Script
Wahrscheinlichkeitstheorie

Inhaltsverzeichnis

1 Wahrscheinlichkeitstheorie

1.1 Zufallsexperimente

1.1.1 Definitionen und Beispiele

Wenn wir eine Münze werfen, so geben wir dem Ergebnis Kopf die Wahrscheinl. 0,5 . Warum? Wir gehen davon aus, dass es zwei mögliche Ergebnisse gibt, Kopf und Zahl.

Mathematisch:

$\Omega = \{\text{Kopf, Zahl}\}$ Ergebnismenge [enthält alle mgl Ergebnisse]

Aus Symmetriegründen geben wir Kopf und Zahl die gleiche Wahrscheinlichkeit. Da der Münzwurf irgendein Ergebnis produziert, und Kopf und Zahl sich gegenseitig ausschließen, muss die Wkt. von Kopf und Zahl je 0,5 sein.
(1,0 $\implies$ Irgendein Ergebnis [= 0,5 Kopf + 0,5 Zahl])

Bemerkung:

Es ist üblich Wkt.en als Bruch oder Dezimalzahl zu schreiben, also hier Wkt. von Kopf ist $\frac{1}{2}$.

Wenn wir das Münzwurfexperiment n-mal durchführen und mit n_K (bzw. n_Z) die Anzahl von Ergebnis Kopf beschrieben, dann ist $\frac{n_K}{n}$ die relative Häufigkeit von Kopf.

Historisch:

Buffon: $n = 4040,$ $n_K = 2048$

Pearson: $n = 24000$, $n_K = 12012$

Aus $\frac{n_K}{n} = 0,5$ in beiden Versuchen schließen wir, dass unser Modell für einen fairen Münzwurf ein gutes Modell ist, da es gute Vorhersagen für die Wahrscheinlichkeiten macht.

Def. 1 :

Eine Menge S heißt abzählbar unendlich, falls es eine bijektion von den natürlichen Zahlen ($\mathbb{N}$) nach S gibt. Eine Menge S heißt abzählbar, falls sie endlich oder abzählbar unendlich ist.

Def. 2 :

Ein diskreter Wkt.-Raum bzw. ein diskretes Wkt.-Modell ist ein Paar (Ω, p) wobei Ω eine abzählbare Ergebnismenge ist und

$$p : \Omega \implies [0,1] \quad \text{mit} \quad \sum_{x \in \Omega} p(x) = 1$$

die Wkt.-funktion auf Ω, d.h. eine Funktion, die den verschiedenen Elementarereignissen $x \in \Omega$ ihre Wkt. zuordnet.

Beispiel :

a) fairer Münzwurf

$$\Omega = \{\text{Kopf, Zahl}\}, \qquad p(Kopf) = p(Zahl) = \frac{1}{2}$$

b) unfairer Münzwurf
Sei $a \in [0,1] \backslash \{\frac{1}{2}\}$ ein Parameter

$$\Omega = \{\text{Kopf, Zahl}\}, \qquad p(Kopf) = a, \qquad p(Zahl) = 1 - a$$

Die Wahl von Ω und Wkt.-funktion p sind Modellierungsannahmen, die gut oder schlecht sein können.
Das Schätzen von Parametern (etwa der Parameter a im obigen Beispiel) anhand von Daten ist Teil der (parametrischen) Statistik.

Def. 3 :

Sei Ω eine endliche Ergebnismenge. Wir nennen $p : \Omega \implies [0,1]$ mit $p(x) = \frac{1}{|\Omega|}$ für alle $x \in \Omega$ die Gleichverteilung auf Ω.

Die Gleichverteilung ist eine gute Wahl, wenn:

- wegen z.B. Symmetrie kein Elementarereignis wahrscheinlicher als ein anderes ist.
- man aufgrund fehlender Informationen kein Argument hat einem Elementarereignis eine
höhere/niedrigere Wkt. zuzuordnen.

Def. 4 :

Für eine Menge S ist $2^S := \{S' \subset S\}$ die Potenzmenge von S, die Menge aller Teilmengen. Teilmengen von (Abzählbaren) Ergebnismengen nennen wir

Ereignisse.

Def. 5 :

Sei (Ω,p) ein diskretes Wkt.modell.Für ein Ereignis $A \subset \Omega$ definieren wir die Wkt. von A als

$$\mathbb{P}(A) = \sum_{x \in A} p(x)$$

Diese Funktion $\mathbb{P} : 2^{\Omega} \implies [0,1]$ heißt Wkts.verteilung oder Wkts.maß auf Ω.

Beispiel :

Das Ereignis, dass ein Würfel ein Ergebnis kleiner gleich 3 hat, wird durch die Menge $\{1,2,3\}$ beschrieben.

$$\mathbb{P}(\{1,2,3\}) = p(1) + p(2) + p(3) = 0.5$$

ist die Wkt. davon.

Bemerkung:

Wir haben $p(x) = \mathbb{P}(\{x\})$, daher werden (diskrete) Wkts. modelle auch über das Paar $(\Omega, \mathbb{P})$ bestimmt.

Proposition 6 :

Sei Ω endlich und p die gleichverteilt auf Ω. Dann ist die Wkt. von einem Ereignis $A \subset \Omega$ genau:

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

.

Bemerkung:

Eine Strategie zur Modellierung ist es, den Ergebnisraum soweit zu verfeinern, bis man eine Gleichverteilung annehmen kann. Danach können Wkt.en durch Zählen bestimmt werden.

Beispiel :

Sei ein Würfel gegeben, der statt einer 6 zwei 1en hat. Was ist die Wkt. von einem Ergebnis ≤ 3?
Modelle:

a) $\Omega = \{1,2,3,4,5\}, \quad p(x) = \frac{1}{5}$ (Gleichvert.)

$$\mathbb{P}(\{1,2,3\}) = \frac{3}{5}$$

$\Longrightarrow$ schlechtes Modell, 1en sollten doppelt so wahrscheinlich wie andere Zahlen sein

b) $\Omega = \{1,2,3,4,5\}, \quad p(1) = \frac{2}{6} = \frac{1}{3}, \quad p(x) = \frac{1}{6}$ für $x \geq 1$

$$\mathbb{P}(\{1,2,3\}) = \frac{4}{6} = \frac{2}{3}$$

c) $\Omega = \{1_a, 1_b, 2, 3, 4, 5\}, \quad p(x) = \frac{1}{6}$, (Gleichvert. wg. Symmetrie des Würfels)

$$\mathbb{P}(\{1_a, 1_b, 2, 3\}) = \frac{4}{6} = \frac{2}{3}$$

Modelle b) und c) sind verschieden, aber liefern beide das Ergebnis $\frac{2}{3}$. Beides sind gute Modelle.
Modell a) liefert $\frac{3}{5}$. Das Ergebnis ist "richtig" in dem Sinne, dass die Wkt. bzgl. Modell a) korrekt berechnet wurde.
Modell a) ist aber ein schlechtes Modell, da es fälschlicher Weise von einer Symmetrie der Ergebnisse 1,2,3,4,5 ausgeht.

1.1.2 Zählen und Kombinatorik

In der Kombinatorik unterscheidet man zwischen geordneten und ungeordneten Mengen

Beispiel :

Teilmengen:

- Geordnete Teilmengen von $\{1, 2, 3\}$ der Größe 2:

$$(1, 2), (1, 3), (2, 1), (2, 3), (3, 1), (3, 2)$$

- ungeordneten Teilmengen von $\{1, 2, 3\}$ der Größe 2:

$$\{1, 2\}, \{1, 3\}, \{2, 3\}$$

Satz 7 (Kombinatorisches Zählen):

Sei S endlich, $|S| = n$:

a) Dei Anzahl der Folgen der Länge k mit den n Symbolen aus S ist n^k

b) Die Anzahl der geordneten Teilmengen von S der Länge $k \leq n$ ist

$$n \cdot (n-1)(n-2)(n-3)...(n-k+1) = \frac{n!}{(n-k)!}$$

Insbesondere hat S genau $n!$ viele verschiedene Reihenfolgen.

c) Die Anzahl der ungeordneten Teilmengen von S der Länge $K \leq n$ ist

$$\binom{n}{k} = \frac{n \cdot (n-1)...(n-k+1)}{k!} = \frac{n!}{(n-k)! \cdot k!}$$

Beweis :

a) Für das erste Element hat man n Möglichkeiten, für das zweite ebenfalls, als n^2 viele für die ersten beiden. Fortsetzten ergibt n^k

b) Für das erste Element hat man n Möglichkeiten, für das zweite noch $(n-1)$, das dritte nur noch $(n-2)$, usw., da das gleiche Element aus S nicht mehrmals gewählt werden kann.
Damit ergibt sich für k Elemente

$$n(n-1)(n-2)(n-3)...(n-k+1)$$

viele Möglichkeiten.

c) Sei $A \subset S$ (ungeord. falls nicht spezifiziert), $|A| = k$
Nach b) gibt es $k!$ viele Möglichkeiten A zu sortieren. Folglich taucht A in all den verschiedenen Reihenfolgen $k!$ mal in der Menge der geordneten Teilmenge der Länge k auf. Da die für jedes solche A gilt muss die Menge der geordneten Teilmengen $k!$ mal größer sein als die Menge der ungeordneten Teilmengen. Mit b) ist also die Anzahl der ungeordneten Teilmengen der Länge k genau

$$\frac{n!}{(n-k)!}/k! = \frac{n!}{(n-k)! \cdot k!}$$

$\square$

Beispiel (Urnenmodelle):

Sei eine Urne mit 8 nummerierten Kugeln gegeben. Ansonsten sind die Kugeln identisch. Man zieht blind eine Kugel

a) **Ziehen mit Zurücklegen:** Ziehen Kugel, notieren die Zahl, legen die Kugel zurück. Dies wiederholen wir 2 mal.
Wkts.modell:

$$S = \{1, 2, 3, 4, 5, 6, 7, 8\}$$

$$\Omega = S^3 = \{(x_1, x_2, x_3) \quad : \quad x_i \in S, \quad i = 1, 2, 3\} (\text{nicht geord. TM})$$

$$p(x) = \frac{1}{|\Omega|} = \frac{1}{8^3} = \frac{1}{512} \qquad (\text{Gleichverteilung})$$

Was ist die Wkt., dass die Summe 4 ist?

Das Ergebnis :

$$A := \{(x_1, x_2, x_3) \in \Omega \quad : \quad x_1 + x_2 + x_3 = 4\} = \{(1,1,2),(1,2,1),(2,1,1)\}$$

beschreibt die Möglichkeiten mit den 3 Zahlen in der Summe 4 zu erhalten. Also ist die Wkt. (in diesm Modell).

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{3}{512} \approx 0,006$$

b) **Ziehen ohne Zurücklegen:** Ziehe Kugel, notiere Zahl, legen Kugel beiseite, 2 mal wiederholen

Fall 1) Mit Beachung der Reihenfolge
$\Omega = \{(x_1, x_2, x_3) \qquad \text{geordnete Teilmenge von S der Länge 3}\}$

$$p(x) = \frac{1}{|\Omega|} = \frac{1}{8 \cdot 7 \cdot 6} = \frac{1}{336} \quad \text{Gleichverteilung}$$

Fall 2) Ohne Betrachtung der Reihenfolge
$\Omega = \{\{x_1, x_2, x_3\} \in 2^S\}$

$$p(x) = \frac{1}{|\Omega|} = \frac{1}{\binom{8}{3}} = \frac{1}{56}$$

1.1.3 Eigenschaften von Wkts verteilungen/-maßen

In diesem Abschnitt seien $(\Omega, \mathbb{P})$ Wkts.modell gegeben
[Erinnerung: Wir können ein Wkts.modell (oder Wkt. raum) durch die Wkts.verteilung $\mathbb{P}$ oder Wkts.funktion p bestimmen]

Def. 8 :

Ereignisse $A_1, A_2, ...$ heißen paarweise disjunkt, falls $A_i \cap A_j = \varnothing$ ist für alle $i \neq j$

Proposition 9 (Eigenschaften von Wkts.verteilungen):

a)
$$\mathbb{P}(\varnothing) = 0, \qquad \mathbb{P}(\Omega) = 1$$

b) Seien $A_1, A_2, ..., A_n$ paarweise disjunkt. Dann ist
$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i)$$

c) Seien $A_1, A_2, ...$ paarweise disjunkt. Dann ist
$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

d)
$$\mathbb{P}\left(A^C\right) = 1 - \mathbb{P}(A)$$

e) Wenn $A \subset B$, dann ist $\mathbb{P}(A) \leq \mathbb{P}(B)$. Genauer:
$$\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \backslash A)$$

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

Beweis :

a)

$$\mathbb{P}(\Omega) = \sum_{\omega\in\Omega} p(\omega) = 1, \qquad \text{nach Def. von p}$$

$$\mathbb{P}(\varnothing) = \sum_{\omega\in\varnothing} p(\omega) = 0, \qquad \text{da die Summe leer ist}$$

b)

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{\omega\in\bigcup_{i=1}^{n} A_i} p(\omega) = \sum_{\omega\in A_1} p(\omega) + \sum_{\omega\in A_2} p(\omega) + ... + \sum_{\omega\in A_n} p(\omega)$$

c)

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{\omega\in\bigcup_{i=1}^{\infty} A_i} p(\omega) = \sum_{\omega\in A_1} p(\omega) + \sum_{\omega\in A_2} p(\omega) + ...$$

d) A und A^C sind disjunkt. Daher

$$1 \overset{a)}{=} \mathbb{P}(\Omega) = \mathbb{P}\left(A \cup A^C\right) \overset{b)}{=} \mathbb{P}(A) + \mathbb{P}\left(A^C\right)$$

e) Da nach definition gilt: $\quad B = A \cup B \backslash A$

$$\mathbb{P}(B) = \sum_{\omega\in A} p(\omega) + \sum_{\omega\in B\backslash A} p(\omega) = \mathbb{P}(A) + \mathbb{P}(B\backslash A)$$

f)

$$\begin{aligned}\mathbb{P}(A\cup B) &= \sum_{\omega\in A\cup B} p(\omega) \\ &= \sum_{\omega\in A} p(\omega) + \sum_{\omega\in B} p(\omega) - \sum_{\omega\in A\cap B} p(\omega) \\ &= \mathbb{P}(A\cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A\cap B)\end{aligned}$$

$\square$

Proposition 10 :

Wir können verschiedene Ereignisse durch Mengenoperationen verknüpfen:

a) A oder B tritt ein $\widehat{=}$ $A \cup B$

b) A und B tritt ein $\widehat{=}$ $A \cap B$

c) A tritt nicht ein $\widehat{=}$ $A^C = \Omega \backslash A$

Beweis :

a) A oder B tritt ein

$$\widehat{=} \{\omega \in \Omega : \omega \in A \text{ oder } \omega \in B\} = \{\omega \in \Omega : \omega \in A \cup B\} = A \cup B$$

b) A und B tritt ein

$$\widehat{=} \{\omega \in \Omega : \omega \in A \text{ und } \omega \in B\} = \{\omega \in \Omega : \omega \in A \cap B\} = A \cap B$$

c) A tritt nicht ein

$$\widehat{=} \{\omega \in \Omega : \omega \notin A\} = \Omega \backslash A = A^C$$

$\square$

1.1.4 Bedingte Wahrscheinlichkeit

Beispiel :

a) Wir werfen eine Münze zwei mal. Sei K_1 bzw K_2 das Ereignis, dass der erste bzw. zweite Wurf Kopf zeigt.
z.B.:

$$\begin{aligned} \Omega &= \{KK, KZ, ZK, ZZ\} \\ K_1 &= \{KK, KZ\} \\ K_2 &= \{KK, ZK\} \end{aligned}$$

Das Eintreten von K_1 sollte keinen Einfluss darauf haben, dass K_2 eintritt oder nicht eintritt. Wir glauben, dass K_1 und K_2 unabhängig (voneinander) sind.

b) Wir werfen einen Würfel. Sei A_3 das Ereignis, dass das Ergebnis 3 ist. Sei $A_{\leq 4}$ das Ereignis, dass das Ergebnis kleiner gleich 4 ist.
z.B.:

$$\begin{aligned} \Omega &= \{1, 2, 3, 4, 5, 6\} \\ A_3 &= \{3\} \\ A_{\leq 4} &= \{1, 2, 3, 4\} \end{aligned}$$

In diesem Fall sind die Ereignisse klar voneinander abhängig:

$A_{\leq 4}$

tritt ein	tritt nicht ein
3 eines von 4 mgl. Erg.	3 als Erg. unmgl.
Wkt. von A_3, gegeben dass $A_{\leq 4}$ eintritt, ist $\frac{1}{4}$?	Wkt. von A_3 gegeben, dass $A_{\leq 4}$ nicht eintritt ist 0?

Empirische Notation:

Nehmen wir an, wir haben zwei beobachtbare Größer, z.B. infiziert / nicht infiziert und Test positiv / negativ.
Uns interessiert die falsch-negativ Rate des Tests, d.h. die Wkt., dass der Test negativ ist, gegeben, dass die Person infiziert ist.
Ereignisse: $I \widehat{=}$ infiziert, $\quad N \widehat{=}$ Test negativ
Mit n Wiederholungen können wir zählen:

$n_I :=$ # infiziert
$n_{I\cap N} :=$ # infiziert und Test negativ

empirische falsch-negativ Rate $= \frac{n_{I\cap N}}{n_I}$

$$= \frac{\frac{n_{I\cap N}}{n}}{\frac{n_I}{n}} \approx \frac{\mathbb{P}\left(I \cap N\right)}{\mathbb{P}\left(I\right)}$$

Def. 11 (Bedingte Wahrscheinlichkeit):

Sei $\Omega, \mathbb{P}$ ein Wkts.modell und $A, B \subset \Omega$ Ereignisse mit $\mathbb{P}\left(B\right) > 0$.
Dann ist die bedingte Wkt. von A gegeben B definiert als:

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

Beispiel (Fortsetztung des Würfelspiels):

Die bedingte Wkt. von A_3 gegeben $A_{\leq 4}$ ist:

$$\mathbb{P}\left(A_3|A_{\leq 4}\right) = \frac{\mathbb{P}\left(A_3 \cap A_{\leq 4}\right)}{\mathbb{P}\left(A_{\leq 4}\right)} = \frac{\mathbb{P}\left(\{3\}\right)}{\mathbb{P}\left(\{1,2,3,4\}\right)} = \frac{\frac{1}{6}}{\frac{4}{6}} = \frac{1}{4}$$

Beispiel (Fortsetztung Münzwürfe):

Die bedingte Wkt. von K_2 gegeben K_1 ist:

$$\mathbb{P}\left(K_2|K_1\right) = \frac{\mathbb{P}\left(K_2 \cap K_1\right)}{\mathbb{P}\left(K_1\right)} = \frac{\mathbb{P}\left(\{KK\}\right)}{\mathbb{P}\left(\{KZ, KK\}\right)} = \frac{\frac{1}{4}}{\frac{2}{4}} = \frac{1}{2} = \mathbb{P}\left(K_2\right)$$

Sei nun M das Ereignis, dass mindestens eine Münze Kopf zeigt, also $M = \{KK, KZ, ZK\}$.
Wkt., dass der zweite Wurf Kopf zeigt, gegeben M?

$$\mathbb{P}\left(K_2|M\right) = \frac{\mathbb{P}\left(K_1 \cap M\right)}{\mathbb{P}\left(M\right)} = \frac{\mathbb{P}\left(\{KK, KZ\}\right)}{\mathbb{P}\left(\{KK, KZ, ZK\}\right)} = \frac{\frac{2}{4}}{\frac{3}{4}} = \frac{2}{3}$$

Beispiel (Geschwister):

Situation: Wir untersuchen Familien mit 2 Kindern
Annahmen:

- Junge/Mädchen sind gleich wahrscheinlich
- Die Geschlechter der Kinder sind unabhängig voneinander

Frage: Gegeben, dass eine Familie mindestens ein Mädchen hat, was ist die Wkt., dass das andere Kind auch ein Mädchen ist?
Wkts.modell:

$$\Omega := \{mm, mj, jm, jj\}$$

p Gleichverteilung (Mädchen/Junge wie Münzwurf)

Ereignisse:
$M_{\geq 1} = \{mm, mj, jm\}$
$M_{\geq 2} = \{mm\}$

$$\mathbb{P}\left(M_{\geq 2} | M_{\geq 1}\right) = \frac{\mathbb{P}\left(M_{\geq 2} \cap M_{\geq 1}\right)}{\mathbb{P}\left(M_{\geq 1}\right)} = \frac{\mathbb{P}\left(\{mm\}\right)}{\mathbb{P}\left(\{mm, mj, jm\}\right)} = \frac{1}{3}$$

Jetzt besuchen wir einen Familie mit mindestens einem Mädchen und klingeln an der Tür. Ein Mädchen macht auf. Was ist die Wkt., dass das andere Kind ein Mädchen ist?
Modell:

$$\Omega = \{m'm, mm', m'j, mj', j'm, jm', j'j, jj'\}$$

p Gleichverteilung (welches Kind die Tür öffnet ist unabh. von Rest und zufällig)
$M_{\geq 1} = \{m'm, mm', m'j, mj', j'm, jm'\}$
$M'_{\geq 1} = \{m'm, mm', m'j, jm'\}$

$$\mathbb{P}\left(M_{\geq 2} | M'_{\geq 1}\right) = \frac{\mathbb{P}\left(M_{\geq 2} \cap M'_{\geq 1}\right)}{\mathbb{P}\left(M'_{\geq 1}\right)} = \frac{\mathbb{P}\left(\{m'm, mm'\}\right)}{\mathbb{P}\left(\{m'm, mm', m'j, jm'\}\right)} = \frac{\frac{2}{8}}{\frac{4}{8}} = \frac{1}{2}$$

Def. 12 :

Eine abzählbare Folge von Ereignissen $B_1, B_2, \ldots$ ist eine Partition von Ω, falls für die $B_1, B_2, \ldots$ paarweise disjunktheit und $\bigcup_i B_i = \Omega$ gilt.

Bemerkung:

Die Folge $B_1, B_2, \ldots$ kann endlich oder abzählbar unendlich sein. Der einfachste Fall ist $B_1 = B, B_2 = B^C = \Omega \backslash B$

Satz 13 (Satz der totalen Wkt.):

Seien A,B Ereignisse mit $\mathbb{P}(B) > 0$. Dann ist

$$\mathbb{P}(A) = \mathbb{P}(A|B) \cdot \mathbb{P}(B) + \mathbb{P}\left(A|B^C\right) \cdot \mathbb{P}\left(B^C\right)$$

Allgemeiner, sei $B_1, B_2, \ldots$ eine Partition von Ω mit $\mathbb{P}(B_i) > 0 \quad \forall i$. Dann ist

$$\mathbb{P}(A) = \sum_i \mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)$$

Beweis :

Sei $A_i = A \cap B_i$, $\quad A_1, A_2, \ldots$ sind paarweise disjunkt, da die $B_1, B_2, \ldots$ eine Partition sind. Außerdem ist $\bigcup_i A_i = A$

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_i A_i\right) \overset{Prop.9}{=} \sum_i \mathbb{P}(A_i)$$

Nach der Def. der bedingten Wkt ist

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

$$\implies \mathbb{P}(A_i) = \mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)$$

□

Satz 14 (Satz von Bayes):

Seien A, B Ereignisse mit $\mathbb{P}(A), \mathbb{P}(B) > 0$.
Dann ist

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

Beweis :

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

□

Bemerkung:

Der Satz von Bayes ist sehr wichtig für viele Anwendungen und kann Ergebnisse liefern, die auf den ersten Blick der Intuition widersprechen.

1.1.5 Unabhängigkeit und Produktverteilung

Def. 15 (Unabhängigkeit):

Zwei Ereignisse A, B heißen unabhängig, falls

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

Bemerkung:

Gilt im allgemeinen **nicht** für alle Ereignisse!!!

Beispiel (Münzwürfe):

Modell:

$\Omega = \{KK, KZ, ZK, ZZ\}, \quad p(\omega) = \frac{1}{4}, \quad \omega \in \Omega, \quad K_1 = \{KK, KZ\}, \quad K_2 = \{ZK, KK\}$

$\mathbb{P}(K_1 \cap K_2) = \mathbb{P}(\{KK\}) = \frac{1}{4}$
$\mathbb{P}(K_1) \cdot \mathbb{P}(K_2) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

Bemerkung:

Falls $\mathbb{P}(B) > 0$, so sind A und B unabh. gdw. $\mathbb{P}(A|B) = \mathbb{P}(A)$

Def. 16 (Produktverteilung):

Seien $(\Omega_1, p_1), ..., (\Omega_n, p_n)$ Wkts.modelle. Sei $\bar{\Omega} = \Omega_1 * ... * \Omega_n = \{(\omega_1, ..., \omega_n) : \omega_i \in \Omega_i\}$
Dann heißt $\bar{p} : \bar{\Omega} \to [0, 1], \quad \bar{p}((\omega_1, ..., \omega_n)) = \prod_{i=1}^{n} p_i(\omega_i)$
Produktverteilungsfunktion und $(\bar{\Omega}, \bar{p})$ Produktraum oder Produktmodell.

Proposition 17 (Produktverteilung):

Sei $(\bar{\Omega}, \bar{p})$ das Produktmodell von $(\Omega_i, p_i), \quad i = 1, ..., n$.
Seien $\bar{A}_1, ..., \bar{A}_n$ Ereignisse in $\bar{\Omega}$, die jeweils nur von der i-ten Koordinate abhängen, d.h.

$$\bar{A}_i = \{(\omega_1, ..., \omega_n) \in \bar{\Omega} : \quad \omega_i \in A_i\}$$

für ein $A_i \subset \Omega_i$.
Dann sind $\bar{A}_1, ..., \bar{A}_n$ unabhängig.

Beispiel (n unfaire Münzwürfe):

Sei $\Omega_i = \{0,1\}$, $p_i(1) = \alpha$, $p_i(0) = 1-\alpha$ das Modell für einen einfachen Münzwurf mit Parameter α.
Dann beschreibt das Produktmodell $(\bar{\Omega}, \bar{p})$ von $(\Omega_1, p_1), ..., (\Omega_n, p_n)$
n identische Münzwürfe, die unabhängig voneinander mit Wkt. α 1 (Kopf) zeigen.
Sei $\mathbb{P}$ die Wkt.verteilung von $\bar{p}$ und B_k das Ereignis, dass genau k Münzen Kopf zeigen. Was ist $\mathbb{P}(B_k)$?

Lemma 18 :

$$\mathbb{P}(B_k) = \binom{n}{k}\alpha^k(1-\alpha)^{n-k}$$

Beweis :

Sei $\bar{\omega} = (\omega_1, ..., \omega_n) \in B_k$.

$$\bar{p}(\bar{\omega}) = \prod_{i=1}^{n} p_i(\omega_i) = \alpha^k(1-\alpha)^{n-k}$$

$\implies \mathbb{P}(B_k) = \sum_{\bar{\omega}\in B_k} \bar{p}(\bar{\omega}) = |B_k|\alpha^k(1-\alpha)^{n-k}$
Beh.: $|B_k| = \binom{n}{k}$ nach Satz 7.c). Zu jedem $\bar{\omega} \in B_k$ gehört genau eine Teilmenge $\{j \in \{1, ..., n\} : \omega_j = 1\} \subseteq \{1, ..., n\}$.
Diese Teilmenge hat genau k Elemente. Umgekehrt bildet jede solche Menge einen Vektor $\bar{\omega}$, bei dem genau diese Koordinate 1 sind.
Nach Prop. 7 c) gibt es genau $\binom{n}{k}$ ungeordnete Teilmengen von $\{1, ..., n\}$ mit k Elementen □

Anwendung:

- Wahlen/-prognosen
- Medizinische Tests im Labor
- Produktionsmaschinen

Def. 19 :

Sei $n \in \mathbb{N}$ und $\alpha \in [0,1]$ gegeben. Dann bilden $\Omega = \{1, ..., n\}$ und Wkts.funktion $p(k) = \binom{n}{k}\alpha^k(1-\alpha)^{n-k}$ die Binomialverteilung mit n Versuchen und Erfolgswkt. α.
Die Binomialverteilung zählt die Anzahl von "Erfolgen", wenn man n Versuche

hat und in jedem unabhängig Erfolgswkt. α.

Anwendung:

- Eine Maschine produziert am Tag 1000 Teile, davon sind normaler Weise 1% defekt.
  $\Longrightarrow$ Binomialverteilung mit $n = 1000, \quad \alpha = 0,01$.
  $\Longrightarrow$ benutzen, um Maschinendefekt/Fehlkallibrierung festzustellen.
  Betrachte zum Beispiel Ereignis, dass es mehr als 15 Defekte Teile gibt.

  $\mathbb{P}(A) \approx$ (Ergebnis aus Tabelle, Bin.vert.rechner im Internet, Matlab,...)

Satz 20 (Gesetz der großen Zahlen (erste Version)):

Sei $\Omega_n = \{0,1\}^n$ mit Produktverteilung $\mathbb{P}_n$ (bzw. Produktverteilungsfunktion $\bar{p}_n$) zu Parameter α wie oben.
Sei

$$A_{n,\epsilon} = \{\omega \in \Omega_n : |\alpha - \frac{1}{n}\sum_{i=1}^{n}\omega_i| \leq \epsilon\}$$

das Ereignis, dass die zufällige Frequenz der 1-Ergebnisse nun nicht mehr als ϵ von α abweicht.
Dann gilt

$$\lim_{n\to\infty} \mathbb{P}_n(A_{n,\epsilon}) = 1 \qquad \text{für alle } \epsilon > 0$$

1.2 Zufallsvariablen

1.2.1 Definitionen

Beispiel (n Würfelwürfe):

(Ω, p) Produktmodell von n einfachen Würfelwürfen ($\Omega_i = \{1, 2, ..., 6\}$, p_i gleichvert. $\implies \Omega = \{1, ..., 6\}^n, p(\omega) = \frac{1}{6^n} \quad \forall \omega \in \Omega$
Sei $X_i : \Omega \to \mathbb{R}, \quad X_i((\omega_1, ..., \omega_n)) = \omega_i$ das Ergebnis des i-ten Würfelwurfs.
Summe der Würfel:

$$S : \Omega \to \mathbb{R}, \qquad S = \sum_{i=1}^{n} X_i$$

Ergebnis, dass die ersten 3 Würfel eine 6 zeigen:

$$B = \{\omega \in \Omega : X_1(\omega) = X_2(\omega) = X_3(\omega) = 6\}.$$

Def. 1 (Zufallsvariablen):

Sei (Ω, p) ein diskretes Wkts.-modell/-raum. Eine Abbildung

$$X : \Omega \to \mathbb{R}$$

heißt Zufallsvariable.

Konventionen:

- Für Zufallsvariablen benutzten wir Großbuchstaben
- Für Ereignisse $\{\omega \in \Omega : X(\omega) = x\}, \quad \{\omega \in \Omega : X(\omega) \geq x\}, ...$ schreiben wir verkürzend:

$$\{X = x\}, \quad \{X \geq x\}, ...$$

- Für Wkten lassen wir oft Mengenklammern weg, also

$$\mathbb{P}\left(X = x\right), \quad \mathbb{P}\left(X \geq x\right), ...$$

Typische Zufallsvariablen:

- Die Größe, die uns letztlich interessiert (Summe der Würfel, die Regenmenge von Morgen, der Sieger der US-Präsidentschaftswahl)
- Bausteine, die mir als Zwischenschritte benötigen (Ergebnisse der einzelnen Würfel, Wolkendichte in Gewitterzellen einer Wettersimulation, Sieger in den einzelnen Bundesstaaten)

Rechenregeln

Seien X, Y Z.Var auf einem Wkts.raum (Ω, p). Dann sind auch Z.Var:

-
$$Z = X + Y, \qquad \text{d.h.} \quad Z(\omega) = X(\omega) + Y(\omega)$$
$$(Z = X \cdot Y, Z = X - Y, Z = min(X, Y), ...)$$

-
$$Z = g(X), \qquad \text{d.h.} \quad Z(\omega) = g(X(\omega)) \quad \text{für } g : \mathbb{R} \to \mathbb{R}$$
$$Z = g(X, Y), \qquad \text{d.h.} \quad Z(\omega) = g(X(\omega), Y(\omega)) \quad \text{für } g : \mathbb{R} \to \mathbb{R}$$

usw.

Lemma 2 :

Eine Z.Var. $X : \Omega \to \mathbb{R}$ auf einer abzählbaren Ergebnismenge Ω nimmt höchstens abzählbar viele verschiedene Werte an. Ihr Wertebereich

$$X(\Omega) := \{X(\omega) : \omega \in \Omega\} \quad \text{ist abzählbar.}$$

Def. 3 :

Die Wkts.funktion einer Z.Var. X auf $(\Omega, \mathbb{P})$ ist

$$p_X : X(\Omega) \to [0, 1], \quad p_X(x) = \mathbb{P}(X = x) . \qquad (= \mathbb{P}(\{\omega \in \Omega : X(\omega) = x\}))$$

Wir sagen eine Z.Var. ist gleichverteilt (binomialverteilt, ...) falls ihre Wkts.funktion p_X eine Gleichverteilung (Binomialverteilung, ...) ist.

Def. 4 :

Sei A ein Ereignis in Ω. Die Z.Var.

$$\mathbb{1}_A(\omega) := \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$$

Bemerkung:

Für A, B Ereignisse gilt:

$$\mathbb{1}_{A \cap B} = \mathbb{1}_A \cdot \mathbb{1}_B$$

$$\mathbb{1}_{A \cup B} \le \mathbb{1}_A + \mathbb{1}_B$$

Def. 5 :

Z.Var. $X_1, ..., X_n$ auf einem Wkts.raum/-modell $(\Omega, \mathbb{P})$ heißen unabhängig, falls die Ereignisse $\{X_1 = x_1\}, ..., \{X_n = x_n\}$ unabhängig sind für $\forall x_i \in X_i(\Omega), \quad i = 1, ..., n$

$\implies \quad X, Y$ unabh., falls $\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y)$ für alle x,y gilt.

Proposition 6 :

Seien $(\Omega, \mathbb{P})$ und $(\Omega', \mathbb{P}')$ Wkts.räume mit unabh. Z.Var. $X_1, ..., X_n$ auf Ω, bzw. $X_1', ..., X_n'$ auf Ω'.

Wenn X_i die gleiche Verteilung wie X_i' hat $\forall i$ (d.h. $p_{X_i} = p_{X_i'} \quad \forall i$), dann gilt

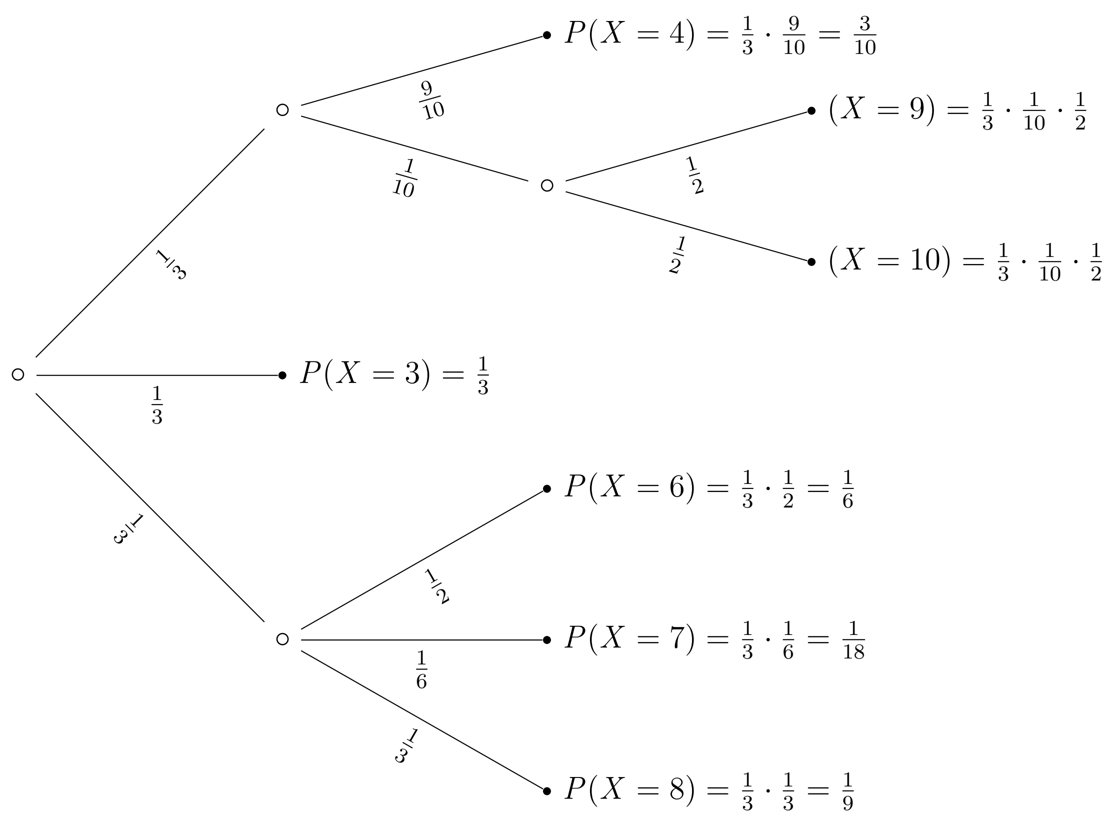
$$\mathbb{P}(f(X_1, ..., X_n) = y) = \mathbb{P}'(f(X_1', ..., X_n') = y) \quad \forall f : \mathbb{R}^n \to \mathbb{R} \text{ und alle } y$$

D.h. die Z.Var $(X_1, ..., X_n)$ und $(X_1', ..., X_n')$ beschreiben das gleiche Modell, auch wenn die Wkts.räume unterschiedlich sind.

$\rightarrowtail$ Es reicht, die Verteilung/Wkts.funktionen der unabhängigen Z.Var anzugeben, die man als Bausteine benutzt. Den Wkts.raum $(\Omega, \mathbb{P})$ muss man dann nicht genauer spezifizieren.

1.2.2 Mehrstufige Experimente und Entscheidungsbäume

Beispiel :



Sei ein endlicher Baum (mit Wurzel) gegeben, sowie Wkt.en an Kanten. Für jeden inneren Knoten müssen sich die Wk.ten zu den Kindern zu 1 summieren.

- In jedem Schritt wählt man zufällig ein Kind nach den gegebenen Wkt.en, und wandert so den Baum hinab, bis man an ein Blatt kommt.
- Das Endergebnis hängt davon ab in welchem Blatt man landet

Satz 7 :

Sei die Spezifikation eines mehrstufigen Experiments gegeben, d.h. ein Baum (mit nummerierten Knoten und Wurzel 0) und Wkt.en p_{ij} für Kanten (i,j) mit

$$\sum_{j \text{ kind von } i} p_{ij} = 1$$

Sei die Z.Var. X der Index des zufälligen Blatts, in dem man landet. Dann gilt

$$\mathbb{P}(X = x) = p_{0x_1} \cdot p_{x_1x_2} \cdot \ldots \cdot p_{x_{n-1}x}, \quad \text{wobei } 0x_1x_2 \ldots x_{n-1}x \text{ der Pfad von 0 nach x ist.}$$

Bemerkung:

(*) nennt man Pfadregel

Beweis :

Die einzelnen Faktoren in der Pfadregel entsprechen den unabhängigen Entscheidungen in jedem Schritt.

$$\begin{aligned}\mathbb{P}(X=x) &= \mathbb{P}(\{0 \to x_1\} \cap \{x_1 \to x_2\} \cap \ldots \cap \{x_{n-1} \to x\}) \\ &\overset{\text{unabh.}}{=} \mathbb{P}(\{0 \to x_1\}) \cdot \mathbb{P}(\{x_1 \to x_2\}) \cdot \ldots \cdot \mathbb{P}(\{x_{n-1} \to x\}) \\ &= p_{0x_1} \cdot p_{x_1x_2} \cdot \ldots \cdot p_{x_{n-1}x},\end{aligned}$$

wobei $\{x_k \to x_{k+1}\}$ das Ereignis ist, dass wir in dem Schritt k nach x_{k+1} gehen, wenn wir in x_k sind. $\square$

Def. 8 :

Eine Z.Var X mit Werten in $\{0,1\}$ und mit Wkts.funktion p_X mit
$p_X(1) = \alpha, \quad p_X(0) = 1 - \alpha$ heißt Benulli-verteilt mit Parameter α.
$\to$ Fachbegriff für einen (unfairen) Münzwurf.

Beispiel :

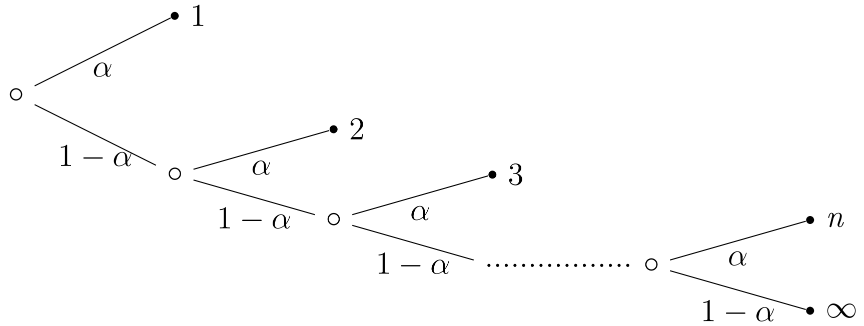
Seien $X_1, ..., X_n$ unabhängige Bernoulli-verteilte Z.Var.en mit Parameter $\alpha \in (0,1]$ auf einem Wkts.raum $(\Omega, \mathbb{P})$

$$S := \sum_{i=1}^{n} X_i = \#\text{Erfolge} \qquad \text{ist Binomial (n,a) -verteilt}$$

$$Y := \begin{cases} min\{i \in \{1,...,n\} : X_i = 1\}, & S > 0 \\ \infty, & S = 0 \end{cases}$$

$=$ #Versuche bis zum ersten Erfolg
Verteilung/Wkts.funktion von Y? $\quad p_Y(k) = \mathbb{P}(Y = k) = ?$

Alternative Betrachtungsweise als mehrstufiges Experiment:



X_i entscheidet, ob man in Stufe i nach links $X_i = 1$ oder rechts $X_i = 0$ läuft.

$$p_Y(k) = (1-\alpha)^{k-1} \cdot \alpha \qquad \text{für} \quad k \in \{1, ..., n\}$$

$$p_Y(\infty) = (1-\alpha)^n$$

Def. 9 :

Eine Z.Var Y ist geometrisch verteilt mit Erfolgsparameter $\alpha \in (0,1]$, falls

$$p_Y(k) = \alpha(1-\alpha)^{k-1} \qquad \text{ist,} \quad k \in \mathbb{N}$$

Die geometrische Verteilung zählt die Anzahl der Versuche bis zum ersten Erfolg in einer unendlich langen Folge von unabh. Bernoulli-Experimenten (Münzwürfe)

Bemerkung:

Manchmal zählen geometrische Vert. die Anzahl der Fehlversuche bis zum ersten Erfolg.

Bemerkung: (nicht prüfungsrelevant)

Eine unendliche Folge von 0-1-Bernoulli-Experimenten würde auf dem Ergebnisraum $\Omega = \{0,1\}^{\mathbb{N}}$ leben. Aber Ω ist nicht abzählbar und für $\alpha \in (0,1)$ ist

$$p(\omega) = \alpha^{\#\text{1en in } \omega} \cdot (1-\alpha)^{\#\text{0en in } \omega} = 0$$

→ Jedes mögliche Ergebnis hat Wkt. 0
→ diskrete Wkts.theorie ist hier nicht gut genug

- diskrete WT geht auf Laplace (1749-1827) zurück
- moderne WT geht auf Kolmogorov (1933) zurück

1.2.3 Erwartungswert und Varianz

In disem Abschnitt nehmen wir einen Wkts.raum $(\Omega, \mathbb{P})$ als gegeben an

Def. 10 :

Der Erwartungswert einer Z.Var X auf Ω ist definiert als

$$\mathbb{E}(X) := \sum_{x \in X(\Omega)} x \cdot \mathbb{P}(X = x), \qquad \text{falls} \quad |X(\Omega)| < \infty$$

So wird zusätzlich gefordert, dass

$$\sum_{x \in X(\Omega), x > 0} x \cdot \mathbb{P}(X = x) < \infty \tag{1}$$

oder

$$\sum_{x \in X(\Omega), x < 0} x \cdot \mathbb{P}(X = x) > -\infty \tag{2}$$

ist.

Bemerkung:

- Der Erwartungswert ist das durchschnittliche Ergebnis der Z.Var.
- Der Erwartungswert kann $+\infty$ (bzw. $-\infty$) sein, falls nur (2) (bzw. nur (1)) gilt.
  Meistens ist der Erwartungswert endlich.
- der Erwartungswert von X hängt nur von der Wkts.funktion p_X ab: $\mathbb{E}(X) = \sum_{x \in X(\Omega)} x p_X(x)$.

Beispiel :

Sei X das Ergebnis eines Würfelwurfs

$$\mathbb{E}(X) = \sum_{x=1}^{6} x \cdot \underbrace{\mathbb{P}(X = x)}_{=6} = \frac{1 + 2 + ... + 6}{6} = 3,5$$

Satz 11 (Rechenregeln):

Seien X, Y Z.Var mit endlichen Erwartungswerten ((1) und (2) gelten), und $a, b \in \mathbb{R}$.

a) $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$

b) $\mathbb{E}(aX+b) = a\mathbb{E}(X) + b$

c) Für $g : \mathbb{R} \to \mathbb{R}$ mit $\mathbb{E}(g(X))$ endlich ist

$$\mathbb{E}(g(X)) = \sum_{x \in X(\Omega)} g(x) \cdot \mathbb{P}(X = x)$$

Beispiel (X wieder Würfelung):

Sei $Y = \underbrace{3 \cdot X}_{\text{Auszahlung des Spiels}} - \underbrace{11}_{\text{Einsatz}}$

$$\mathbb{E}(Y) = \mathbb{E}(3X - 11) = 3\mathbb{E}(X) - 11 = 3 \cdot 3,5 - 11 = -0,5$$

Beispiel (Glücksspiel basierend auf Münzwurf):

Kopf: 11€ Gewinn, Zahl: 10€ Verlust
$\Omega = \{0, 1\}$, $\mathbb{P}$ Gleichverteilung

$$X(\omega) = \begin{cases} 11, & \omega = 1 \\ -10, & \omega = 0 \end{cases}$$

$$\mathbb{E}(X) = 11 \cdot \mathbb{P}(X = 11) - 10 \cdot \mathbb{P}(X = -10) = 11 \cdot \frac{1}{2} - 10 \cdot \frac{1}{2} = 0.5$$

Ein gelangweilter Millionär ändert die Regeln:
Kopf: 1001€ Gewinn, Zahl: 1000€ Verlust

$$Y(\omega) = \begin{cases} 10^3 + 1, & \omega = 1 \\ -10^3, & \omega = 0 \end{cases}$$

$$\mathbb{E}(Y) = ... = 0.5$$

Def. 12 :

Sei X eine Z.Var mit endlichem Erwartungswert μ. Dann ist die **Varianz** von X definiert als

$$Var(X) := \mathbb{E}((X - \mu)^2) = \sum_{x \in X(\Omega)} (x - \mu)^2 \mathbb{P}(X = x)$$

Die **Standardabweichung** $\sigma(X)$ ist definiert als

$$\sigma(X) = \sqrt{Var(X)}$$

Bemerkung:

- Die Standardabweichung ist ein Maß dafür, wie weit eine Z.Var von ihrem Erwartungswert abweichen kann.
- $Var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$

Beispiel (Fortsetztung von oben):

$$\begin{aligned} Var(X) &= \mathbb{E}((X - \frac{1}{2})^2) \\ &= (X - \frac{1}{2})^2 \mathbb{P}(X = 11) + (-10 - \frac{1}{2})^2 \mathbb{P}(X = -10) \\ &= (10,5)^2 \cdot \frac{1}{2} + (-10,5)^2 \cdot \frac{1}{2} \\ &= 110,25 \approx 10^2 \end{aligned}$$

$$\sigma(X) \approx 10^1 \qquad \rightarrow \frac{1}{2} \pm 10^1$$

$$\begin{aligned} Var(Y) &= \mathbb{E}((Y - \frac{1}{2})^2) \\ &= (10^3 + 1 - \frac{1}{2})^2 \cdot \frac{1}{2} + (-10^3 - \frac{1}{2})^2 \cdot \frac{1}{2} \\ &= 1001000,25 \approx 10^6 \end{aligned}$$

$$\sigma(Y) \approx 10^3 \qquad \rightarrow \frac{1}{2} \pm 10^3$$

Bemerkung:

- $X \approx \mathbb{E}(X) \pm \sigma(X)$ ist eine mathematisch falsche, aber hilfreiche Darstellung, um das Verhalten von X zu veranschaulichen

Def. 13 :

Seien X, Y Z.Var mit endlichem Erwartungswert. Die **Kovarianz** von X und Y ist definiert als

$$Cov(X,Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}(X) \cdot \mathbb{E}(Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y)))$$

Zwei Z.Var heißen unkorreliert, falls $Cov(X,Y) = 0$.

Satz 14 (Rechenregeln):

Seien X, Y Z.Var mit endlichem Erwartungswert, $a, b \in \mathbb{R}$

a) $Var(aX + b) = a^2 \, Var(X)$

b) $Var(X + Y) = Var(X) + Var(Y) + 2 \cdot Cov(X,Y)$

Satz 15 :

Sei X, Y **unabhängige** Z.Var. mit endlichem Erwartungswert. Dann gilt

$$\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$$

Insbesondere sind X und Y unkorreliert und $Var(X+Y) = Var(X) + Var(Y)$.

Beweis :

$$\begin{aligned}
\mathbb{E}(X \cdot Y) &= \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} x \cdot y \cdot \mathbb{P}(\{X = x\} \cap \{Y = y\}) \\
&\overset{unabh.}{=} \sum_x \sum_y x \cdot y \cdot \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y) \\
&= (\sum_x x \cdot \mathbb{P}(X = x)) \cdot (\sum_y y \cdot \mathbb{P}(Y = y)) \\
&= \mathbb{E}(X) \cdot \mathbb{E}(Y)
\end{aligned}$$

□

Satz 16 :

Sei X eine Z.Var. mit nicht-negativen Werten. Dann gilt für jedes $a > 0$

$$\mathbb{P}(X \geq a) \leq \frac{1}{a} \cdot \mathbb{E}(X)$$

Beispiel :

$$\mathbb{P}\left(X \geq 4 \cdot \mathbb{E}\right) \leq \frac{1}{4}$$

Beweis :

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{x \in X(\Omega)} x \cdot \mathbb{P}\left(X = x\right) \\
&\overset{x \geq 0}{\geq} \sum_{x \geq a} x \cdot \mathbb{P}\left(X = x\right) \\
&\geq \sum_{x \geq a} a \cdot \mathbb{P}\left(X = x\right) = a \cdot \mathbb{P}\left(X \geq a\right)
\end{aligned}$$

$\square$