze:*

$$\overline{A \cup B} = \overline{A} \cap \overline{B} \text{ und } \overline{A \cap B} = \overline{A} \cup \overline{B}$$

Aufgaben 1.2.1 *(Fortsetzung):*

4) Werfen von zwei Würfeln:

$$S = \{(1,1),(1,2),\ldots,(1,6),(2,1),(2,2),\ldots,(6,6)\}$$

und

A: „Mindestens ein Würfel zeigt eine Sechs“

B: „Die Augenzahl beider Würfel ist gleich“

C: „Beide Würfel zeigen ungerade Zahlen“

Aufgaben 1.2.1 *(Fortsetzung):*

Als Teilmenge von S:

$A = ?$

$B = ?$

$C = ?$

$\bar{A} = ?$

$\bar{C} = ?$

Aufgaben 1.2.1 *(Fortsetzung):*

$$B \cap C \quad = ?$$

$$B \setminus C \quad = ?$$

$$\overline{A \cup C} \quad = ?$$

$$\overline{A} \cap \overline{B} \cap \overline{C} = ?$$

Aufgabe 1.3.1:

Welcher Wahrscheinlichkeitsbegriff?

(1) Die Wahrscheinlichkeit, dass man beim Lotto drei oder mehr richtige Zahlen tippt, beträgt knapp 2%.

(2) Die Wahrscheinlichkeit, dass es im nächsten Juni in Frankfurt schneien wird, ist geringer als 5%.

(3) Der Bewerber X wird bei der Ausschreibung der Stelle Y mit einer Wahrscheinlichkeit von 80% zum Bewerbungsgespräch eingeladen.

Aufgabe 1.5.1:

“Werfen von zwei Würfeln”
(Siehe Aufgabe 1.2.1 (4))

A: “Mindestens ein Würfel zeigt eine Sechs”
B: “Die Augenzahl beider Würfel ist gleich”
C: “Beide Würfel zeigen ungerade Zahlen”

Berechnen Sie die Wahrscheinlichkeit der verschiedenen Ereignisse.

Aufgabe 1.5.2:

Krankmeldungen:

Wahrscheinlichkeiten für die Krankmeldungen von drei Mitarbeitern X, Y und Z (statistisch):

E_i	{-}	{X}	{Y}	{Z}	{XY}	{XZ}	{YZ}	{XYZ}	
$P(E_i)$	0.751	0.1	0.063	0.061	0.011	0.008	0.005	0.001	Σ=1

Berechne:

→ P[X krank] = ?

→ P[Y krank] = ? →

P[X und Y krank] = ?

→ P[X oder Y krank] = ?

Aufgabe 1.6.1:

Berechnen Sie die Wahrscheinlichkeit der folgenden Ereignisse:

a) A: “Bei vier Würfen mit einem Würfel tritt mindestens eine Sechs auf.“

b) B: “Bei 24 Würfen mit zwei Würfeln tritt mindestens eine Zwölf auf.“

Aufgabe 1.6.2:

$$U = \left\{ (a,b) \mid a \in [0,3],\ \ b < 1 - \frac{a}{6} \right\}$$

$$P(U) = ?$$

Aufgabe 1.7.1:
Krankenmeldungen (siehe Aufgabe1.5.2)

→ Sind die Krankenmeldungen der Mitarbeiter X, Y und Z paarweise stochastisch unabhängig?

I) P(X und Y krank) = ?

II) P(X und Z krank) = ?

III) P(Y und Z krank) = ?

=> Wie verändert eine Krankmeldung von X die Wahrscheinlichkeit einer Krankmeldung von Y?

P(Y krank | X krank) = ?

Aufgabe 1.7.2:

In einem Bürohaus befinden sich zwei technisch und funktional gleiche, unabhängig voneinander operierende Aufzüge A und B.
Die Wahrscheinlichkeit, dass sich Aufzug A bzw. B zu einem bestimmten Zeitpunkt im Erdgeschoss befindet, beträgt jeweils 0.2.

Aufgabe 1.7.2 *(Fortsetzung):*

→ Wie gross ist die Wahrscheinlichkeit, dass ein zu einem zufälligen Zeitpunkt eintreffender Besucher....

I) beide Aufzüge im Erdgeschoss vorfindet?
II) mindestens einen Aufzug im Erdgeschoss vorfindet?
III) genau einen Aufzug im Erdgeschoss vorfindet?

→ Die Aufzüge haben jeweils eine Ausfall-Wahrscheinlichkeit von 5% (befinden sich nicht im Erdgeschoss). Wie gross ist die Wahrscheinlichkeit, dass...

I) Aufzug A ausgefallen ist, wenn er sich nicht im Erdgeschoss befindet?

Aufgabe 1.7.3

Die Übertragung einer Nachricht von A nach B erfolgt über 3 voneinander unabhängige Kanäle, die mit Wahrscheinlichkeit “p“ ausfallen können.

P[“Übertragung kommt zustande“] = ?

Aufgabe 1.8.1: *Ziehung der Zusatzzahl im Lotto*

Beim Zahlenlotto werden insgesamt sieben Kugeln, ohne Zurücklegen, aus einer Urne mit 49 durchnummerierten Kugeln gezogen, wobei die letzte Ziehung die Zusatzzahl bestimmt.

Wie gross ist die Wahrscheinlichkeit von
A: “Zusatzzahl 1 wird gezogen“?

Aufgabe 1.8.2:

Ein Fahrzeughersteller stattet seine Fahrzeuge mit Klimaanlagen aus, die er von drei verschiedenen Zulieferern A, B und C bezieht.

Zulieferer	Lieferanteil	Mängelquote
A	50%	5%
B	30%	9%
C	20%	24%

M: “Zufällig ausgewähltes Fahrzeug besitzt eine mangelhafte Klimaanlage“

→ $P(M) = ?$

→ M ist beobachtet: $P(A|M) = ? \;=\; P(A) = 0.5\ ?$

$P(B|M) = ?$

$P(C|M) = ?$

Aufgabe 1.9.1:

Zulieferer mit Qualitätsunterschieden
(Siehe Aufgabe 1.8.2, Folgeaufgabe)

$$P\left(A \middle| M\right) = ?$$

$$P\left(B \middle| M\right) = ?$$

$$P\left(C \middle| M\right) = ?$$

Aufgabe 1.9.2: Urne

Gegeben seien 2 Urnen U_1 und U_2.

U_1 enthalte 5 weisse und 7 rote Kugeln.

U_2 eine weisse und 5 rote Kugeln.

Aus einer zufällig ausgewählten Urne wird ebenfalls zufällig eine Kugel entnommen. Diese sei rot. Mit welcher Wahrscheinlichkeit wurde U_1 gewählt?

Aufgabe 3.0.1:

→ Anzahl der Kunden in einem Geschäft:

→ Brenndauer einer Glühbirne (Stunden):

Aufgabe 3.1.1: *Doppelwürfel*

$S = \{(i, j) \mid i , j \in \{1,...,6\}\}$

X = Summe der beiden Augenzahlen: $(i, j) \to i+j$

$W = \{2,...,12\}$

Verteilungsfunktion ?

Aufgabe 3.1.2:

Eine Maschine produziert defekte Elemente mit einer Wahrscheinlichkeit von 5%. Der Produktion werden zufällig 4 Elemente entnommen.

X = # defekter Elemente in der Stichprobe

W = $\{0, 1, 2, 3, 4\}$

Verteilungsfunktion ?

Aufgabe 3.2.1:

Eine Zufallsvariable besitze die Wahrscheinlichkeitsfunktion

$$
f_x(x) = \begin{cases} K & \text{für} \quad x = 0 \\ 2K & \text{für} \quad x = 1 \\ 3K & \text{für} \quad x = 2 \\ 5K & \text{für} \quad x = 3 \\ 0 & \text{sonst.} \end{cases}
$$

Aufgabe 3.2.1 *(Fortsetzung):*

→ Man bestimme K:

→ $P(1 < X \leq 3) = ?$

$P(X > 1) = ?$

$P(X = 1) = ?$

→ Welches ist der kleinste Wert von X, für den gilt

$$P(X \leq x) = F_x(x) \geq 0.5?$$

Aufgabe 3.2.2:

Umsatz bei unsicherer Auftragslage (Fortsetzung)
(Siehe Beispiel 3.0.5)

Die Geschäftsleitung interessiert sich für die Wahrscheinlichkeit, dass

1. der Umsatz im nächsten Jahr höchstens 30 Mio beträgt.
2. die absolute Abweichung vom Umsatzziel 36 Mio im nächsten Jahr höchstens 6 Mio beträgt.

Aufgabe 3.3.1:

Eine Zufallsvariable X besitzt die Dichtefunktion

$$f(x) = \begin{cases} 3x^2 & , \quad 0 \leq x \leq \frac{1}{2} \\ \frac{3x}{2} & , \quad \frac{1}{2} < x \leq c \\ 0 & , \quad \text{sonst.} \end{cases}$$

a) Ermitteln Sie die obere Grenze c, so dass f eine Dichtefunktion ist. Zeichnen Sie die Dichte.

b) Bestimmen Sie die Verteilungsfunktion von X.

Aufgaben 3.4.1:

1. Zufallsvariable X mit

$$f(x) = \begin{cases} 3x^2 & , \quad 0 \le x \le \dfrac{1}{2} \\ \dfrac{3}{2}x & , \quad \dfrac{1}{2} < x \le c \\ 0 & , \quad \text{sonst.} \end{cases}$$

$E[X] = ?$

2. Wartezeit an der S-Bahn-Station (Fortsetzung)

   (Siehe Beispiel 3.3.2)

   $E[X] = ?$

3. Zwei Spieler A und B würfeln abwechselnd um Geld.

   Der würfelnde Spieler erhält von seinem Gegenspieler

   - 3 Euro wenn er eine 1 oder 2 würfelt;
   - 6 Euro bei einer 6; und
   - 0 Euro wenn er eine 3, 4 oder eine 5 würfelt.

Berechnen Sie den Erwartungswert der Zufallsvariable

X = "Gewinn von A" wenn jeder Spieler einmal würfelt.

Aufgabe 3.5.1:

Eine Zufallsvariable X besitzt die Dichtefunktion

$$f(x) = \begin{cases} \frac{3}{8}x^2 , & 0 \leq x \leq c \\ 0 , & \text{sonst.} \end{cases}$$

Berechnen Sie E[X] und V(X).

Aufgabe 3.5.2:

Man bestimme den Erwartungswert und die Varianz der Zufallsvariablen $Y = 3X + 2$, wobei X die Wahrscheinlichkeitsmassenfunktion

X	1	2	5
$f(x)$	0.2	0.3	0.5

besitzt.

Aufgabe 3.5.3:

Eine Rohrleitung besteht aus 20 Segmenten. Da die Ausflussmenge kleiner ist als die Zuflussmenge, muss irgendwo ein Leck bestehen. Wir nehmen an, dass es genau ein Leck gibt und dass es mit Wahrscheinlichkeit 1/20 in einem bestimmten Segment liegt. Wir möchten das defekte Segment mit möglichst wenigen Inspektionen (d.h. Messen der Durchflussmenge an einer Segmentgrenze) ausfindig machen.

Aufgabe 3.5.3 *(Fortsetzung):*

a) Bestimmen Sie die Verteilung der Anzahl Inspektionen X, wenn man sukzessive jede Segmentgrenze inspiziert. Berechnen Sie ferner E(X) und σ_x^2.

b) Bessere Strategie? (günstigere)

Aufgabe 4.3.1:

Qualitätskontrolle

Bei der Produktion von hochwertigen Trinkgläsern beträgt die Ausschussquote 20%. Im Rahmen einer Qualitätskontrolle werden nach einem Zufallsprinzip vier Gläser zur Prüfung entnommen (mit Zurücklegen).

X: # der fehlerhaften Gläser in der Stichprobe
Y: # der einwandfreien Gläser in der Stichprobe

Berechnen Sie die Wahrscheinlichkeit, dass

(1) in der Stichprobe genau ein Glas defekt ist;

(2) in der Stichprobe mindestens zwei Gläser defekt sind;

(3) in der Stichprobe genau ein Glas einwandfrei ist.

(4) E[X], E[Y], V(X), V(Y)?

Aufgabe 4.4.1:
Kunden an einem Bankschalter

An einen Bankschalter kommen zu unvorhersehbaren Zeitpunkten *vormittags* (8-12 Uhr) im Durchschnitt 12 Kunden pro Stunde und *nachmittags* (14-16 Uhr) im Durchschnitt 10 Kunden pro Stunde (Kunden kommen unabhängig voneinander gleichmässig über den Vor- bzw. Nachmittag verteilt).

Berechnen Sie die Wahrscheinlichkeit,

1) dass an einem Tag zwischen 09.00h und 09.15h kein Kunde kommt;
2) dass an einem Tag zwischen 15.00h und 15.15h kein Kunde kommt;
3) dass an einem Tag zwischen 15.30h und 16.00h mehr als 6 Kunden kommen.

Aufgabe 4.6.1:

Kunden an einem Bankschalter (Siehe Aufgabe 4.4.1, Fortsetzung)

→ vormittags (8-12): 12 Kunden pro Stunde

→ nachmittags (14-16): 10 Kunden pro Stunde

Wie gross ist an einem beliebigen Zeitpunkt am Vormittag bzw. Nachmittag die Wahrscheinlichkeit, dass in den nächsten fünf Minuten ein Kunde an den Schalter kommt?

→ X_{vor} = Zeit bis zur Ankunft des nächsten Kunden (vormittags) ~ $f_{Ex}(x; 12)$

→ X_{nach} = Zeit bis zur Ankunft des nächsten Kunden (nachmittags) ~ $f_{Ex}(x; 10)$

Aufgabe 4.7.1:
Arbeiten mit beliebiger Normalverteilung N(μ, σ²)

Sei $X \sim f_N(x\,;\,2{,}16)$

$$\rightarrow \quad \mathrm{P}[X \le 0] = ?$$

$$\rightarrow \quad \mathrm{P}\big[|X| \le 2\big] = ?$$

Nun sei $X \sim f_N(x\,;\,5,\,100)$. Finden Sie q, so dass:

$$\rightarrow \quad \mathrm{P}\big[X \le \mathrm{q}\big] = 0.25$$

$$\rightarrow \quad \mathrm{P}\big[X \le \mathrm{q}\big] = 0.75$$

Universität St.Gallen

Aufgabe 5.1.1:

Aus einer Schachtel mit 2 weissen, 3 schwarzen und 1 blauen Kugel werden 2 Kugeln *mit* Zurücklegen gezogen.

X = Anzahl weisse Kugeln

Y = Anzahl blaue Kugeln

Finden Sie die gemeinsame (bivariate) Verteilung und die Randverteilungen.

Aufgabe 5.4.1:

$\{X_1, X_2, ...\}$ sei eine Zufallsstichprobe (iid) mit $E(X_i) = \mu$

und $V(X_i) = \sigma^2$. Sei $Z = \frac{1}{22} \cdot \sum_{i=1}^{2} \left(4X_{2i} + 3X_i + 2\right)$

Berechnen Sie E[Z] und V(Z).

Aufgabe 5.4.2:

X und Y seien zwei unabhängige, poissonverteilte Zufallsvariablen. Es gelte $V(X) + V(Y) = 5$.

Bestimmen Sie $P(X+Y \leq 2)$.

Hinweis:

Die Poisson-Verteilung besitzt die reproduktive Eigenschaft:

$$\left.\begin{array}{lll} X & : & f_{Po} \\ Y & : & f_{Po} \\ X,Y & & \text{unabhängig} \end{array}\right\} \Rightarrow \left(X+Y\right) \quad : \quad f_{Po}$$

Universität St.Gallen

Aufgabe 6.1:

Die Wahrscheinlichkeit, dass es an einem Tag im Juni regnet, beträgt in einem mediterranen Urlaubsort 0.08, unabhängig davon welche Wetterlage an den übrigen Juni-Tagen dort herrscht.

a) Wie ist die Anzahl der Regentage in einer Juni-Woche (X_7), bzw. im gesamten Monat Juni (X_{30}) verteilt?

b) Bestimmen Sie den Erwartungswert und die Varianz von X_7.

c) Wie gross ist die Wahrscheinlichkeit, dass es ...
   - in einer Woche nicht regnet?
   - in einer Woche an mindestens drei Tagen regnet?
   - im ganzen Juni höchstens zwei Regentage gibt?

Aufgabe 6.1 *(Fortsetzung):*

d) Am gleichen Ort ist die Sonnenscheindauer an einem Juni-Tag normalverteilt mit $\mu = 10$ [Stunden] und $\sigma^2 = 10.8$ [Stunden^2]. Wie ist die Gesamt-Sonnenscheindauer im Juni (Y_{30}), bzw. die durchschnittliche tägliche Sonnenscheindauer im Juni ($\overline{Y}_{30}$) verteilt?

e) Wie gross ist die Wahrscheinlichkeit, dass die Sonne im Juni eines Jahres durchschnittlich mehr als 11 Stunden täglich scheint?

Aufgabe 6.2:

100 Zufallsziffern von 1 bis 5 werden gewählt und addiert. Wie gross ist die Wahrscheinlichkeit, dass die Summe

a) höchstens den Wert 250 annimmt?

b) zwischen den Werten 275 und 305 (inkl. Grenzen) liegt?

Aufgabe 6.3:

Ein Würfel wird 300 Mal geworfen. Sei X die Anzahl geworfener 3er.

Bestimme $P(50 < X \leq 53)$ und $P(X < 40)$.

Aufgabe 7.4.1:

Gegeben sind die Ergebnisse einer Prüfungsklausur (23 Studenten)

6.2; 4.82; 2.96; 6.18; 6.52; 7.9; 9.62, 6.22; 0.42;

9.06; 11.7; 6.54; 3.14; 4.74; 2.66; 7.04; 7.78; 11.8;

9.44; 20.76; 2.9; 8.42; 8.02

Boxplot?   QQ-Plot?

Aufgabe 8.2.1 :

Schätzung für λ im Falle einer Poissonverteilung.

Betrachten wir:

$$T_1 = \overline{X}_n \quad \text{und} \quad T_2 = \frac{n}{n-1} S^2$$

für den Parameter *λ* einer Poissonverteilung.

i) Sind T_1 und T_2 erwartungstreu für *λ* ?

ii) Ist T_1 konsistent für *λ* ?

iii) Ist T_1 effizient gegenüber T_2?

Aufgabe 8.2.2 :

Seien zwei Schätzfunktionen

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i \quad \text{und} \quad \overline{X}^{|}=\frac{1}{n+1}\left[2X_1+\sum_{i=2}^{n}X_i\right]$$

für E[X]=μ gegeben.

Zeigen Sie

1. Beide sind erwartungstreu
2. Berechnen Sie die Varianz von beiden
3. Welche Schätzfunktion ist effizient (zwischen diesen beiden)?

Aufgabe 8.3.1:

n identische Münzen werden je solange geworfen, bis erstmals Kopf erscheint.

Finde ML-Schätzer für $p = P[\text{"Kopf"}]$ aufgrund einer Stichprobe $X_1, \ldots, X_n$, wobei X_i die Anzahl Würfe vor dem ersten Erfolg für die Münze *i* bezeichnet.

Aufgabe 9.2.1:

Ein Sektor besteht aus N=12.100 Einzelunternehmen. Wir betrachten eine Zufallsstichprobe von Größe n=225. Die Zielgröße ist P = „Jahresgewinn" (in Schweizer Franken). Deskriptive Statistiken für die Ergebnisse im Jahr 2006 sind:

$$\bar{p}_{225} = 600,000.-; \quad s_P = 90,000.-$$

Bilden Sie:

1. ein Konfidenzintervall für den durchschnittlichen Jahresgewinn mit Signifikanzniveau α=4,55%;
2. ein Konfidenzintervall für den gesamten Jahresgewinn des Sektors mit Signifikanzniveau α=4,55%.

Aufgabe 9.2.2:

Im Juni 1986 veröffentlichte *Consumer Reports* einige Daten über den Kaloriengehalt von Rinder-Hotdogs. Hier sind die Kalorienzahlen für 20 verschiedene Hotdog-Marken (in kcal):

186, 181, 176, 149,184, 190, 158, 139, 175, 148,
152, 111, 141, 153,190, 157, 131, 149, 135, 132.

Nehmen Sie an, dass diese Zahlen die beobachteten Werte aus einer Zufallsstichprobe von zwanzig unabhängigen normalverteilten Zufallsvariablen mit unbekanntem Mittelwert μ und unbekannter Standardabweichung σ sind.

Finden Sie ein 90% Konfidenzintervall für den Erwartungswert der Kalorien μ.

Aufgabe 10.4.1:

Ein Psychologe behauptet, dass der durchschnittliche IQ einer Population 100 beträgt. Eine Zufallsstichprobe von 30 Individuen aus dieser Population hat einen durchschnittlichen IQ von 102. Nehmen Sie an, dass der IQ der Population normalverteilt ist und die Standardabweichung 15 beträgt.

Testen Sie die Behauptung des Psychologen mit Signifikanzniveau von 1%.

1. Formulieren Sie die Null- und Alternativhypothese.
2. Berechnen Sie die Teststatistik.
3. Bestimmen Sie den p-Wert.
4. Treffen Sie eine Entscheidung bezüglich der Nullhypothese.

Aufgabe 10.4.2:

Ein Autohersteller behauptet, dass sein neues Modell im Durchschnitt 30 Meilen pro Gallone (mpg) auf der Autobahn erreicht. Eine Zufallsstichprobe von 25 Autos zeigte einen Mittelwert von 28 mpg. Nehmen Sie an, dass die Population normalverteilt ist und die Standardabweichung 4 mpg beträgt.

Konstruieren Sie einen einseitigen Test für die Behauptung des Herstellers mit Signifikanzniveau von 10%.

1. Formulieren Sie die Null- und Alternativhypothese.
2. Berechnen Sie die Teststatistik.
3. Bestimmen Sie den p-Wert.
4. Treffen Sie eine Entscheidung bezüglich der Nullhypothese.