universität wien

DIPLOMARBEIT

STOCHASTIK IN POLITISCHEN FRAGESTELLUNGEN

angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag. rer.nat.)

Verfasser	Markus Dorn
Matrikel-Nummer	9804366
Studienrichtung (lt. Studienblatt)	A 190 406 313
Betreuer	ao. Univ.-Prof. Dr. Peter Raith

Wien, am 10.August 2010

Inhaltsverzeichnis

VORWORT

Das Wort Stochastik kommt aus dem Altgriechischen und leitet sich von στοχαστικη τεχνη („stochastike techne“) ab. Übersetzt bedeutet das die Kunst des Vermutens, womit schon angedeutet ist, dass sich die Stochastik mit Ereignissen befasst, die vom Zufall abhängen. Die Stochastik ist ein Teilgebiet der Mathematik, das sich aus den beiden Bereichen Wahrscheinlichkeitsrechnung und Statistik zusammensetzt.

Als Lehramtsstudent ist mir die Anwendbarkeit von Mathematik ein großes Anliegen, zumal gerade SchülerInnen oftmals den Bezug der Mathematik zur Realität vermissen und sich fragen, welchen Sinn die Auseinandersetzung mit Mathematik eigentlich hat.

Ich bin der Meinung, dass gerade die Stochastik für LehrerInnen besonders geeignet ist, SchülerInnen für die Welt der Mathematik zu begeistern und einen interessanten, abwechslungsreichen Mathematikunterricht zu gestalten. Kaum ein anderer mathematischer Bereich lässt sich mit so vielen verschiedenen anderen Schulgegenständen kombinieren wie die Stochastik.

Mein Ziel war es spannende Beispiele aus der Politik auszuwählen, wo die Stochastik eine wichtige Rolle spielt. Die Beispiele sind größtenteils mit gewöhnlichem Maturawissen nachvollziehbar und könnten somit bereits auch im Schulunterricht angesprochen werden.

Ich habe alle Lebensdaten der vorkommenden Personen angegeben, sofern sie recherchierbar waren. Damit wollte ich eine Einordnung der besprochenen Inhalte in einen historischen Kontext ermöglichen sowie zeigen, dass Mathematik eine Wissenschaft ist, die bis in die Gegenwart hinein von neuen Impulsen angetrieben wird.

Bei den Grundlagen habe ich die verwendete Literatur nur im Literaturverzeichnis angegeben. Bei einigen speziellen Themen (Kapitel 3.4: Simpsonsche Paradoxon, Kapitel 9: Mandatsverteilungen, Kapitel 10: Machtindizes und Kapitel 14: Randomized-Response-Methoden) erschien es mir jedoch sinnvoll die verwendete Literatur auch gleich im Anschluss an dieses Kapitel angegeben.

Die Diplomarbeit wurde mit Microsoft Word 2003 verfasst. Die Formeln habe ich mit dem Zusatzprogramm MathType 6.5 erstellt. Alle Berechnungen, die für diese Diplomarbeit notwendig waren, habe ich mit den beiden Computerprogrammen Derive 6 bzw. mit Mathematica 6 ausgeführt. Die Grafiken wurden großteils mit Derive 6 erstellt und anschließend mit diversen Bildbearbeitungsprogrammen (hauptsächlich Paint 5.1 und Microsoft Office Picture Manager 2003) überarbeitet, wobei mir mein Freundeskreis dabei eine sehr große Hilfe war.

Mein Dank gilt einerseits Univ.-Prof. Dr. Peter Raith für seine hervorragende Betreuung beim Erstellen dieser Diplomarbeit. Er stand mir jederzeit mit Rat und Tat zur Seite und hat all meine Fragen mit mir erörtert, die sich im Laufe dieser Diplomarbeit ergeben haben.

Andererseits möchte ich auch meinen Eltern Waltraud und Walter Dorn danken, die mir mein Studium ermöglicht haben und mich all die Jahre hindurch unterstützt haben.

Ich widme diese Diplomarbeit meiner Mutter, die nach langer, schwerer Krankheit im März 2010 verstorben ist und das Ende meines Studiums nicht mehr erleben konnte, ohne die ich aber diese Diplomarbeit nicht verwirklichen hätte können.

TEIL 1

WAHRSCHEINLICHKEITSRECHNUNG

1. GRUNDLAGEN DER WAHRSCHEINLICHKEITSRECHNUNG

Zunächst einmal möchte ich einen mengentheoretischen Einstieg in die Wahrscheinlichkeitstheorie geben und die wichtigsten Begriffe anhand von einigen anschaulichen Beispielen einführen.

Ich werde zwei intuitive Zugänge zum Wahrscheinlichkeitsbegriff erörtern. Zum einen die naive Wahrscheinlichkeit, die der Österreicher Mises 1931 angegeben hat und die den Wahrscheinlichkeitsbegriff als einen Grenzwert der relativen Häufigkeiten auffasst. Diese Wahrscheinlichkeit wird dann in Kapitel 12 durch die Gesetze der großen Zahlen präzisiert werden. Und zum anderen die Laplace-Wahrscheinlichkeit, die die Wahrscheinlichkeit als Quotient der Anzahl der günstigen Fälle und der Anzahl aller möglichen Fälle definiert.

Im Zuge dessen werde ich als Anwendungsbeispiel zeigen, wie sich die Wahrscheinlichkeit bestimmen lässt, dass eine Partei in einem Drei-Parteien-System die absolute Mehrheit bei einer Wahl erreicht. Hierbei wird die Laplace-Wahrscheinlichkeit geometrisch als Fläche gedeutet.

Den Abschluss dieses Einführungskapitels bilden dann die Axiome von Kolmogorow, welche das Fundament der Wahrscheinlichkeitsrechnung bilden. Erst durch diese Axiome wurde die Wahrscheinlichkeitsrechnung wissenschaftlich anerkannt. Kolmogorow verwendete dieses Axiomensystem erstmals in seinem 1933 erschienenen Buch „Grundbegriffe der Wahrscheinlichkeitsrechnung“.

1.1 Grundbegriffe und Mengenlehre

Für die Wahrscheinlichkeitstheorie ist ein grundlegendes Verständnis der Mengenlehre notwendig.

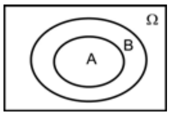
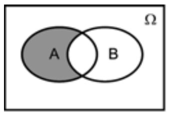
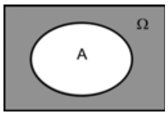
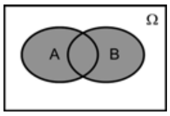
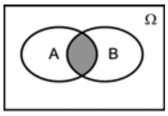
Operation	Erklärung	Graphik
$A = \Omega$	Ereignis *A* ist ein sicheres Ereignis und tritt daher stets ein.	
$A = \emptyset$	Ereignis *A* ist ein unmögliches Ereignis und kann folglich nie eintreten.	
$A \subseteq B$	Ereignis *B* beinhaltet Ereignis *A*. Wenn also *A* eintritt, dann auch *B*.	
$A \setminus B$	Ereignis *A* ist eingetreten, nicht jedoch Ereignis *B*.	
$A^c = \Omega \setminus A$	A^c ist das Komplementereignis von *A* bezüglich Ω.	
$A \cup B$ $\left(\text{allg. } \bigcup_{i \in I} A_i\right)$	Entweder Ereignis *A* oder Ereignis *B* oder beide Ereignisse treten ein.	
$A \cap B$ $\left(\text{allg. } \bigcap_{i \in I} A_i\right)$	Sowohl Ereignis *A* als auch Ereignis *B* treten gleichzeitig ein. Falls $A \cap B = \emptyset$, dann sind die beiden Ereignisse disjunkt.	

Tabelle 1 Mengentheoretische Operationen für Ereignisse von Ω

Ein *Zufallsexperiment* ist ein beliebig oft wiederholbarer Vorgang mit den folgenden beiden Eigenschaften:

- ♠ Alle möglichen Ergebnisse sind im Voraus bekannt.
- ♠ Der tatsächliche Ausgang ist jedoch im Voraus unbekannt.

Die möglichen einelementigen Ergebnisse eines Zufallsexperiments nennt man *Elementarereignisse* und bezeichnet sie mit ω. Die Menge aller Elementarereignisse ω bildet den *Ereignisraum* Ω. Jede Teilmenge von Ω ist ein *Ereignis* und wird mit einem gewöhnlichen Großbuchstaben angegeben.

Nicht jedes Ereignis muss *zulässig* sein. Wir nennen ein Ereignis zulässig, falls eine Wahrscheinlichkeit berechnet werden kann.

Sei $A \subseteq \Omega$ ein beliebiges Ereignis und ω ein Elementarereignis. Dann bedeutet, $\omega \in A$ dass *A* eingetreten ist.

KOROLLAR 1.1 *Mengentheoretische Rechengesetze*

Seien $A, B, C \in \Omega$. Dann gelten folgende Gesetze:

(i) *Kommutativgesetze:*

$$A \cap B = B \cap A, A \cup B = B \cup A$$

(ii) Assoziativgesetze:

$$(A \cap B) \cap C = A \cap (B \cap C), (A \cup B) \cup C = A \cup (B \cup C)$$

(iii) *Distributivgesetze:*

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

(iv) *Adjunktivitätsgesetze:*

$$A \cap (A \cup B) = A, A \cup (A \cap B) = A$$

(v) *Idempotenzgesetze:*

$$A \cap A = A, A \cup A = A$$

(vi) *Identitätsgesetze:*

$$A \cap \emptyset = \emptyset, A \cap \Omega = A$$

$$A \cup \emptyset = A, A \cup \Omega = \Omega$$

(vii) *Komplementgesetze:*

$$A \cap A^c = \emptyset, A \cup A^c = \Omega$$

$$\left(A^c\right)^c = A, \Omega^c = \emptyset, \emptyset^c = \Omega$$

(viii) *Gesetze von De Morgan:*

$$(A \cap B)^c = A^c \cup B^c, (A \cup B)^c = A^c \cap B^c$$

KOROLLAR 1.2 *Verallgemeinerung der Gesetze von De Morgan*

Seien $A, B, C \in \Omega$. Dann gelten folgende Gesetze:

(i) $\left(\bigcap_{i \in I} A_i\right)^c = \bigcup_{i \in I} A_i^c$

(ii) $\left(\bigcup_{i \in I} A_i\right)^c = \bigcap_{i \in I} A_i^c$

BEISPIEL 1.3 *Beispiele für Elementarereignisse und Ereignisräume*

- ♠ Zeitpunkt der Wahlentscheidung eines Wählers:
  $\Omega = \{$kurz vorher, 1-2 Wochen vorher, mehr als 2 Wochen vorher$\}$
- ♠ Wahlverhalten eines Wählers: $\Omega = \{$Stammwähler, Wechselwähler$\}$
- ♠ Geschlecht eines Wählers: $\Omega = \{$männlich, weiblich$\}$
- ♠ Schulbildung eines Wählers:
  $\Omega = \{$Pflichtschule, Fachschule, Matura, Studienabschluss$\}$
- ♠ Beruf eines Wählers:
  $\Omega = \{$Beamter, Angestellter, selbständig, Pensionist, sonstiges$\}$

BEISPIEL 1.4 *Ankreuzen eines Stimmzettels*

- ♠ $\Omega = \{$SPÖ, ÖVP, FPÖ, Grüne, sonstige Parteien, ungültig$\} =$

  $= \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$
- ♠ $\omega_i = \{$eine bestimmte Partei wird angekreuzt$\}$
- ♠ Mögliche Ereignisse (Teilmengen von Ω) sind etwa:

  $A = \{$Wähler entscheidet sich für die große Koalition$\} = \{\omega_1, \omega_2\}$

  $B = \{$Wähler wählt nicht FPÖ$\} = \{\omega_1, \omega_2, \omega_4, \omega_5, \omega_6\}$

  $C = \{$Wähler wählt keine der vier größten Parteien$\} = \{\omega_5, \omega_6\}$

Sei Ω ein beliebiger Ereignisraum eines Zufallsexperiments. Dann bezeichnet Ω *das sichere Ereignis*, welches immer eintreten muss, und $\emptyset$ *das unmögliche Ereignis*, welches nie eintreten kann.

DEFINITION 1.5 Sei $\Omega \neq \emptyset$. Ein System $\mathcal{A}$ von Teilmengen von Ω heißt *σ-Algebra, falls*

(i) $\emptyset \in \mathcal{A}$

(ii) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$

(iii) $(A_i)_{i\in\mathbb{N}} \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$

PROPOSITION 1.6 *Eigenschaften der σ-Algebra $\mathcal{A}$*

(i) $\Omega \in \mathcal{A}$

(ii) $(A_i)_{i\in\mathbb{N}} \in \mathcal{A} \Rightarrow \bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$

BEWEIS:

(i) $\emptyset \in \mathcal{A} \underset{1.5\text{ (ii)}}{\overset{\text{Def.}}{\Rightarrow}} \emptyset^c \in \mathcal{A} \underset{1.1\text{ (vii)}}{\overset{\text{Kor.}}{\Rightarrow}} \Omega \in \mathcal{A}$

(ii) $A_i^c \in \mathcal{A} \underset{1.5\text{ (iii)}}{\overset{\text{Def.}}{\Rightarrow}} \bigcup_{i=1}^{\infty} A_i^c \in \mathcal{A} \underset{1.5\text{ (ii)}}{\overset{\text{Def.}}{\Rightarrow}} \Omega \setminus \bigcup_{i=1}^{\infty} A_i^c \in \mathcal{A}$

$\underset{1.2\text{ (i)}}{\overset{\text{Kor.}}{\Rightarrow}} \Omega \setminus \left(\bigcap_{i=1}^{\infty} A_i\right)^c \in \mathcal{A} \underset{1.1\text{ (vii)}}{\overset{\text{Kor.}}{\Rightarrow}} \bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$

□

1.2 Intuitive Zugänge zum Wahrscheinlichkeitsbegriff

1.2.1 Naive Wahrscheinlichkeit

Sei $r_n(A) = \frac{h_n(A)}{n}$ die *relative Häufigkeit* für den Anteil von *n* Versuchen (etwa Befragungen von Personen), bei denen das Ereignis *A* eingetreten ist. Die Wahrscheinlichkeit für das Eintreten eines bestimmten Ereignisses kann durch dessen relative Häufigkeit abgeschätzt werden. Für große *n* nähern sich die relativen Häufigkeiten immer mehr einem festen Zahlenwert an,

Definition 1.7 *Naive Wahrscheinlichkeit*

$$P(A) := \lim_{n\to\infty} \frac{h_n(A)}{n}$$

Probleme der Definition der naiven Wahrscheinlichkeit:

- ♠ Der Grenzwert lässt sich nicht bestimmen, weil das Experiment nicht unendlich oft durchgeführt werden kann.
- ♠ Selbst wenn das Experiment unendlich oft durchgeführt werden könnte, so müsste der Grenzwert nicht einmal existieren.
- ♠ Aus verschiedenen Erhebungen würde man für dasselbe Ereignis im allgemeinen verschiedene Wahrscheinlichkeiten erhalten.

Proposition 1.8 *Eigenschaften der naiven Wahrscheinlichkeit*

(i) $0 \le P(A) \le 1$

(ii) $A \le B \Rightarrow P(A) \le P(B)$

(iii) $P(\varnothing) = 0$

(iv) $P(\Omega) = 1$

(v) $A \cap B = \varnothing \Rightarrow P(A \cup B) = P(A) + P(B)$

(vi) $P(A^c) = 1 - P(A)$

Beweis:

(i)
$$0 \le h_n(A) \le n \overset{:n}{\Rightarrow} 0 \le \underbrace{\frac{h_n(A)}{n}}_{\to P(A)} \le 1$$

(ii) Sei $h_n(A) \le h_n(B)$

$$\Rightarrow P(A) = \lim_{n\to\infty} \frac{h_n(A)}{n} \le \lim_{n\to\infty} \frac{h_n(B)}{n} = P(B)$$

(iii)
$$h_n(\varnothing) = 0 \Rightarrow P(\varnothing) = \lim_{n\to\infty} \underbrace{\frac{h_n(\varnothing)}{n}}_{=0} = 0$$

(iv)
$$h_n(\Omega) = n \Rightarrow P(\Omega) = \lim_{n\to\infty} \underbrace{\frac{h_n(\Omega)}{n}}_{=1} = 1$$

(v) Sei $h_n(A \cup B) = h_n(A) + h_n(B)$

$$\Rightarrow P(A \cup B) = \lim_{n\to\infty} \frac{h_n(A \cup B)}{n} = \lim_{n\to\infty}\left(\frac{h_n(A)}{n} + \frac{h_n(B)}{n}\right) =$$

$$= \lim_{n\to\infty} \frac{h_n(A)}{n} + \lim_{n\to\infty} \frac{h_n(B)}{n} = P(A) + P(B)$$

(vi) $1 \underset{1.8\text{ (iv)}}{\overset{\text{Prop.}}{=}} P(\Omega) \underset{1.1\text{ (vii)}}{\overset{\text{KOR.}}{=}} P(A \cup A^c) \underset{1.8\text{ (v)}}{\overset{\text{Prop.}}{=}} P(A) + P(A^c)$

$\Rightarrow P(A^c) = 1 - P(A)$

□

1.2.2 LAPLACE-WAHRSCHEINLICHKEIT

Ein *Laplace-Experiment* ist ein Zufallsexperiment mit den beiden folgenden Eigenschaften:

- ♠ Der Ereignisraum Ω ist endlich.
- ♠ Alle Elementarereignisse ω_i sind gleichwahrscheinlich.

DEFINITION 1.9 *Laplace-Wahrscheinlichkeit*

Sei n die Anzahl des Ereignisraumes Ω, k die Anzahl der Elementarereignisse der Teilmenge A und $P(A)$ die Wahrscheinlichkeit, dass das Ereignis A eintritt. Dann gilt

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

Alle Elementarereignisse ω_i haben die Wahrscheinlichkeit

$$P(\omega_i) = \frac{1}{n}.$$

PROPOSITION 1.10 *Eigenschaften der Laplace-Wahrscheinlichkeit*

(i) $P(\emptyset) = 0$

(ii) $P(\Omega) = 1$

(iii) $P(A^c) = 1 - P(A)$

BEWEIS:

(i) $P(\varnothing) = \frac{0}{n} = 0$

(ii) $P(\Omega) = \frac{n}{n} = 1$

(iii) $P(A^c) = \frac{n-k}{n} = 1 - \frac{k}{n} = 1 - P(A)$

□

Die Laplace-Wahrscheinlichkeit ist eine Funktion $P(\Omega) \to [0,1]$, die man als Wahrscheinlichkeitsmaß bezeichnet.

BEISPIEL 1.11 *Die österreichische Bundespräsidentenwahl von 2004*[1]

Stimmen	Anzahl
Wahlberechtigte	6030982
abgegebene Stimmen	4318439
gültige Stimmen	4136016
abgegebene Stimmen für Dr. Heinz Fischer	2166690

TABELLE 2 Ausgang der Bundespräsidentenwahl von 2004

a) Wie hoch war die Wahlbeteiligung? (=Wie groß ist die Wahrscheinlichkeit, dass ein Österreicher zur Wahl ging?)
b) Wie viel Prozent der gültigen Stimmen entfielen auf Dr. Heinz Fischer? (=Wie wahrscheinlich war es, dass sich ein Wähler für Dr. Heinz Fischer entschieden hat?)

a) $P(\text{Wahlberechtigter ging wählen}) = \frac{4318439}{6030982} = 0{,}7160 \Rightarrow \underline{\underline{71{,}60\,\%}}$

b) $P(\text{Dr. Heinz Fischer wurde gewählt}) = \frac{2166690}{4136016} = 0{,}5239 \Rightarrow \underline{\underline{52{,}39\,\%}}$

Die Laplace-Wahrscheinlichkeit kann auch als Flächeninhalt gedeutet werden, indem man das Verhältnis der Teilfläche zur Gesamtfläche betrachtet.

DEFINITION 1.12 *Geometrische Wahrscheinlichkeit*

Sei $\Omega \subseteq \mathbb{R}^2$ ein Gebiet mit dem endlichen Flächeninhalt $F(\Omega)$ und sei $A \subseteq \Omega$ ein Teilgebiet mit dem Flächeninhalt $F(A)$. Dann gilt

$$P(A) = \frac{F(A)}{F(\Omega)} = \frac{\text{Flächeninhalt von } A}{\text{Flächeninhalt von } \Omega}.$$

[1] http://www.bmi.gv.at/wahlen/wahldownloads/bpwahl04_ErgE.pdf [29.12.2008]

BEISPIEL 1.13 *Absolute Mehrheit in einem Drei-Parteien-System*[2]

In einem Land treten drei Parteien zur Wahl an. Wie wahrscheinlich ist es, dass eine der Parteien die absolute Mehrheit erreichen kann, wenn aufgrund fehlender Zusatzinformationen alle möglichen Wahlausgänge als gleichwahrscheinlich angesehen werden?

Ausgangspunkt für unsere Überlegung ist ein gleichseitiges Dreieck mit der Höhe 1, wobei jeder Partei eine Dreiecksseite zugeteilt wird. Nun ist aus der Geometrie bekannt, dass für jeden Punkt innerhalb eines gleichseitigen Dreiecks die Summe der Abstände zu den drei Seiten gleich der Dreieckshöhe ist.

Es ist daher naheliegend, dass diese Abstände den prozentuellen Stimmanteil p_i der i-ten Partei wiedergeben, da $p_1 + p_2 + p_3 = 1$. Somit kann jedem möglichen Wahlergebnis genau ein Punkt innerhalb des Dreiecks zugeordnet werden.

Offensichtlich entsprechen die Wahlergebnisse mit absoluter Mehrheit genau der Menge der Dreieckspunkten, die einen Abstand von mindestens ½ zu einer der Dreiecksseiten haben. Man erhält vier gleich große Dreiecke (siehe Abbildung 1 rechts). Aufgrund von Definition 1.12 ist die gesuchte Wahrscheinlichkeit somit der Quotient aus der schraffierten Fläche und der Gesamtfläche und beträgt ¾.

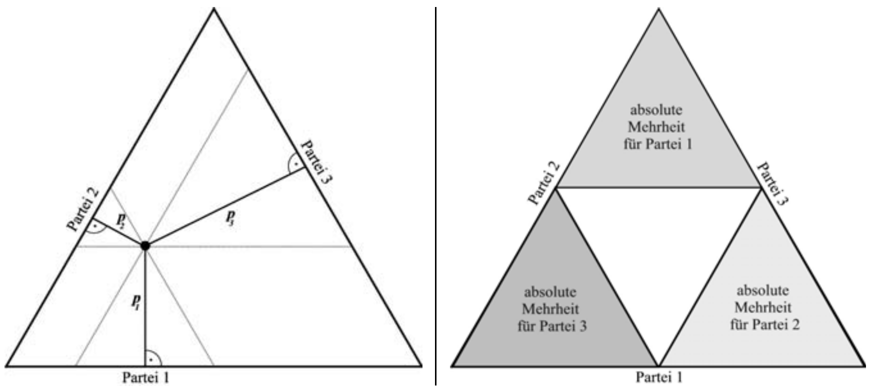


ABBILDUNG 1 Graphische Darstellung des Wahlergebnisses in einem Drei-Parteien-System (links) sowie der absoluten Mehrheit in diesem System (rechts)

[2] HESSE, Christian / MEISTER, Alexander, Übungsbuch zur angewandten Wahrscheinlichkeitstheorie. Aufgaben und Lösungen, Wiesbaden 2005, S.105ff.

1.3 Axiomatischer Aufbau

Definition 1.14 *Axiome von Kolmogorow*

Sei $\Omega \neq \emptyset$ und $\mathcal{A}$ eine σ-Algebra auf Ω. Dann heißt eine Funktion $P : \mathcal{A} \to \mathbb{R}$ *Wahrscheinlichkeit(smaß)*, falls

(i) $P(A) \geq 0$ (*Nichtnegativität*)

(ii) $P(\Omega) = 1$ (*Normierung*)

(iii) $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ für paarweise disjunkte Ereignisse $(A_i)_{i\in\mathbb{N}}$

(σ-*Additivität*)

Das Wort Wahrscheinlichkeit wird in zweifacher Hinsicht verwendet. Zum einen für die Funktion *P*, zum anderen aber auch für den Wert $P(A)$.

Definition 1.15 *Wahrscheinlichkeitsraum*

Sei $\Omega \neq \emptyset$, $\mathcal{A}$ eine σ-Algebra und $P : \mathcal{A} \to \mathbb{R}$ ein Wahrscheinlichkeitsmaß. Dann nennt man $(\Omega, \mathcal{A}, P)$ einen Wahrscheinlichkeitsraum.

Satz 1.16 *Eigenschaften des Wahrscheinlichkeitsmaßes P*

Seien $A, B \in \Omega$. Dann gilt:

(i) $P(\emptyset) = 0$

(ii) $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$ für paarweise disjunkte Ereignisse $(A_i)_{i\in\mathbb{N}}$

(*endliche Additivität*)

(iii) $P(A^c) = 1 - P(A)$

(iv) $A \subseteq B \Rightarrow P(A) \leq P(B)$

(v) $0 \leq P(A) \leq 1$

(vi) $P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i)$ für beliebige Ereignisse

Beweis:

(i) Setze $A_i := \emptyset \; \forall i$

$$\Rightarrow P(\emptyset) = P\left(\bigcup_{i=1}^{\infty} A_i\right) \underset{1.14\text{ (iii)}}{\overset{\text{Def.}}{=}} \sum_{i=1}^{\infty} P(A_i) = P(\emptyset) + P(\emptyset) + \ldots$$

$$\Rightarrow P(\emptyset) = 0$$

(ii) $P(\emptyset) \underset{1.16\text{ (i)}}{\overset{\text{Satz}}{=}} 0$

Setze $A_i := \emptyset \; \forall i \geq n+1$ in Definition 1.14 (iii)

(iii) $1 \underset{1.14\text{ (ii)}}{\overset{\text{Def.}}{=}} P(\Omega) \underset{1.3\text{ (vii)}}{\overset{\text{Kor.}}{=}} P\left(A \cup A^c\right) \underset{1.16\text{ (ii)}}{\overset{\text{Satz}}{=}} P(A) + P\left(A^c\right)$

$\Rightarrow P\left(A^c\right) = 1 - P(A)$

(iv) Setze $B := \bigcup_{i=1}^{\infty} A_i,\ A_1 := A,\ A_2 := B \setminus A,\ A_i := \varnothing\ \forall i \geq 3$

$$\Rightarrow P(B) = P\left(\bigcup_{i=1}^{\infty} A_i\right) \underset{1.14\text{ (iii)}}{\overset{\text{Def.}}{=}} \sum_{i=1}^{\infty} P(A_i) =$$

$$= P(A_1) + P(A_2) = P(A) + \underbrace{P(B \setminus A)}_{\geq 0 \text{ (wegen Def. 1.14 (i))}} \geq P(A)$$

(v) Sei $A \subseteq \Omega$

$\Rightarrow 0 \underset{1.14\text{ (i)}}{\overset{\text{Def.}}{\leq}} P(A) \underset{1.16\text{ (iv)}}{\overset{\text{Satz}}{\leq}} P(\Omega) \underset{1.14\text{ (ii)}}{\overset{\text{Def.}}{=}} 1$

(vi) Sei $\bigcup_{i=1}^{\infty} A_i = A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus (A_2 \cup A_1)) \cup \ldots$

$$\Rightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = P(A_1) + \underbrace{P(A_2 \setminus A_1)}_{\leq P(A_2) \text{ (wegen Satz 1.16 (iv))}} + \ldots \leq \sum_{i=1}^{\infty} P(A_i)$$

□

SATZ 1.17 *Additionsssatz von Sylvester*

Für zwei beliebige Ereignisse $A_1, A_2 \in \mathscr{A}$ gilt

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

BEWEIS:

$$P(A \cup B) = P\left(A \cup \left(A^c \cap B\right)\right) \underset{1.16\text{ (ii)}}{\overset{\text{Satz}}{=}} P(A) + P\left(A^c \cap B\right)$$

$$P(B) = P\left((A \cap B) \cup \left(A^c \cap B\right)\right) \underset{1.16\text{ (ii)}}{\overset{\text{Satz}}{=}} P(A \cap B) + P\left(A^c \cap B\right)$$

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

□

SATZ 1.18 *Stetigkeitseigenschaft*

Sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum und $(A_i)_{i \in \mathbb{N}}$ eine Folge in $\mathscr{A}$.

(i) Sei $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$, dann gilt $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n \to \infty} P(A_n)$

(ii) Sei $A_1 \supseteq A_2 \supseteq A_3 \supseteq \ldots$, dann gilt $P\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{n \to \infty} P(A_n)$

BEWEIS:

(i) Setze $B_1 := A_1$ und $B_i := A_i \setminus A_{i-1}\ \forall i > 1$

Es gelte $(B_i)_{i\in\mathbb{N}}$ sind paarweise disjunkt,

$$\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i \text{ und } A_n = \bigcup_{i=1}^{n} B_i\ \forall n \in \mathbb{N}$$

$$\Rightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = P\left(\bigcup_{i=1}^{\infty} B_i\right) \underset{1.14\text{ (iii)}}{\overset{\text{Def.}}{=}} \sum_{i=1}^{\infty} P(B_i) =$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) \underset{1.16\text{ (ii)}}{\overset{\text{Satz}}{=}} \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} B_i\right) = \lim_{n\to\infty} P(A_n)$$

(ii) $A_1 \supseteq A_2 \supseteq A_3 \supseteq \ldots \Rightarrow A_1^c \subseteq A_2^c \subseteq A_3^c \subseteq \ldots$

$$\Rightarrow \underbrace{P\left(\bigcup_{i=1}^{\infty} A_i^c\right)}_{\underset{1.2\text{ (i)}}{\overset{\text{Kor.}}{=}} P\left(\bigcap_{i=1}^{\infty} A_i\right)^c} \underset{1.18\text{ (i)}}{\overset{\text{Satz}}{=}} \lim_{n\to\infty} \underbrace{P\left(A_n^c\right)}_{\underset{1.16\text{ (iii)}}{\overset{\text{Satz}}{=}} 1-P(A_n)}$$

$$\underset{1.16\text{ (iii)}}{\overset{\text{Satz}}{\Rightarrow}} 1 - P\left(\bigcap_{i=1}^{\infty} A_i\right) = 1 - \lim_{n\to\infty} P(A_n)$$

$$\Rightarrow P\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} P(A_n)$$

□

2. Kombinatorik

Mit Hilfe der Kombinatorik können die möglichen Anordnungen aller Elemente einer Menge *n* bzw. einer Teilmenge daraus ermittelt werden. Die Kombinatorik ist Grundlage für die Berechnung von Machtindizes in Kapitel 10.

Insgesamt unterscheidet man drei mögliche Arten des Zählens, die man als Permutation, Variation und als Kombination bezeichnet. Eine Anordnung von *n* Elementen einer Menge, wobei die Reihenfolge wesentlich ist, wird Permutation (lat. *permutare* = vertauschen) genannt. Unter einer Variation (lat. *variare* = verändern) versteht man eine Auswahl von Elementen, bei denen die Reihenfolge der Anordnung wesentlich ist, während bei einer Kombination (lat. *combinare* = verbinden) die Reihenfolge der ausgewählten Elemente keine Rolle spielt. Somit sind Variationen geordnete Stichproben, Kombinationen dagegen ungeordnete Stichproben.

Im Zuge dieses Kapitels treten erstmals die Stirlingsche Formel sowie der Binomialkoeffizient auf, die im weiteren Verlauf meiner Arbeit noch benötigt werden. Ich werde hier den Zusammenhang des Binomialkoeffizienten mit dem sogenannten Pascalschen Dreieck darlegen. Die Stirlingsche Formel wird später dann in Kapitel 7.3 beim Beweis des Grenzwertsatzes von De Moivre–Laplace auftauchen (Beweis von Lemma 7.13). Der Binomialkoeffizient spielt bei kombinatorischen Überlegungen eine entscheidende Rolle. Er wird uns dann noch einmal in Kapitel 7.2 bei der Binomialverteilung begegnen.

Ein schönes Anwendungsbeispiel für kombinatorische Überlegungen stellt das sogenannte Stimmzettelproblem dar, das ich am Ende dieses Kapitels besprechen werde. Hierbei geht es um die Berechnung der Wahrscheinlichkeit, dass der Wahlgewinner während der gesamten Stimmauszählung in Führung lag.

2.1 Permutationen

2.1.1 Permutationen ohne Wiederholung

Definition 2.1 *n-Faktorielle*

Das Produkt aller natürlichen Zahlen von 1 bis n bezeichnet man mit $n!$ (sprich „n-Faktorielle" bzw. „n-Fakultät")

$$\prod_{i=1}^{n} i = n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 3 \cdot 2 \cdot 1, n \in \mathbb{N}$$

Weiters gilt:

$$0! = 1.$$

Rekursive Darstellung von n-Faktorielle:

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

Proposition 2.2 *Anordnung von n Elementen*

Man kann n verschiedene Elemente auf

$$n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 3 \cdot 2 \cdot 1$$

verschiedene Arten anordnen, wobei die Reihenfolge wesentlich ist.

Beweis:

Für das erste Element gibt es n mögliche Positionen, für das nächste Element verbleiben dann noch $(n-1)$ Positionen, für das dritte Element gibt es noch $(n-2)$ Positionen, usw. Für das letzte Element bleibt dann schließlich nur noch eine Position übrig.

□

Beispiel 2.3 *Ministersitze*

Im österreichischen Parlament gibt es eine eigene Sitzreihe für die zwölf Minister. Auf wie viele unterschiedlichen Arten können die Sitzplätze von den Ministern eingenommen werden?

Es gibt

$$12! = 12 \cdot 11 \cdot 10 \cdot \ldots \cdot 3 \cdot 2 \cdot 1 = 479\,001\,600$$

verschiedene Sitzanordnungen.

Man sieht, dass $n!$ selbst für kleine n sehr rasch recht große Werte annimmt, die ohne Computer nur recht mühsam berechnet werden können. Um $n!$ auch für große n problemlos berechnen zu können, verwendet man die sogenannte Stirlingsche Formel, mit der sich die benötigen Rechenschritte stark reduzieren lassen.

DEFINITION 2.4 *asymptotisch gleich*

Seien a_n und b_n zwei reelle Folgen positiver Zahlen. Dann nennt man a_n und b_n *asymptotisch gleich* (in Symbolen $a_n \sim b_n$), wenn

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 1$$

LEMMA 2.5 *Stirlingsche Formel*

Es existiert die folgende Näherung für *n*-Fakultät

$$n! \sim C \cdot n^{n+\frac{1}{2}} \cdot e^{-n}$$

bzw.

$$\lim_{n\to\infty} \frac{n!}{C \cdot n^{n+\frac{1}{2}} \cdot e^{-n}} = 1$$

wobei

$$C = \sqrt{2\pi}\,.$$

BEWEIS:

Aufgrund der Stetigkeit der Exponentialfunktion und des Logarithmus sind die beiden folgenden Aussagen äquivalent

$$\lim_{n\to\infty} \frac{n!}{C \cdot n^{n+\frac{1}{2}} \cdot e^{-n}} = 1 \Leftrightarrow d_n := \log n! - \left(n + \frac{1}{2}\right) \cdot \log n + n \underset{n\to\infty}{\to} C$$

Unser Interesse gilt nun der Differenzen-Folge $d_n - d_{n+1}$

$$\begin{aligned} d_n - d_{n+1} &= \log n! - \left(n + \frac{1}{2}\right) \cdot \log n \cancel{+n} - \\ &\quad - \log(n+1)! + \left(n + \frac{1}{2} + 1\right) \cdot \log(n+1) \cancel{-n} - 1 = \\ &= \underbrace{\log \frac{n! \cdot (n+1)}{(n+1)!}}_{=0} + \left(n + \frac{1}{2}\right) \cdot \log \frac{n+1}{n} - 1 = \\ &= (2n+1) \cdot \frac{1}{2} \cdot \log \frac{n+1}{n} - 1 \end{aligned}$$

Elementare Umformungen führen zu

$$\frac{n+1}{n} = \frac{2n+2}{2n} = \frac{\frac{2n+1+1}{2n+1}}{\frac{2n+1-1}{2n+1}} = \frac{1 + \frac{1}{2n+1}}{1 - \frac{1}{2n+1}}$$

Wir wenden nun die Potenzreihenenwicklung von

$$\frac{1}{2}\log\frac{1+x}{1-x}=x+\frac{1}{3}x^3+\frac{1}{5}x^5+\dots$$

für $-1<x<1$ auf die Differenzenfolge an und erhalten

$$d_n-d_{n+1}=(2n+1)\cdot\frac{1}{2}\cdot\log\frac{n+1}{n}-1=(2n+1)\cdot\frac{1}{2}\cdot\log\frac{1+\frac{1}{2n+1}}{1-\frac{1}{2n+1}}-1=$$

$$=(2n+1)\cdot\left[\frac{1}{2n+1}+\frac{1}{3}\cdot\left(\frac{1}{2n+1}\right)^3+\frac{1}{5}\cdot\left(\frac{1}{2n+1}\right)^5+\dots\right]-1=$$

$$=\frac{1}{3}\cdot\left(\frac{1}{2n+1}\right)^2+\frac{1}{5}\cdot\left(\frac{1}{2n+1}\right)^4+\dots>0$$

Folglich muss d_n streng monoton fallend sein. Wir erhalten die folgende Abschätzung

$$0<d_n-d_{n+1}<\frac{1}{3}\cdot\left[\left(\frac{1}{2n+1}\right)^2+\left(\frac{1}{2n+1}\right)^4+\left(\frac{1}{2n+1}\right)^6+\dots\right]$$

Nach kurzer Umformung erkennt man, dass es sich bei der oberen Grenze der Differenzenfolge um eine geometrische Reihe mit $q=\left(\frac{1}{2n+1}\right)^2$ handelt:

$$\frac{1}{3}\cdot\left[\left(\frac{1}{2n+1}\right)^2+\left(\frac{1}{2n+1}\right)^4+\left(\frac{1}{2n+1}\right)^6+\dots\right]=$$

$$=\frac{1}{3}\cdot\left(\frac{1}{2n+1}\right)^2\cdot\left[1+\left(\frac{1}{2n+1}\right)^2+\left(\frac{1}{2n+1}\right)^4+\dots\right]\overset{\text{geometrische}}{\underset{\text{Reihe}}{=}}$$

$$=\frac{1}{3}\cdot\left(\frac{1}{2n+1}\right)^2\cdot\frac{1}{1-\left(\frac{1}{2n+1}\right)^2}=$$

$$=\frac{1}{3}\cdot\left(\frac{1}{2n+1}\right)^2\cdot\frac{1}{\left(\frac{2n+1}{2n+1}\right)^2-\left(\frac{1}{2n+1}\right)^2}=$$

$$=\frac{1}{3}\cdot\cancel{\left(\frac{1}{2n+1}\right)^2}\cdot\frac{1}{\cancel{\left(\frac{1}{2n+1}\right)^2}\cdot\left[(2n+1)^2-1\right]}=$$

$$=\frac{1}{3}\cdot\frac{1}{4n^2+4n}=\frac{1}{12}\cdot\frac{n+1-n}{n\cdot(n+1)}=$$

$$=\frac{1}{12}\cdot\left(\frac{\cancel{n+1}}{n\cdot\cancel{(n+1)}}-\frac{\cancel{n}}{\cancel{n}\cdot(n+1)}\right)=\frac{1}{12}\cdot\left(\frac{1}{n}-\frac{1}{n+1}\right)$$

Somit gilt

$$0 < d_n - d_{n+1} < \frac{1}{12} \cdot \left(\frac{1}{n} - \frac{1}{n+1} \right)$$

und

$$d_n - \frac{1}{12n}$$

ist streng monoton steigend.

Für jedes beliebige n gilt daher

$$d_1 - \frac{1}{12} \leq d_n - \frac{1}{12n} < d_n$$

Deshalb ist $d_1 - \frac{1}{12}$ eine untere Schranke von d_n. Also ist d_n monoton fallend und nach unten beschränkt. Somit konvergiert d_n und der Grenzwert existiert mit der Eigenschaft, dass

$$\lim_{n\to\infty} d_n = \lim_{n\to\infty} \left(d_n - \frac{1}{12n} \right) = C$$

Nun muss noch gezeigt werden, dass $C = \sqrt{2\pi}$ gilt. Üblicherweise beweist man dies mit Hilfe des Wallischen Produktes $\frac{\pi}{2} = \lim_{k\to\infty} \frac{2^2 \cdot 4^2 \cdots (2k)^2}{1^2 \cdot 3^2 \cdots (2k-1)^2} \cdot \frac{1}{2k}$.

Ein Beweis findet sich in jedem gängigen Analysis-Lehrbuch. Ich werde jedoch später in Kapitel 7.3 im Zuge des Beweises des Satzes von De Moivre-Laplace (Satz 7.14) einen alternativen Beweis geben, der auf der Ungleichung von Tschebyschew aus Kapitel 5.4 sowie auf der Standardnormalverteilung aus Kapitel 6.2. beruht.

□

Schon für kleine *n* ist die Stirlingsche Formel eine recht gute Näherung, wie die folgende Tabelle zeigt, da der Quotient von *n*! und der Stirlingschen Formel nahe bei 1 liegt:

n	exakter Wert	Näherungswert	Quotient
2	2	1,919	1,042
3	6	5,836	1,028
4	24	23,506	1,021

TABELLE 3 Vergleich von *n*! und der Stirlingschen Formel

BEISPIEL 2.6 *EU-Flaggen*

Im Jahr 2006 gab es 25 EU-Mitgliedsstaaten: Belgien, Dänemark, Deutschland, Estland, Finnland, Frankreich, Griechenland, Irland, Italien, Lettland, Litauen, Luxemburg, Malta, Niederlande, Polen, Österreich, Portugal, Schweden, Slowakei, Slowenien, Spanien, Tschechische Republik, Ungarn, Vereinigtes Königreich und Zypern. Bei einem Treffen der EU-Kommissäre sollen die Flaggen aller Mitgliedsstaaten sowie die EU-Flagge aufgestellt werden. Wie viele mögliche Anordnungen gibt es?

Es gibt

$$26! = 26 \cdot 25 \cdot 24 \cdot \ldots \cdot 3 \cdot 2 \cdot 1 = 403\,291\,461\,126\,605\,635\,584\,000\,000$$

Möglichkeiten um 26 Flaggen aufzustellen.

Mit Hilfe der Stirlingschen Näherungsformel erhält man

$$26! \approx \left(\frac{26}{e}\right)26 \cdot \sqrt{52\pi} = 402\,000\,993\,000\,000\,000\,000\,000\,000$$

Für den Quotient dieser beiden Werte gilt

$$\frac{403\,291\,461\,126\,605\,635\,584\,000\,000}{402\,000\,993\,000\,000\,000\,000\,000\,000} \cong 1{,}003$$

BEISPIEL 2.7 *Kongress*

An einem politischen Kongress nehmen 4 Nationen teil: 2 Österreicher, 5 Deutsche, 4 Briten und 3 Franzosen.

a) Auf wie viele Arten können die Kongressteilnehmer in einer Reihe sitzen, wenn Personen mit gleicher Nationalität nebeneinander sitzen sollen?

b) Jetzt findet das Meeting an einem runden Tisch statt. Wie viele verschiedene Sitzordnungen gibt es dieses Mal, wenn erneut Personen mit gleicher Nationalität nicht getrennt werden?

a) Die vier Nationalitäten lassen sich auf 4! verschiedene Arten anordnen. Nun muss man noch zusätzlich die unterschiedlichen möglichen Sitzanordnungen der einzelnen Landsleute berücksichtigen: Die Österreicher haben 2!, die Deutschen 5!, die Briten 4! und die Franzosen 3! mögliche Plätze innerhalb ihrer Fraktion. Insgesamt erhält man also $4! \cdot 2! \cdot 5! \cdot 4! \cdot 3! = 829\,440$ Möglichkeiten.

b) Hierbei handelt es sich um eine so genannte Ringpermutation. Die Abordnung einer Nation kann sich zusammen irgendwo hinsetzen. Für die restlichen drei Nationen verbleiben somit nur noch 3! Möglichkeiten. Da für die Sitzordnungen der Personen innerhalb einer Nation dasselbe gilt, wie in a), erhält man dieses Mal als Ergebnis $3! \cdot 2! \cdot 5! \cdot 4! \cdot 3! = 207\,360$.

2.1.2 PERMUTATIONEN MIT WIEDERHOLUNG

Bei einer Permutation ohne Wiederholung tritt jedes Element einmal auf, bei einer Permutation mit Wiederholung treten einige Objekte mehrmals auf.

PROPOSITION 2.8 *Multinomialkoeffizient*

Von n Objekten seien $k_1, k_2, ..., k_r$ gleich. Somit lässt sich n als Summe der einzelnen Teilgruppen darstellen

$$k_1 + k_2 + ... + k_r = k$$

Man kann nun diese n Elemente auf

$$\frac{n!}{k_1! \cdot k_2! \cdot ... \cdot k_r!}$$

verschiedene Arten anordnen, wobei die Reihenfolge wesentlich ist. Dieser Ausdruck wird als Multinomialkoeffizient bezeichnet.

BEWEIS:

Sei x die gesuchte Anzahl aller Anordnungsmöglichkeiten. Insgesamt können n Elemente auf $n!$ Arten angeordnet werden. Jede Teilgruppe k_i kann ihrerseits in jeder der möglichen Anordnungen von x ebenfalls auf $k_i!$-Arten angeordnet werden. Somit erhält man

$$x \cdot k_1! \cdot k_2! \cdot ... \cdot k_r! = n!$$

bzw.

$$x = \frac{n!}{k_1! \cdot k_2! \cdot ... \cdot k_r!}$$

□

BEISPIEL 2.9 *Wahlkampf*

Für einen Wahlkampf werden einzelne Teams gebildet und in verschiedene Städte geschickt. Wie viele Möglichkeiten gibt es, wenn 15 Personen in drei gleich große Teams aufgeteilt werden sollen?

Jedes Team besteht aus 5 Personen. Somit gibt es

$$\frac{15!}{5! \cdot 5! \cdot 5!} = 756756$$

verschiedene Team-Zusammenstellungen.

2.2 Variationen

2.2.1 Variationen mit Wiederholung

Proposition 2.10 *Variationen mit Wiederholung*

Aus n Objekten werden k Stück ausgewählt, wobei die Reihenfolge wesentlich ist. Dann gibt es beim Ziehen mit Zurücklegen

$$n^k$$

Auswahlmöglichkeiten.

Beweis:

Es gibt n Möglichkeiten, das erste Objekt zu ziehen. Da es sich um ein Ziehen mit Zurücklegen handelt, gibt es ebenfalls n Möglichkeiten für das zweite Element. Auch für jedes folgende Element gibt es n Möglichkeiten. Daher gibt es insgesamt

$$\underbrace{n \cdot n \cdot \ldots \cdot n}_{k-\text{mal}} = n^k$$

Möglichkeiten.

□

Beispiel 2.11 *Fragebogen*

Mit Hilfe eines Fragebogens soll die aktuelle Beliebtheit von 15 Politikern herausgefunden werden. Zu jeder Frage gibt es immer nur zwei Antwortmöglichkeiten, nämlich „sympathisch“ und „unsympathisch“. Wie viele Möglichkeiten gibt es, diesen Fragebogen auszufüllen?

Insgesamt kann man den Fragebogen auf

$$2^{15} = 32768$$

verschiedene Arten ausfüllen.

2.2.2 Variationen ohne Wiederholung

Proposition 2.12 *Variationen ohne Wiederholung*

Aus n Objekten werden k Stück ausgewählt, wobei die Reihenfolge wesentlich ist. Dann gibt es beim Ziehen ohne Zurücklegen

$$n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}, 1 \le k \le n$$

Auswahlmöglichkeiten.

BEWEIS:

Es gibt n Möglichkeiten, das erste Objekt zu ziehen. Da es sich um ein Ziehen ohne Zurücklegen handelt, gibt es für das zweite Element nur noch $(n-1)$ Möglichkeiten, usw. Für das k-te Element gibt es schlussendlich nur noch $(n-k+1)$ Möglichkeiten. Daher gibt es insgesamt

$$n\cdot(n-1)\cdot(n-2)\cdot\ldots\cdot(n-k+1)=$$

$$=\frac{\overbrace{n\cdot(n-1)\cdot(n-2)\cdot\ldots\cdot(n-k+1)\cdot(n-k)\cdot(n-k-1)\cdot\ldots\cdot3\cdot2\cdot1}^{=n!}}{\underbrace{(n-k)\cdot(n-k-1)\cdot\ldots\cdot3\cdot2\cdot1}_{=(n-k)!}}=\frac{n!}{(n-k)!}$$

Möglichkeiten.

□

BEISPIEL 2.13 *Auswahl von Bewerbern*

Aus zehn Bewerbern sollen drei ausgewählt und in der Reihenfolge ihrer Eignung angegeben werden. Wieviele Möglichkeiten gibt es?

Aus den 10 Bewerbern können auf

$$\frac{10!}{(10-3)!}=\frac{\overset{10\cdot9\cdot8}{\cancel{10!}}}{\cancel{7!}}=720$$

Arten drei Bewerber ausgewählt werden.

PROPOSITION 2.14 *Spezialfall von Proposition 2.12*

Falls $k=n$ ist, so handelt es sich bei diesem Spezialfall um eine Permutation ohne Wiederholung.

BEWEIS:

Für eine Permutation ohne Wiederholung von n Elementen gilt:

$$\frac{n!}{(n-k)!}=\frac{n!}{(n-n)!}=\frac{n!}{0!}=n!$$

□

2.3 Kombinationen

2.3.1 Binomialkoeffizient und binomischer Lehrsatz

Definition 2.15 *Binomialkoeffizient*

Der Binomialkoeffizient wird definiert als

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!},\ 0 \le k \le n,\ k,n \in \mathbb{N}.$$

Das Symbol

$$\binom{n}{k}$$

wird gelesen als *„n über k"*.

Proposition 2.16 *Spezialfall von Proposition 2.8*

Der Binomialkoeffizient

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

ist ein Spezialfall des Multinomialkoeffizienten

$$\frac{n!}{k_1! \cdot k_2! \cdot \ldots \cdot k_r!}.$$

Beweis:

Sei $r = 2$. Setzt man $k_1 = k$ und $k_2 = n - k_1 = n - k$, so erhält man

$$\frac{n!}{k_1! \cdot k_2!} = \frac{n!}{k! \cdot (n-k)!} = \binom{n}{k}$$

□

Lemma 2.17 *Additionssatz*

$$\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}$$

Beweis:

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

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

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

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

□

Die Binomialkoeffizienten stehen im engen Zusammenhang zu dem sogenannten *Pascalschen Dreieck*, das durch die folgenden zwei Eigenschaften charakterisiert wird:

- ♠ Die erste und letzte Zahl einer jeden Zeile ist 1.
- ♠ Jede andere Zahl erhält man als Summe der beiden darüberliegenden Zahlen. (Diese Eigenschaft entspricht dem Additionssatz 2.17.)

Die Bestimmung der Binomialkoeffizienten ist nun sehr leicht:

- ♠ n ist die Anzahl der Zeilen, von oben nach unten gezählt und bei 0 beginnend.
- ♠ k ist der Faktor in jeder Zeile, den man erhält, wenn man von links nach rechts zählt und erneut bei 0 beginnt.

n	Pascalsches Dreieck	Binomialkoeffizienten	Zeilen-summe
0	1	$\binom{0}{0}$	2^0
1	1 1	$\binom{1}{0}$ $\binom{1}{1}$	2^1
2	1 2 1	$\binom{2}{0}$ $\binom{2}{1}$ $\binom{2}{2}$	2^2
3	1 3 3 1	$\binom{3}{0}$ $\binom{3}{1}$ $\binom{3}{2}$ $\binom{3}{3}$	2^3

TABELLE 4 Zusammenhang zwischen dem Pascalschem Dreieck und den Binomialkoeffizienten

SATZ 2.18 *Der Binomische Lehrsatz* $\left(n \in \mathbb{N}, a,b \in \mathbb{R}^+\right)$

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k =$$

$$= \binom{n}{0} a^n + \binom{n}{1} a^{n-1} b + \binom{n}{2} a^{n-2} b^2 + \ldots + \binom{n}{n-1} a b^{n-1} + \binom{n}{n} b^n$$

BEWEIS: *Vollständige Induktion*

Induktionsanfang: $n = 0$

$$\underbrace{(a+b)^0}_{=1} = \underbrace{\sum_{k=0}^{0} \binom{0}{k} a^{0-k} b^k}_{=1}$$

Induktionsannahme:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

Induktionsschritt:

$$(a+b)^{n+1} = (a+b)\cdot(a+b)^n \underset{\text{annahme}}{\overset{\text{Induktions-}}{=}} (a+b)\cdot\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k \underset{\text{multiplizieren}}{\overset{\text{aus-}}{=}}$$

$$= \underbrace{\sum_{k=0}^{n}\binom{n}{k}a^{n-k+1}b^k}_{=a^{n+1}+\sum_{k=1}^{n}\binom{n}{k}a^{n-k+1}b^k} + \underbrace{\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^{k+1}}_{=\sum_{k=0}^{n-1}\binom{n}{k}a^{n-k}b^{k+1}+b^{n+1}} \underset{\text{verschiebung}}{\overset{\text{Index-}}{=}}$$

$$= a^{n+1} + \sum_{k=1}^{n}\binom{n}{k}a^{n-k+1}b^k + \sum_{k=1}^{n}\binom{n}{k-1}a^{n-k+1}b^k + b^{n+1} =$$

$$= a^{n+1} + \sum_{k=1}^{n}\left[\binom{n}{k}+\binom{n}{k-1}\right]a^{n-k+1}b^k + b^{n+1} \underset{2.17}{\overset{\text{LEMMA}}{=}}$$

$$= a^{n+1} + \sum_{k=1}^{n}\binom{n+1}{k}a^{n-k+1}b^k + b^{n+1} =$$

$$= \sum_{k=0}^{n}\binom{n+1}{k}a^{n-k+1}b^k + b^{n+1} = \sum_{k=0}^{n+1}\binom{n+1}{k}a^{n-k+1}b^k$$

□

2.3.2 KOMBINATIONEN OHNE WIEDERHOLUNG

PROPOSITION 2.19 *Kombinationen ohne Wiederholung*

Aus n Objekten werden k Stück ausgewählt, wobei die Reihenfolge ohne Bedeutung ist. Dann gibt es beim Ziehen ohne Zurücklegen

$$\binom{n}{k} = \frac{n!}{k!\cdot(n-k)!}, 1 \le k \le n$$

Auswahlmöglichkeiten.

BEWEIS:

Ausgangspunkt ist die Formel für die Variation ohne Wiederholung:

$$\frac{n!}{(n-k)!}$$

Bei der Variation kommt es ja bekanntlich auf die Anordnung der einzelnen Elemente an. Daher kommen dort auch alle Anordnungen der k Elemente vor. Diese k Elemente lassen sich auf *k!* Möglichkeiten anordnen, da hier eine Permutation vorliegt. Dividiert man nun die Formel für die Variation ohne Wiederholung durch $k!$ so erhält man die Anzahl aller Möglichkeiten für ungeordnete Stichproben.

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

□

BEISPIEL 2.20 *Delegation*

Für eine Delegation sollen aus einer Gruppe von 25 Personen 5 ausgewählt werden. Wie viele Möglichkeiten gibt es?

Aus den 25 Personen können auf

$$\binom{25}{5} = \frac{\cancelto{25\cdot 24\cdot 23\cdot 22\cdot 21}{25!}}{5!\cdot \cancel{20!}} = \frac{\cancelto{5}{25}\cdot \cancel{24}\cdot 23\cdot 22\cdot 21}{\cancel{5}\cdot \cancel{4\cdot 3\cdot 2}\cdot 1} = 53130$$

Arten fünf Personen ausgewählt werden.

BEISPIEL 2.21 *Untersuchungsausschuss*

Es soll ein Untersuchungsausschuss aus 3 Männern und 2 Frauen gebildet werden. Wie viele Möglichkeiten gibt es, wenn insgesamt 13 Männer und 12 Frauen zur Verfügung stehen?

Aus den 13 Männern können

$$\binom{13}{3} = \frac{\cancelto{13\cdot 12\cdot 11}{13!}}{3!\cdot \cancel{10!}} = \frac{13\cdot \cancelto{2}{12}\cdot 11}{\cancel{6}} = 286$$

Männer ausgewählt werden.

Analog können aus den 12 Frauen

$$\binom{12}{2} = \frac{\cancelto{12\cdot 11}{12!}}{2!\cdot \cancel{10!}} = \frac{\cancelto{6}{12}\cdot 11}{\cancel{2}} = 66$$

Frauen ausgewählt werden.

Insgesamt gibt es also

$$\binom{13}{3}\cdot\binom{12}{2} = 286\cdot 66 = 18876$$

Möglichkeiten einen Ausschuss zu bilden.

2.3.3 KOMBINATIONEN MIT WIEDERHOLUNG

PROPOSITION 2.22 *Kombinationen mit Wiederholung*

Aus n Objekten werden k Stück ausgewählt, wobei die Reihenfolge ohne Bedeutung ist. Dann gibt es beim Ziehen mit Zurücklegen

$$\binom{n+k-1}{k}, 1 \le k \le n$$

Auswahlmöglichkeiten.

BEWEIS:

Es werden k Objekte ausgewählt, wobei die Reihenfolge egal ist. Jedes Objekt darf mehrmals ausgewählt werden. Das 1.Objekt werde k_1-mal ausgewählt, das zweite Objekt k_2-mal, usw. Das n-te Objekt werde k_n-mal ausgewählt. Die Summe aller Objekte muss natürlich k ergeben.

$$k_1 + k_2 + \ldots + k_n = k,\ 0 \le k_i \le k$$

Betrachten wir nun k Objekte, dargestellt als Sterne, die mit Hilfe von Strichen in n Gruppen eingeteilt werden können.

$$\underbrace{*\,*\,*\,*\,*}_{k_1-\text{mal Objekt } 1} \Big| \underbrace{*\,*\,*}_{k_2-\text{mal Objekt } 2} \Big| \ldots \Big| \underbrace{*\,*\,*\,*\,*\,*}_{k_n-\text{mal Objekt n}}$$

Die k Sterne und die $n-1$ Striche sind in einer Reihe angeordnet. Zusammen nehmen sie $n+k-1$ Positionen ein. Die Frage nach der Anzahl aller möglichen Anordnungen ist gleichbedeutend mit der Frage nach der Anzahl der Möglichkeiten, aus $n+k-1$ möglichen Plätzen für Sterne und Striche $n-1$ Positionen für die Striche auszuwählen. Wir erhalten

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

□

BEISPIEL 2.23 *Stimmabgabe*

Bei einer Wahl stellen sich sechs KandidatInnen. Ein Wähler hat drei Stimmen zur Verfügung, die er beliebig verteilen darf. Es ist auch erlaubt, einem/einer KandidatIn mehrere Stimmen zu geben. Wie viele Möglichkeiten gibt es, die Stimmen abzugeben?

Die Reihenfolge der getätigten Stimmen spielt keine Rolle. Daher kann man auf

$$\binom{6+(3-1)}{3} = \binom{8}{3} = \frac{\overset{8\cdot 7\cdot 6}{\cancel{8!}}}{3! \cdot \cancel{5!}} = \frac{8 \cdot 7 \cdot \cancel{6}}{\cancel{6}} = 56$$

Arten die Stimmen verteilen.

2.4 DAS STIMMZETTELPROBLEM[3]

Bei einer (Stich)Wahl treten zwei KandidatInnen *A* und *B* gegeneinander an. Um den Sieger ermitteln zu können, müssen alle gültigen Stimmzettel der Reihe nach ausgezählt werden, so dass zu jedem beliebigen Zeitpunkt feststeht, welcher der beiden KandidatInnen im Moment in Führung liegt.

KandidatIn *A* hat nun die Wahl für sich entscheiden können. Wie groß ist dann die Wahrscheinlichkeit, dass *A* während der gesamten Stimmauszählung in Führung lag?

BEISPIEL 2.24 *ein konkretes Beispiel*

KandidatIn *A* habe fünf Stimmen und KandidatIn *B* nur zwei Stimmen erhalten. Der allgemeine Fall folgt danach anschließend ganz analog.

Man kann sich den Ablauf der Stimmauszählung leicht graphisch überlegen. Dazu betrachtet man die einzelnen Punkte im Koordinatensystem, wobei man immer von einem Gitterpunkt (x,y) nach $(x+1,y)$ geht, wenn KandidatIn *A* eine zusätzliche Stimme erhalten hat, bzw. von (x,y) nach $(x,y+1)$, wenn der Stimmzettel für KandidatIn *B* abgegeben wurde.

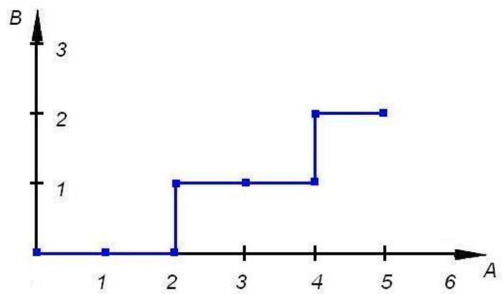


ABBILDUNG 2 Graphische Darstellung des Pfades AABAABA im Gitter $\mathbb{N}^2$ von (0,0) nach (5,2)

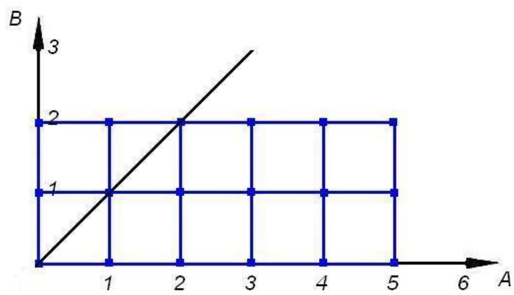


ABBILDUNG 3 Graphische Darstellung aller möglichen Wege sowie der Geraden $g: y = x$

Die Menge aller Wege von (0,0) nach (5,2) ist dann der Ergebnisraum Ω, der nun berechnet wird.

---

[3] DEHLING Herold / Beate HAUPT, Einführung in die Wahrscheinlichkeitstheorie und Statistik, Berlin / Heidelberg [2]2004, S.37ff.

Insgesamt gibt es die folgenden 21 verschiedenen Möglichkeiten der Stimmauszählung:

BBAAAAA	ABBAAAA	AABBAAA	AAABBAA
BABAAAA	ABABAAA	AABABAA	AAABABA
BAABAAA	ABAABAA	AABAABA	AAABAAB
BAAABAA	ABAAABA	AABAAAB	
BAAAABA	ABAAAAB		
BAAAAAB		AAAABBA	AAAAABB
		AAAABAB	

Mathematisch gesehen handelt es sich hierbei um eine Kombination ohne Wiederholung, da jede Stimme nur einmal ausgezählt wird und die einzelnen *A*s und *B*s nicht zu unterscheiden sind.

Insgesamt möchte man also fünf (bzw. zwei) Stimmen auf sieben möglichen Plätzen anordnen.

$$|\Omega| = \binom{7}{5} = \frac{\overset{7\cdot 6}{\cancel{7!}}}{\cancel{5!}\cdot 2!} = \frac{7\cdot \overset{3}{\cancel{6}}}{\cancel{2}} = 21$$

Man kann die folgenden drei disjunkten Ereignisse von Ω definieren:

- ♠ E_1: KandidatIn *B* erhält die erste Stimme. $\Rightarrow$ Menge aller Wege, die durch den Punkt (0,1) gehen.
- ♠ E_2: KandidatIn *A* erhält die erste Stimme, aber wird zu bestimmten Zeitpunkten von KandidatIn B eingeholt. $\Rightarrow$ Menge aller Wege, die durch den Punkte (1,0) gehen, aber nicht gänzlich unterhalb der Geraden $g: y = x$ verlaufen.
- ♠ E_3: KandidatIn *A* liegt zu jedem Zeitpunkt bei der Stimmauszählung voran. $\Rightarrow$ Menge aller Wege, die stets unterhalb der Geraden $g: y = x$ verlaufen.

Beschäftigen wir uns nun etwas genauer mit diesen Ereignissen. Kennen wir nämlich E_1 und E_2, so können wir problemlos E_3 berechnen, da

$$|E_3| = |\Omega| - |E_1| - |E_2|$$

gilt.

Insgesamt erhält KandidatIn *B* sechsmal zuerst eine Stimme: Es sind dies die folgenden Pfade

BBAAAAA	BAAABAA
BABAAAA	BAAAABA
BAABAAA	BAAAAAB

Da die erste Position fix vorgegeben ist, gilt es nun nur noch fünf (bzw. eine) Stimme(n) auf sechs mögliche Positionen aufzuteilen

$$|E_1| = \binom{6}{5} = \frac{\overset{6}{\cancel{6!}}}{\cancel{5!}\cdot 1!} = 6$$

Auch E_2 besteht aus sechs Möglichkeiten, und zwar

ABBAAAA	ABAAABA
ABABAAA	ABAAAAB
ABAABAA	AABBAAA

Dass die beiden Mengen E_1 und E_2 gleich groß sind, ist kein Zufall, sondern lässt sich leicht mit Hilfe des sogenannten Reflexionsprinzip begründen: Ein beliebiger Punkt aus E_1 beginnt bei (0,1) und muss auf seinem Weg zu (5,2) die Gerade $g: y = x$ mindestens einmal schneiden. Beim ersten Schnittpunkt wird nun dieser Pfad an der Geraden gespiegelt. Dieser Weg enthält den Punkt (1,0) und liegt offensichtlich nicht ganz unterhalb der Geraden. Somit ist dieser Pfad Bestandteil von E_2.

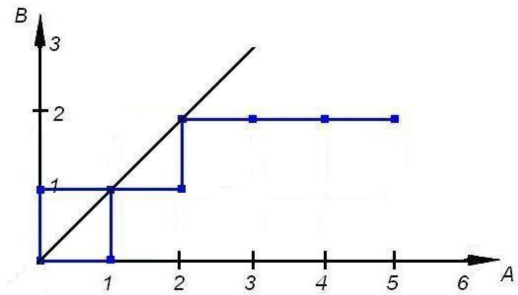


ABBILDUNG 4 Spiegelt man den Pfad BAABAAA an der Geraden $g: y = x$, so erhält man den Weg ABABAAA

Diese Spiegelung kann man sich bei allen sechs Ereignissen von E_1 überlegen.

BBAAAAA $\Leftrightarrow$ AABBAAA
BABAAAA $\Leftrightarrow$ ABBAAAA
BAABAAA $\Leftrightarrow$ ABABAAA
BAAABAA $\Leftrightarrow$ ABAABAA
BAAAABA $\Leftrightarrow$ ABAAABA
BAAAAAB $\Leftrightarrow$ ABAAAAB

Das Ereignis E_3 setzt sich schließlich aus folgenden neun Pfaden zusammen.

AABABAA	AAABAAB
AABAABA	AAAABBA
AABAAAB	AAAABAB
AAABBAA	AAAAABB
AAABABA	

Rein rechnerisch würde man auf diese Lösung kommen, indem man die Ereignisse E_1 und E_2 von der Gesamtzahl aller möglichen Wege abzieht.

$$\begin{aligned} |E_3| &= |\Omega| - |E_1| - |E_2| \\ &= 21 - 6 - 6 = 9 \end{aligned}$$

Die gesuchte Wahrscheinlichkeit, dass KandidatIn A während der gesamten Stimmauszählung vor KandidatIn B gelegen ist, lautet also

$$P(A \text{ liegt stets voran}) = P(E_3) = \frac{9}{21} = \frac{3}{7} \cong 0{,}4286$$

Anhand dieses anschaulichen Beispiels ist der allgemeine Fall kein Problem mehr.

SATZ 2.25 *Das Stimmzettelproblem*

KandidatIn A habe a Stimmen erhalten, KandidatIn B b Stimmen. Dann beträgt die Wahrscheinlichkeit, dass KandidatIn A während der gesamten Stimmauszähung in Führung liegt

$$P(A \text{ liegt stets voran}) = \frac{a-b}{a+b}.$$

BEWEIS:

$$|\Omega| = \binom{a+b}{a}$$

$$|E_1| = |E_2| = \binom{a+b-1}{a}$$

$$|E_3| = |\Omega| - |E_1| - |E_2|$$
$$= \binom{a+b}{a} - 2 \cdot \binom{a+b-1}{a}$$

$$P(E_3) = \frac{\binom{a+b}{a} - 2 \cdot \binom{a+b-1}{a}}{\binom{a+b}{a}} = 1 - 2 \cdot \frac{\binom{a+b-1}{a}}{\binom{a+b}{a}} =$$

$$= 1 - 2 \cdot \frac{\dfrac{(a+b-1)!}{\cancel{a}! \cdot (\cancel{a} + b - 1 - \cancel{a})!}}{\dfrac{(a+b)!}{\cancel{a}! \cdot (\cancel{a} + b - \cancel{a})!}} = 1 - 2 \cdot \frac{\cancel{(a+b-1)!} \cdot \overset{b}{\cancel{b!}}}{\cancel{(b-1)!} \cdot \underset{a+b}{\cancel{(a+b)!}}} =$$

$$= \frac{a+b}{a+b} - 2 \cdot \frac{b}{a+b}$$

Und somit erhält man die gesuchte Wahrscheinlichkeit

$$P(A \text{ liegt stets voran}) = \frac{a-b}{a+b}$$

□

3. Bedingte Wahrscheinlichkeit und Unabhängigkeit

In diesem Abschnitt beschäftige ich mich jetzt mit der Frage, wie sich eine Wahrscheinlichkeit für ein Ereignis *A* berechnen lässt, wenn bekannt ist, dass ein anderes Ereignis *B* bereits eingetreten ist. Dies nennt man bedingte Wahrscheinlichkeit. Wir betrachten Ereignisse, die von anderen Ereignissen beeinflusst werden und solche, wo sich die Wahrscheinlichkeit trotz Zusatzinformation nicht verändert. Wenn das Ereignis *B* keinen Einfluss auf das Eintreten des Ereignisses *A* hat, so spricht man von unabhängigen Ereignissen.

Die wichtigsten Resultate dieses Kapitels sind der Satz von der totalen Wahrscheinlichkeit und der Satz von Bayes. Beide Formeln werden beim Umgang mit bedingten Wahrscheinlichkeiten benötigt. Sind die bedingten Wahrscheinlichkeiten bekannt und interessiert man sich für die Einzelwahrscheinlichkeiten, so verwendet man den Satz von der totalen Wahrscheinlichkeit. Mit dem Satz von Bayes kann man dagegen die bedingten Wahrscheinlichkeiten berechnen, wenn man die Einzelwahrscheinlichkeiten kennt.

Das Kapitel endet mit dem Simpsonschen Paradoxon, bei dem man bei der Zusammenfassung verschiedener Gruppen ein anderes Resultat erhält, als wenn man die Gruppen einzeln betrachten würde. Grund dafür ist die unterschiedliche Gewichtung der Einzelereignisse. Erstmals setzte sich der Statistiker Edward Hugh Simpson 1951 im „Journal of the Royal Statistical Society“ mit diesem Phänomen auseinander[4], das später dann nach ihm benannt wurde. Ich möchte das Simpsonsche Paradoxon anhand der Zusammensetzung der Geschworenen-Jurys in Neuseeland vorstellen.

[4] SIMPSON, Edward Hugh, The Interpretation of Interaction in Contingency Tables, In: Journal of the Royal Statistical Society Series B, 13(1951), S.238-241

3.1 BEDINGTE WAHRSCHEINLICHKEIT

BEISPIEL 3.1 *Nationalratswahl in Österreich vom 24.November 2002*

	ÖVP	SPÖ	FPÖ	Grüne	Sonstige[5]
Österreich	2076833	1792499	491 328	464980	84005
Wien	261496	373436	67975	129141	21194
Burgenland	81127	87660	12163	9009	1335
Niederösterreich	485351	373179	70208	73177	12762
Oberösterreich	363497	316009	88790	74043	11226
Steiermark	340185	281965	73540	53011	13897
Salzburg	138924	91674	31949	30848	4418
Kärnten	105894	133131	82002	21495	4764
Tirol	203228	95855	39059	45585	8083
Vorarlberg	97131	39590	25642	28671	6326

TABELLE 5 Wahlergebnis der Nationalratswahl in Österreich vom 24.November 2002[6]

Wie groß ist nun die Wahrscheinlichkeit, dass
a) ein beliebiger Österreicher ÖVP gewählt hat?
b) ein Vorarlberger ÖVP gewählt hat?
c) ÖVP gewählt wurde unter der Bedingung, dass wir uns für den Wahlausgang nur in Wien interessieren?

a) $2076833 + 1792499 + 491328 + 464980 + 84005 = 4909645$

$$P(\text{ÖVP}) = \frac{2076833}{4909645} \cong \underline{\underline{42,30\%}}$$

b) $97131 + 39590 + 25642 + 28671 + 6326 = 197360$

$$P(\text{Vorarlberger wählt ÖVP}) = \frac{97131}{197360} \cong \underline{\underline{49,22\%}}$$

c) $261496 + 373436 + 67975 + 129141 + 21194 = 853242$

$$P(\text{Wiener wählt ÖVP}) = \frac{261496}{853242} \cong \underline{\underline{30,65\%}}$$

Die Zusatzinformation Vorarlberg bzw. Wien bewirkt eine Vergrößerung bzw. Verkleinerung des Prozentsatzes. Der Wahlausgang ist also von Bundesland zu Bundesland unterschiedlich. Dies müsste jedoch nicht sein. Es wäre durchaus möglich, dass eine Zusatzinformation den Prozentsatz nicht verändern würde.

---

[5] KPÖ, LIF, Demokraten, CWG und SLP

[6] Bundesministerium für Inneres, Nationalratswahl 2002. Endergebnis und Berechnung nach d`Hondt, http://www.bmi.gv.at/wahlen/nrw02info.asp [31.12.2008]

DEFINITION 3.2 *bedingte Wahrscheinlichkeit*

Sei P eine Wahrscheinlichkeit auf Ω und sei B ein zulässiges Ereignis mit $P(B) \neq 0, B \subseteq \Omega$. Dann nennt man

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

die bedingte Wahrscheinlichkeit von A unter der Bedingung B.

Beispiel 3.1 b) und c) lassen sich auch wie folgt anschreiben

$$P(\text{ÖVP}|\text{Vorarlberg}) = \frac{97131}{197360} \cong \underline{\underline{49{,}22\%}}$$

$$P(\text{ÖVP}|\text{Wien}) = \frac{261496}{853242} \cong \underline{\underline{30{,}65\%}}$$

SATZ 3.3 Die bedingte Wahrscheinlichkeit erfüllt alle Eigenschaften einer Wahrscheinlichkeit. Daher gelten die folgenden Forderungen

(i) $P(\Omega|B) = 1$

(ii) $P(A|B) \geq 0 \ \forall A \subseteq \Omega$

(iii) $P\left(\bigcup_{i=1}^{\infty} A_i \middle| B\right) = \sum_{i=1}^{\infty} P(A_i|B)$, A_i paarweise disjunkt

sogenannte *σ-Additivität*

BEWEIS:

(i)
$$P(\Omega|B) \underset{3.2}{\overset{\text{Def.}}{=}} \frac{P(\Omega \cap B)}{P(B)} \underset{1.1\text{ (vi)}}{\overset{\text{Kor.}}{=}} \frac{P(B)}{P(B)} = 1$$

(ii)
$$P(A|B) \underset{3.2}{\overset{\text{Def.}}{=}} \frac{P(A \cap B)}{P(B)} \underset{14.1\text{ (i)}}{\overset{\text{Def.}}{\geq}} 0$$

(iii)
$$P\left(\bigcup_{i=1}^{\infty} A_i \middle| B\right) \underset{3.2}{\overset{\text{Def.}}{=}} \frac{P\left(\left(\bigcup_{i=1}^{\infty} A_i\right) \cap B\right)}{P(B)} = \frac{P\left(\bigcup_{i=1}^{\infty} (A_i \cap B)\right)}{P(B)} \underset{1.14\text{ (iii)}}{\overset{\text{Def.}}{=}}$$

$$= \frac{\sum_{i=1}^{\infty} P(A_i \cap B)}{P(B)} \underset{3.2}{\overset{\text{Def.}}{=}} \sum_{i=1}^{\infty} P(A_i \cap B|B) = \sum_{i=1}^{\infty} P(A_i|B)$$

□

PROPOSITION 3.4 *Eigenschaften der bedingten Wahrscheinlichkeit*

(i) $P(A|B) = 1$ falls $B \subseteq A$

(ii) $P(A|B) = 0$ falls $A \cap B = \emptyset$

(iii) $P(A|B) = P(C|B)$ falls $A \cap B = B \cap C$

BEWEIS:

(i) $P(A|B) \overset{\text{Def.}}{\underset{3.2}{=}} \dfrac{P(A \cap B)}{P(B)} \overset{B \subseteq A}{=} \dfrac{P(B)}{P(B)} = 1$

(ii) $P(A|B) \overset{\text{Def.}}{\underset{3.2}{=}} \dfrac{P(A \cap B)}{P(B)} \overset{A \cap B = \emptyset}{=} \dfrac{P(\emptyset)}{P(B)} = 0$

(iii) $P(A|B) \overset{\text{Def.}}{\underset{3.2}{=}} \dfrac{P(A \cap B)}{P(B)} \overset{A \cap B = B \cap C}{=} \dfrac{P(B \cap C)}{P(B)} \overset{\text{Def.}}{\underset{3.2}{=}} P(C|B)$

□

SATZ 3.5 *Multiplikationssatz für bedingte Wahrscheinlichkeiten*

(i) $P(A \cap B) = P(A|B) \cdot P(B)$ falls $P(B) \neq 0$

(ii) $P(A \cap B) = P(B|A) \cdot P(A)$ falls $P(A) \neq 0$

BEWEIS:

(i) $P(A|B) \overset{\text{Def.}}{\underset{3.2}{=}} \dfrac{P(A \cap B)}{P(B)} \Leftrightarrow P(A|B) \cdot P(B) = P(A \cap B)$

(ii) $P(B|A) \overset{\text{Def.}}{\underset{3.2}{=}} \dfrac{P(A \cap B)}{P(A)} \Leftrightarrow P(B|A) \cdot P(A) = P(A \cap B)$

□

SATZ 3.6 *Allgemeiner Multiplikationssatz*

Seien $A_1, ..., A_n$ Ereignisse mit $P\left(\bigcap_{i=1}^{n-1} A_i\right) \neq 0$. Dann gilt

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P\left(A_i \,\middle|\, \bigcap_{j=1}^{i-1} A_j\right)$$

BEWEIS:

$$\prod_{i=1}^{n} P\left(A_i \,\middle|\, \bigcap_{j=1}^{i-1} A_j\right) = P(A_1) \cdot P(A_2|A_1) \cdot \ldots \cdot P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}) \overset{\text{Def.}}{\underset{3.2}{=}}$$

$$= \cancel{P(A_1)} \cdot \frac{\cancel{P(A_1 \cap A_2)}}{\cancel{P(A_1)}} \cdot \frac{\cancel{P(A_1 \cap A_2 \cap A_3)}}{\cancel{P(A_1 \cap A_2)}} \cdot \cancel{\ldots} \cdot$$

$$\cancel{\ldots} \cdot \frac{\cancel{P(A_1 \cap A_2 \cap \ldots \cap A_{n-1})}}{\cancel{P(A_1 \cap A_2 \cap \ldots \cap A_{n-2})}} \cdot \frac{P(A_1 \cap A_2 \cap \ldots \cap A_n)}{\cancel{P(A_1 \cap A_2 \cap \ldots \cap A_{n-1})}} =$$

$$= P(A_1 \cap A_2 \cap \ldots \cap A_n) = P\left(\bigcap_{i=1}^{n} A_i\right)$$

□

BEISPIEL 3.7 *Politverdrossenheit*

Angenommen bei der letzten Nationalratswahl hätten 20% der Wahlberechtigten nicht von ihrem Wahlrecht Gebrauch gemacht und 15% wären bei der letzten Gemeinderatswahl nicht wählen gegangen. 10% hätten bei keiner der beiden Wahlen einen Stimmzettel abgegeben. Wie groß ist die Wahrscheinlichkeit, dass dann

a) ein Wähler bei der letzten Gemeinderatswahl nicht zur Wahlurne gegangen ist, unter der Bedingung, dass er auch schon bei der letzten Nationalratswahl nicht gewählt hat?

b) ein Wähler, der bei der letzten Nationalratswahl wählen gegangen ist, dann auch noch bei der letzten Gemeinderatswahl gewählt hat?

Sei *N* das Ereignis, dass ein Wähler bei der Nationalratswahl gewählt hat und N^c, dass er nicht gewählt hat. Analog bezeichne *G* das Ereignis, dass ein Wähler bei der Gemeinderatswahl seine Stimme abgegeben hat und G^c, dass er nicht zur Wahl gegangen ist.

Aus der Angabe folgen unmittelbar diese Wahrscheinlichkeiten

$$P(N) = 1 - P(N^c) = 1 - 0{,}2 = 0{,}8$$

$$P(G) = 1 - P(G^c) = 1 - 0{,}15 = 0{,}85$$

$$P(G^c \cap N^c) \underset{1.1\text{ (viii)}}{\overset{\text{Kor.}}{=}} P\left((G \cup N)^c\right) = 0{,}1$$

a) Wir erhalten

$$P(G^c | N^c) = \frac{P(G^c \cap N^c)}{P(N^c)} = \frac{0{,}1}{0{,}2} = 0{,}5 \mathrel{\hat{=}} \underline{\underline{50\,\%}}$$

b) Es folgt

$$P(G|N) = \frac{P(G \cap N)}{P(N)} \underset{1.17}{\overset{\text{Satz}}{=}} \frac{P(G) + P(N) - P(G \cup N)}{P(N)} =$$

$$= \frac{P(G) + P(N) - \left[1 - P\left((G \cup N)^c\right)\right]}{P(N)} =$$

$$= \frac{0{,}85 + 0{,}8 - (1 - 0{,}1)}{0{,}8} = \frac{0{,}75}{0{,}8} = 0{,}9375 \mathrel{\hat{=}} \underline{\underline{93{,}75\,\%}}$$

3.2 TOTALE WAHRSCHEINLICHKEIT UND SATZ VON BAYES

SATZ 3.8 *Satz von der totalen Wahrscheinlichkeit*

Sei $B_1,...,B_n$ eine disjunkte Zerlegung von Ω (d.h. $\Omega = B_1 \cup ... \cup B_n$ und $B_i \cap B_j = \emptyset$ für $i \neq j$). Sei $P(B_i) > 0 \ \forall i$. Dann gilt für alle Ereignisse $A \subseteq \Omega$

$$P(A) = \sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)$$

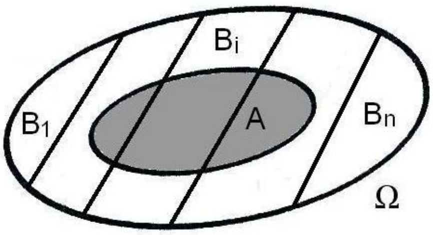


ABBILDUNG 5 Graphische Veranschaulichung des Satzes von der totalen Wahrscheinlichkeit

BEWEIS:

$$\sum_{i=1}^{n} P(A|B_i) \cdot P(B_i) \underset{3.2}{\overset{\text{Def.}}{=}} \sum_{i=1}^{n} \frac{P(A \cap B_i)}{P(B_i)} \cdot P(B_i) \underset{1.16\text{ (ii)}}{\overset{\text{Satz}}{=}} P\left(\bigcup_{i=1}^{n} (A \cap B_i)\right) \underset{1.1\text{ (iii)}}{\overset{\text{KOR.}}{=}}$$

$$= P\left(A \cap \bigcup_{i=1}^{n} B_i\right) = P(A \cap \Omega) \underset{1.1\text{ (vi)}}{\overset{\text{Kor.}}{=}} P(A)$$

□

SATZ 3.9 *Satz von Bayes*

Seien $A, B \in \Omega$ mit $P(A) > 0$. Sei $B_1,...,B_n$ eine disjunkte Zerlegung von Ω mit $P(B_i) > 0 \ \forall i$. Dann gilt für alle Ereignisse $A \subseteq \Omega$

$$P(B_i|A) = \frac{P(A|B_i) \cdot P(B_i)}{\sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)}$$

BEWEIS:

$$P(B_i|A) \underset{3.2}{\overset{\text{Def.}}{=}} \frac{P(B_i \cap A)}{P(A)} \underset{3.5\text{ (i)}}{\overset{\text{Satz}}{=}} \frac{P(A|B_i) \cdot P(B_i)}{P(A)} \underset{3.8}{\overset{\text{Satz}}{=}} \frac{P(A|B_i) \cdot P(B_i)}{\sum_{i=1}^{n} P(A|B_i) \cdot P(B_i)}$$

□

BEISPIEL 3.10 *Nationalratswahl in Österreich vom 24.November 2002*

Beschäftigen wir uns erneut mit der Nationalratswahl in Österreich vom 24.November 2002 (siehe Beispiel 3.1). Wie groß ist nun die Wahrscheinlichkeit, dass

a) ein beliebiger Österreicher SPÖ gewählt hat?

b) eine Person, die SPÖ gewählt hat, aus Wien kommt?

a) $P(\text{SPÖ})$ kann auf zwei Arten berechnet werden

<u>1.Art:</u> Ablesen aus der Tabelle 5

$$P(\text{SPÖ}) = \frac{1792499}{4909645} \cong \underline{\underline{36{,}51\%}}$$

<u>2.Art:</u> Satz von der totalen Wahrscheinlichkeit

$$P(\text{SPÖ}|\text{Wien}) \cdot P(\text{Wien}) = \frac{373436}{\cancel{853242}} \frac{\cancel{853242}}{4909645} \cong 0{,}0761$$

$$P(\text{SPÖ}|\text{Burgenland}) \cdot P(\text{Burgenland}) = \frac{87660}{\cancel{191294}} \frac{\cancel{191294}}{4909645} \cong 0{,}0179$$

$$P(\text{SPÖ}|\text{NÖ}) \cdot P(\text{NÖ}) = \frac{373179}{\cancel{1014677}} \frac{1014677}{4909645} \cong 0{,}0760$$

$$P(\text{SPÖ}|\text{OÖ}) \cdot P(\text{OÖ}) = \frac{316009}{\cancel{853565}} \frac{\cancel{853565}}{4909645} \cong 0{,}0644$$

$$P(\text{SPÖ}|\text{Stmk}) \cdot P(\text{Stmk}) = \frac{281965}{\cancel{762598}} \frac{\cancel{762598}}{4909645} \cong 0{,}0574$$

$$P(\text{SPÖ}|\text{Sbg}) \cdot P(\text{Sbg}) = \frac{91674}{\cancel{297813}} \frac{\cancel{297813}}{4909645} \cong 0{,}0187$$

$$P(\text{SPÖ}|\text{Kärnten}) \cdot P(\text{Kärnten}) = \frac{133131}{\cancel{347286}} \frac{\cancel{347286}}{4909645} \cong 0{,}0271$$

$$P(\text{SPÖ}|\text{Tirol}) \cdot P(\text{Tirol}) = \frac{95855}{\cancel{391810}} \frac{391810}{4909645} \cong 0{,}0195$$

$$P(\text{SPÖ}|\text{Vorarlberg}) \cdot P(\text{Vorarlberg}) = \frac{39590}{\cancel{197360}} \frac{\cancel{197360}}{4909645} \cong 0{,}0081$$

$$P(\text{SPÖ}) = \sum_{k=1}^{9} P(\text{SPÖ}|\text{Bundesland } k) \cdot P(\text{Bundesland } k) =$$

$$= 0{,}0761 + 0{,}0179 + 0{,}0760 + 0{,}0644 + 0{,}0574 +$$

$$+ 0{,}0187 + 0{,}0271 + 0{,}0195 + 0{,}0081 \cong \underline{\underline{36{,}51\%}}$$

b) Mit dem Satz von Bayes erhalten wir

$$P(\text{Wien}|\text{SPÖ}) = \frac{P(\text{SPÖ}|\text{Wien}) \cdot P(\text{Wien})}{\sum_{k=1}^{9} P(\text{SPÖ}|\text{Bundesland } k) \cdot P(\text{Bundesland } k)} =$$

$$= \frac{0{,}0761}{0{,}3651} \cong \underline{\underline{20{,}83\%}}$$

3.3 Unabhängige Ereignisse

Beispiel 3.11 *Nationalratswahl in Österreich vom 24.November 2002*

Was fällt bei einem Vergleich der FPÖ-Wähler bei der *Nationalratswahl in Österreich vom 24.November 2002* österreichweit und in Eichberg, Bezirk Hartberg, Steiermark[7], auf?

	gültige Stimmen	FPÖ Wähler
Österreich	4909645	491328
Eichberg	779	78

Tabelle 6 FPÖ-Wähler österreichweit und in Eichberg bei der Nationalratswahl am 24.November 2002

$$P(\text{FPÖ}) = \frac{491328}{4909645} \cong 10{,}01\%$$

$$P(\text{FPÖ}|\text{Eichberg}) = \frac{78}{779} \cong 10{,}01\%$$

Offensichtlich bewirkt hier die Zusatzinformation Eichberg (nahezu) keine Veränderung des Prozentsatzes. Man spricht dann von unabhängigen Ereignissen.

Definition 3.12 *unabhängige Ereignisse*

Seien A und B beliebige Ereignisse. Dann nennt man A und B *unabhängig*, falls

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

Proposition 3.13 Sei $0 < P(B) < 1$. Dann sind folgende Aussagen äquivalent

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

(ii) $P(A|B) = P(A|B^c)$

(iii) $P(A|B) = P(A)$

Beweis:

(i) → (ii)

$$P(A|B^c) \underset{3.2}{\overset{\text{Def.}}{=}} \frac{P(A \cap B^c)}{P(B^c)} \underset{3.13\,(i)}{\overset{\text{Prop.}}{=}} \frac{P(A) \cdot P(B^c)}{P(B^c)} =$$

$$= P(A) = \frac{P(A) \cdot P(B)}{P(B)} \underset{3.13\,(i)}{\overset{\text{Prop.}}{=}} \frac{P(A \cap B)}{P(B)} \underset{3.2}{\overset{\text{Def.}}{=}} P(A|B)$$

[7] Bundesministerium für Inneres, Nationalratswahl 2002. Endergebnis und Berechnung nach d`Hondt, http://www.bmi.gv.at/wahlen/nrw02info.asp [31.12.2008]

(ii) $\rightarrow$ (iii) $P(A) \underset{3.8}{\overset{\text{Satz}}{=}} P(A|B) \cdot P(B) + P(A|B^c) \cdot P(B^c) \underset{3.13\text{ (ii)}}{\overset{\text{Prop.}}{=}}$

$$= P(A|B) \cdot P(B) + P(A|B) \cdot P(B^c) =$$

$$= P(A|B) \cdot \left[P(B) + P(B^c)\right] \underset{1.16\text{ (iii)}}{\overset{\text{Satz}}{=}} P(A|B)$$

(iii) $\rightarrow$ (i) $P(A \cap B) \underset{3.5\text{ (i)}}{\overset{\text{Satz}}{=}} P(A|B) \cdot P(B) \underset{3.13\text{ (iii)}}{\overset{\text{Prop.}}{=}} P(A) \cdot P(B)$

□

PROPOSITION 3.14 Seien *A* und *B* unabhängige Ereignisse. Dann sind auch die folgenden Ereignisse unabhängig

(i) A^c und B
(ii) A^c und B^c

BEWEIS:

(i) $P(A^c \cap B) = P(B \setminus A) = P(B) - P(A \cap B) \underset{3.12}{\overset{\text{Def.}}{=}} P(B) - P(A) \cdot P(B) =$

$$= P(B) \cdot \left[1 - P(A)\right] \underset{1.16\text{ (iii)}}{\overset{\text{Satz}}{=}} P(A^c) \cdot P(B)$$

(ii) $P(A^c \cap B^c) \underset{1.1\text{ (viii)}}{\overset{\text{Kor.}}{=}} P\left((A \cup B)^c\right) \underset{1.16\text{(iii)}}{\overset{\text{Satz}}{=}} 1 - P(A \cup B) \underset{1.17}{\overset{\text{Satz}}{=}}$

$$= 1 - P(A) - P(B) + P(A \cap B) \underset{3.12.}{\overset{\text{DEF.}}{=}}$$

$$= 1 - P(A) - P(B) + P(A) \cdot P(B) =$$

$$= \left[1 - P(A)\right] - P(B) \cdot \left[1 - P(A)\right] =$$

$$= \left[1 - P(A)\right] \cdot \left[1 - P(B)\right] \underset{1.16\text{(iii)}}{\overset{\text{Satz}}{=}} P(A^c) \cdot P(B^c)$$

□

3.4 Das Simpsonsche Paradoxon

Das Simpsonsche Paradoxon kann bei der Zusammenlegung von Daten auftreten, wenn durch das Gesamtergebnis eine Aussage plötzlich in ihr Gegenteil verwandelt wird.

Beispiel 3.15 *Geschworenen-Jurys in Neuseeland*

Um festzustellen, ob die Maori, die Ureinwohner Neuseelands, von der neuseeländischen Justiz benachteiligt würden, wurde in einer Studie im Herbst 1993 der Anteil der Maori in der Bevölkerung anhand der Volkszählung von 1991 und der Anteil als Geschworene in Gerichtsprozessen von September bis Oktober 1993 verglichen. Dabei ergab sich folgendes überraschende Bild:

Anteil der Maori (in Prozent)			
Distrikt	Bevölkerung (Alter 20-64)	Jury	Unter-repräsentanz
Whangarei	17,0	16,8	0,2
Auckland	9,2	9,0	0,2
Hamilton	13,5	11,5	2,0
Rotorua	27,0	23,4	3,6
Gisborne	32,2	29,5	2,7
Napier	15,5	12,4	3,1
New Plymouth	8,9	4,1	4,8
Palmerston North	8,9	4,3	4,6
Wellington	8,7	7,5	1,2
Nelson	3,9	1,7	2,2
Christchurch	4,5	3,3	1,2
Dunedin	3,3	2,4	0,9
Invercargill	8,4	4,8	3,6
Alle Distrikte	9,5	10,1	-0,6

Tabelle 7 Prozentueller Anteil der Maori in den einzelnen Distrikten Neuseelands

In jedem einzelnen Distrikt betrachtet waren die Maori unterrepräsentiert. Trotzdem hatten sie gesamt gesehen sogar eine Überpräsenz. Wie war das möglich?

Um zu verstehen, wie es zu diesem scheinbaren Widerspruch kommen konnte, betrachten wir die beiden Distrikte Rotorua und Nelson, annähernd gleich große Verwaltungseinheiten, jedoch mit zwei erheblichen Unterschieden. Zum einen hatten diese Distrikte einen wesentlich anderen Anteil an Maori. Zum anderen unterschieden sie sich aber auch hinsichtlich der Zahl an Geschworenen.

Während in Rotorua die Maori 27 Prozent der Bevölkerung ausmachten, waren es in Nelson lediglich knapp 4 Prozent. Außerdem war die Zahl der Jurymitglieder in Rotorua fast sechsmal so groß wie in Nelson.

In beiden Distrikten waren die Maori als Geschworene unterrepräsentiert. Betrachtet man jedoch Rotorua und Nelson zusammen, so entsteht der Eindruck, dass die Maori mit 5 Prozent übervertreten waren.

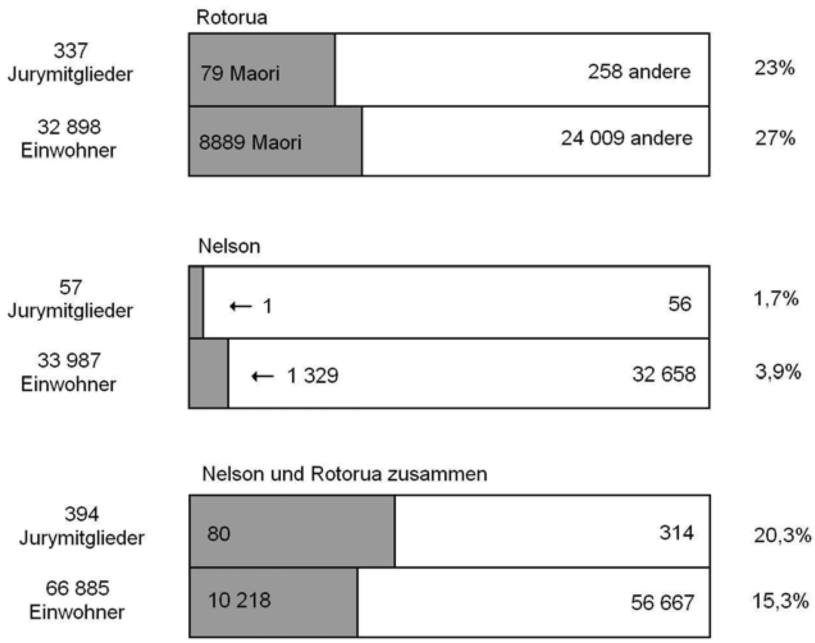


ABBILDUNG 6 Prozentueller Anteil der Maori in der Bevölkerung und als Geschworene anhand der beiden neuseeländischen Distrikte Rotorua und Nelson

Die Ursachen für diesen parodoxen Sachverhalt liegen also in dem unterschiedlichen prozentuellen Anteil der Maori in der Bevölkerung und an Geschworenen in den beiden Bezirken begründet. Diese vereinfachte Situation lässt sich auf ganz Neuseeland ausweiten.

BEMERKUNG 3.16 *Mathematische Beschreibung des Simpsonschen Paradoxons*

Seien $D_1, ..., D_n$ paarweise disjunkte Ereignisse mit $P(D_i) > 0 \ \forall i$ und $\bigcup_{i=1}^{n} D_i = \Omega$. Weiters sei B ein Ereignis mit $P(B \cap D_i) > 0$ und $P(B^c \cap D_i) > 0 \ \forall i$. Dann tritt das Simpsonsche Paradoxon genau dann auf, falls das folgende Ungleichungssystem erfüllt wird (A und B können auch vertauscht werden):

$$\left.\begin{array}{l} P(A|B) < P(A|B^c) \\ P(A|B \cap D_i) > P(A|B^c \cap D_i) \end{array}\right\}$$

Es handelt sich deshalb um ein Paradoxon, da man intuitiv annimmt, dass

$$P(A|B) = \sum_{i=1}^{n} P(D_i|B) \cdot P(A|B \cap D_i)$$

und

$$P(A|B^c) = \sum_{i=1}^{n} P(D_i|B^c) \cdot P(A|B^c \cap D_i)$$

Mittelwerte seien und der Mittelwert von kleineren Werten eigentlich kleiner sein müsste als der Mittelwert von größeren Werten. Dabei übersieht man jedoch, dass es sich bei $P\left(A\middle|B\right)$ und $P\left(A\middle|B^c\right)$ um gewichtete Mittel handelt.

Seien B und D_i unabhängige Ereignisse. Dann kann das Simpsonsche Paradoxon nicht auftreten, da dann

$$P\left(D_i\middle|B\right) = P\left(D_i\middle|B^c\right)\ \forall i$$

gelten würde.

Sei G ein Geschworener und M ein Maori bzw. $\neg G$ kein Geschworener. Weiters bezeichne R einen Bewohner aus Rotorua und N einen Bewohner von Nelson. Dann erhalten wir

$$P\left(M\middle|\neg G\right) = \frac{10138}{66491} \cong 0{,}1525 < P\left(M\middle|G\right) = \frac{80}{394} \cong 0{,}2030$$

$$P\left(M\middle|\neg G \cap R\right) = \frac{8810}{32561} \cong 0{,}2706 > P\left(M\middle|G \cap R\right) = \frac{79}{337} \cong 0{,}2344$$

$$P\left(M\middle|\neg G \cap N\right) = \frac{1328}{33930} \cong 0{,}0391 > P\left(M\middle|G \cap N\right) = \frac{1}{57} \cong 0{,}0175$$

BEMERKUNG 3.17 *Auflösung des Paradoxons von Beispiel 3.15*

Um entscheiden zu können, ob die Maori nun unter- oder überrepräsentiert waren, müssen wir einen Blick auf die Jurorenplätze werfen. Wir erkennen, dass insgesamt 50 der 350 Geschworenen in ganz Neuseeland fehlbesetzt waren und somit die Maori tatsächlich unterrepräsentiert waren.

Anteil der Maori (in Prozent)			
Distrikt	Tatsächlich	Erwartet	Differenz
Whangarei	28	28	0
Auckland	74	76	2
Hamilton	23	27	4
Rotorua	79	91	12
Gisborne	23	25	2
Napier	15	19	4
New Plymouth	4	9	5
Palmerston North	7	14	7
Wellington	28	33	5
Nelson	1	2	1
Christchurch	11	15	4
Dunedin	4	6	2
Invercargill	3	5	2
Alle Distrikte	300	350	50

TABELLE 8 Übersicht über die Anzahl der fehlbesetzten Geschworenen in Neuseeland

Die dargestellte Sichtweise ist nur für dieses konkrete Problem sinnvoll. In anderen Problemstellungen können jedoch auch andere Betrachtungsweisen geeigneter sein. Nicht immer darf dem Teilresultat mehr Vertrauen geschenkt werden als dem Gesamtergebnis.

Verwendete Literatur zum Simpsonschen Paradoxon:

- ♠ BLYTH, Colin, On Simpson's Paradox and the Sure-Thing Principle, In: Journal of the American Statistical Association Vol. 67(1972), Nummer 338, S.364-366
- ♠ BÜCHTER, Andreas / HENN, Hans-Wolfgang, Elementare Stochastik. Eine Einführung in die Mathematik des Zufalls, Berlin / Heidelberg [2]2006, S.143-147
- ♠ DUBBEN, Hans-Hermann / BECK-BORNHOLDT, Hans-Peter, Mit an Wahrscheinlichkeit grenzender Sicherheit. Logisches Denken und Zufall, Reinbek bei Hamburg [2]2006, S.137-156
- ♠ IRLE, Albrecht, Wahrscheinlichkeitstheorie und Statistik, Stuttgart 2001, S.43f.
- ♠ JAHNKE, Thomas, Das Simpsonsche Paradoxon verstehen – ein Beitrag des Mathematikunterrichts zur Allgemeinbildung, In: Journal für Mathematik-Didaktik Vol. 14(1993), Heft 3/4, S.221-242
- ♠ KOHN, Wolfgang, Statistik. Datenanalyse und Wahrscheinlichkeitsrechnung, Berlin / Heidelberg 2005, S.131-133
- ♠ KRÄMER, Walter, Wie lügt man mit Statistik?, In: Stochastik in der Schule 1/1991, S.3-24 (vor allem S.11-12)
- ♠ KÜNZEL, Erhard, Über Simpsons Paradoxon, In: Stochastik in der Schule 1/1991, S.54-62
- ♠ SZEKELEY, Gábor, Paradoxa. Klassische und neue Überraschungen aus Wahrscheinlichkeitsrechung und mathematischer Statistik, Frankfurt am Main 1990, S.63f.
- ♠ WESTBROOKE, Ian, Simpson's Paradox. An example in a New Zealand Survey of Jury Composition. Research Report #1 (revised), Wellington 1997[8]

[8] downloadbar unter http://www.stats.govt.nz/NR/rdonlyres/5DFEE5C8-A969-4684-B95F-FA2AFEDA649A/0/Simpdox.pdf [01.01.2009]

4. Zufallsvariable und Verteilungen

In der Wahrscheinlichkeitstheorie interessiert man sich vor allem für die Ergebnisse bestimmter Ereignisse. Dabei treten Zahlen auf, die bei einer Wiederholung stets einen anderen Wert liefern können. Zufallsvariable werden dazu verwendet um jedem Ausgang eines Ereignisses eine Zahl zuordnen zu können.

In diesem Kapitel möchte ich neben dem Begriff Zufallsvariable auch noch die Verteilungs-, die Wahrscheinlichkeits- sowie die Dichtefunktion einführen, diskrete und stetige Zufallsvariable behandeln und kurz auf mehrdimensionale Zufallsvariable eingehen.

Die Verteilungsfunktion der Zufallsvariablen X gibt die Wahrscheinlichkeit an, dass die Zufallsvariable höchstens einen Wert x annimmt.

Ich werde Zufallsvariable in zwei Klassen einteilen, in die der diskreten Zufallsvariablen und in die der stetigen Zufallsvariablen. Während eine diskrete Zufallsvariable nur endlich viele oder abzählbar unendlich viele Werte annehmen kann, so können stetige Zufallsvariable alle Werte aus einem bestimmten Intervall annehmen, also überabzählbar unendlich viele Werte.

Die Wahrscheinlichkeitsfunktion ordnet jedem Wert einer diskreten Zufallsvariablen eine Wahrscheinlichkeit zu. Die Wahrscheinlichkeitsfunktion kann man mit Hilfe eines Stabdiagramms graphisch darstellen.

Für stetige Zufallsvariable ist der Dichtebegriff essentiell. Hierbei handelt es sich um eine positive Funktion, deren gesamter Flächeninhalt 1 ergibt.

Die Verteilungsfunktion ist monoton wachsend. Der Graph einer diskreten Zufallsvariablen ist eine monoton steigende Treppenfunktion. Die Verteilungsfunktion einer stetigen Zufallsvariablen ist die Stammfunktion der Dichtefunktion.

Mehrdimensionale Zufallsvariable werden benötigt, wenn man sich nicht nur für ein einziges Merkmal interessiert, sondern wenn man Zufallsexperimente betrachtet, bei denen gleichzeitig mindestens zwei Zufallsvariable auftreten.

4.1 DER BEGRIFF EINER ZUFALLSVARIABLEN

DEFINITION 4.1 *Zufallsvariable*

Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum. Eine Zufallsvariable *X* ist eine Funktion, die jedem Ereignis ω des Zufallsexperiments eine reelle Zahl $X(\omega) = x$ zuordnet.

$$X : \Omega \to \mathbb{R}, \omega \mapsto X(\omega)$$

Normalerweise werden für Variable kleine Buchstaben verwendet. Zufallsvariable werden jedoch mit Großbuchstaben gekennzeichnet, deren angenommene Werte mit Kleinbuchstaben. Dies hat einen plausiblen Grund, nämlich um zwischen normalen Variablen und Zufallsvariablen unterscheiden zu können. Die Aussage

- ♠ $x = 4$ … ist richtig oder falsch
- ♠ $X = 4$ … ist mit einer bestimmten Wahrscheinlichkeit richtig oder falsch

Tritt beim Zufallsexperiment das Elementarereignis ω ein, dann nimmt die Zufallsvariable *X* den Wert $X(\omega)$ an. Wir nennen dann $X(\omega)$ die Realisierung der Zufallsvariablen *X*. Das ist also der Wert der Zufallsvariablen *X*, den wir tatsächlich beobachten.

BEISPIEL 4.2 *Einige Zufallsvariablen und deren Realisierungen*

- ♠ Interessiert man sich nur dafür, ob ein Ereignis *A* eintritt oder ob nicht, so nimmt die Zufallsvariable lediglich die beiden Werte 0 oder 1 an

$$X = \begin{cases} 1 \text{ falls } A \text{ eintritt} \\ 0 \text{ falls } A \text{ nicht eintritt} \end{cases}$$

- ♠ Bei einem Fragebogen gibt es 4 Fragen mit jeweils den Antwortmöglichkeiten „ja“ und „nein“. Sei die Zufallsvariable *X* die Anzahl der Fragen, die mit „ja“ beantwortet wurden. Dann sind die angenommenen Werte 0, 1, 2, 3 und 4.
- ♠ Bei einer Wahl treten sechs Parteien an. Jeder Partei wird nun eine Zahl von 1 bis 6 zugeordnet.

Notation und Kurzschreibweise für Ereignisse und Wahrscheinlichkeiten mit der Zufallsvariablen *X:*

- ♠ $\{X = a\} := \{\omega \in \Omega : X(\omega) = a\}$
- ♠ $P(X = a) := P(\{\omega \in \Omega : X(\omega) = a\})$
- ♠ $P(a \leq X \leq b) := P(\{\omega \in \Omega : a \leq X(\omega) \leq b\})$
- ♠ analog: $P(X < a), P(X \leq a), P(X > a), P(X \geq a)$

4.2 VERTEILUNGSFUNKTION

DEFINITION 4.3 *Verteilungsfunktion*

Die Funktion

$$F : \mathbb{R} \to [0,1],$$
$$x \mapsto F(x) = P(\{\omega \in \Omega : X(\omega) \leq x\}) = P(]-\infty, x]) = P(X \leq x)$$

nennt man die *Verteilungsfunktion* der Zufallsvariablen X

Die Verteilungsfunktion der Zufallsvariablen X ordnet also jeder Zahl x die Wahrscheinlichkeit des Ereignisses zu, dass X maximal den Wert x annimmt. Dieses Ereignis lässt sich schreiben als

$$\{X \leq x\} = \{\omega \in \Omega : X(\omega) \leq x\}$$

Die dazugehörige Wahrscheinlichkeit ist von der Form

$$P(X \leq x) = P(\{\omega \in \Omega : X(\omega) \leq x\})$$

SATZ 4.4 *Eigenschaften der Verteilungsfunktion F*

(i) F ist monoton wachsend: $x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$

(ii) F wächst von null bis eins: $\lim\limits_{x \to -\infty} F(x) = 0, \lim\limits_{x \to \infty} F(x) = 1$

(iii) F ist rechtsseitig stetig: $\lim\limits_{\substack{x_n \to x \\ x_n > x}} F(x_n) = F(x)$

BEWEIS:

(i) Seien x_1, x_2 zwei reelle Zahlen mit $x_1 < x_2$

$$\Rightarrow]-\infty, x_1] \subseteq]-\infty, x_2]$$
$$\Rightarrow F(x_1) = P(X \leq x_1) \leq P(X \leq x_2) = F(x_2)$$

(ii) Sei (x_n) eine streng monoton steigende Folge mit $\lim\limits_{n \to \infty} x_n = \infty$

$$\bigcup_{n=1}^{\infty}]-\infty, x_n] = \mathbb{R}$$

$$\lim_{n \to \infty} F(x_n) = \lim_{n \to \infty} P(]-\infty, x_n]) \underset{1.18\text{ (i)}}{\overset{\text{Satz}}{=}} P\left(\bigcup_{n=1}^{\infty}]-\infty, x_n]\right) = P(\mathbb{R}) = 1$$

Sei (x_n) eine streng monoton fallende Folge mit $\lim\limits_{n \to \infty} x_n = -\infty$

$$\bigcap_{n=1}^{\infty}]-\infty, x_n] = \varnothing$$

$$\lim_{n \to \infty} F(x_n) = \lim_{n \to \infty} P(]-\infty, x_n]) \underset{1.18\text{ (ii)}}{\overset{\text{Satz}}{=}} P\left(\bigcap_{n=1}^{\infty}]-\infty, x_n]\right) = P(\varnothing) = 0$$

(iii) Sei (x_n) eine streng monoton fallende Folge mit $\lim\limits_{n\to\infty} x_n = x$

$$\bigcap_{n=1}^{\infty} \left]-\infty, x_n\right] = \left]-\infty, x\right]$$

$$\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} P\left(\left]-\infty, x_n\right]\right) \underset{1.18\text{ (ii)}}{\overset{\text{Satz}}{=}} P\left(\bigcap_{n=1}^{\infty} \left]-\infty, x_n\right]\right) =$$
$$= P\left(\left]-\infty, x\right]\right) = F(x)$$

□

Umgekehrt kann man zu jeder Funktion F mit diesen drei Eigenschaften eine Zufallsvariable X derart konstruieren, dass F die Verteilungsfunktion von X ist.[9]

Bemerkung 4.5 Den linksseitigen Grenzwert der Verteilungsfunktion F bezeichnet man mit

$$F(x-) = \lim_{\substack{\varepsilon\to x \\ \varepsilon<x}} F(x-\varepsilon)$$

Bemerkung 4.6 Berechnung einzelner Wahrscheinlichkeiten mit Hilfe der Verteilungsfunktion:

(i) $P(X \le a) = F(a)$

(ii) $P(X < a) = F(a-)$

(iii) $P(X = a) = F(a) - F(a-)$

(iv) $P(X > a) = 1 - P(X \le a) = 1 - F(a)$

(v) $P(X \ge a) = 1 - F(a-)$

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

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

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

[9] Einen Beweis findet man etwa in IRLE, Albrecht, Wahrscheinlichkeitstheorie und Statistik. Grundlagen – Resultate – Anwendungen, Stuttgart 2001, S.66f.

4.3 DISKRETE ZUFALLSVARIABLE

DEFINITION 4.7 *diskrete Zufallsvariable*

Eine Zufallsvariable X heißt diskret, falls sie nur endlich viele oder abzählbar unendlich viele Werte annehmen kann.

BEISPIEL 4.8 *Beispiele für diskrete Zufallsvariable*

- ♠ X = Anzahl der Wähler einer bestimmten Partei
- ♠ Y = Anzahl der Befragungen bis zum ersten Wähler einer bestimmten Partei

DEFINITION 4.9 *Wahrscheinlichkeitsfunktion*

Sei X eine diskrete Zufallsvariable mit dem Wertebereich $\{x_1, x_2, ...\}$. Dann nennt man die Funktion

$$p : X(\Omega) \mapsto \mathbb{R}, p(x_i) = \begin{cases} P(X = x_i) \text{ falls } x \in X(\Omega) \\ 0 \text{ sonst} \end{cases}$$

die Wahrscheinlichkeitsfunktion von X.

PROPOSITION 4.10 *Eigenschaften einer Wahrscheinlichkeitsfunktion*

(i) $p(x_i) \geq 0$

(ii) $\sum_{i=1}^{\infty} p(x_i) = 1$

BEWEIS:

(i) Laut Definition 1.14 (i) sind alle Wahrscheinlichkeiten $p(x_i) \geq 0$

(ii) $$\sum_{i=1}^{\infty} p(x_i) \underset{4.9}{\overset{\text{Def.}}{=}} \sum_{i=1}^{\infty} P(X = x_i) \underset{1.14\text{ (iii)}}{\overset{\text{Def.}}{=}} P\left(\bigcup_{i=1}^{\infty} \{X = x_i\}\right) = P(\Omega) \underset{1.14\text{ (ii)}}{\overset{\text{Def.}}{=}} 1$$

□

DEFINITION 4.11 *Verteilungsfunktion einer diskreten Zufallsvariablen X*

Unter der Verteilungsfunktion einer diskreten Zufallsvariablen X versteht man die für jedes $x \in \mathbb{R}$ durch

$$F(x) = P(X \leq x) = \sum_{i:x_i \leq x} P(X = x_i)$$

definierte Funktion F.

KOROLLAR 4.12 Eigenschaften der diskreten Verteilungsfunktion *F* einer diskreten Zufallsvariablen *X*

(i) *F* ist eine monoton steigende Treppenfunktion, die von 0 auf 1 anwächst.
(ii) Die Sprungstellen von *F* sind die Werte x_i aus dem Wertebereich von *X*.
(iii) Die Sprunghöhen von *F* entsprechen den Wahrscheinlichkeiten $p_i = P\left(X = x_i\right)$.
(iv) *F* ist rechtsseitig stetig $\forall x$. Außerhalb des Wertebereichs von *X* ist *F* überall stetig.

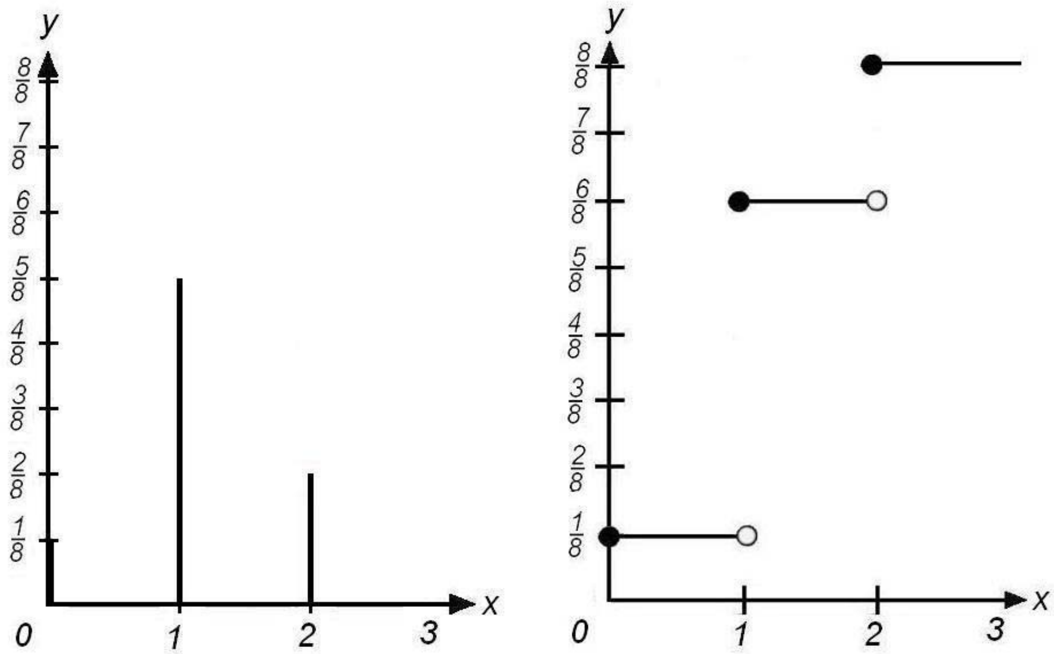


ABBILDUNG 7 Graphische Veranschaulichung und Zusammenhang zwischen Wahrscheinlichkeitsfunktion (links) und Verteilungsfunktion (rechts)

4.4 Stetige Zufallsvariable

Definition 4.13 Eine *Dichte(funktion)* $f(x)$ ist eine auf ganz $\mathbb{R}$ integrierbare Funktion mit den beiden folgenden Eigenschaften

(i) $f(x) \geq 0 \; \forall x \in \mathbb{R}$

(ii) $\int\limits_{-\infty}^{\infty} f(x)\,dx = 1$

Definition 4.14 *stetige Zufallsvariable*

Eine Zufallsvariable X heißt stetig, falls sie überabzählbar unendlich viele Werte annehmen kann und falls es eine Dichtefunktion $f(x)$ derart gibt, dass für alle reelle Zahlen $a < b$

$$P\left(a \leq X \leq b\right) = \int\limits_{a}^{b} f(x)\,dx$$

gilt.

Anschaulich ist klar, dass die Wahrscheinlichkeit für $x \in [a, b]$ gleich dem Flächeninhalt ist, den die Dichte $f(x)$ mit der x-Achse einschließt.

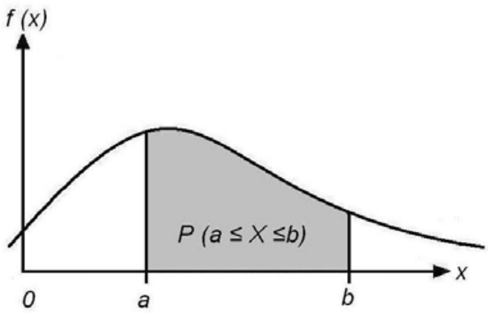


Abbildung 8 Graphischer Zusammenhang zwischen Wahrscheinlichkeit und Flächeninhalt

Beispiel 4.15 *Beispiele für stetige Zufallsvariable*

- ♠ X = Antwortmöglichkeiten auf einer stetigen Rating-Skala
- ♠ Y = Dauer einer Regierungsbildung

Definition 4.16 Sei f die Dichte einer stetigen Zufallsvariablen X. Dann versteht man unter der *Verteilungsfunktion einer stetigen Zufallsvariablen X* die für jedes $x \in \mathbb{R}$ durch

$$F\left(x\right) = P\left(X \leq x\right) = \int\limits_{-\infty}^{x} f(t)\,dt$$

definierte Funktion F.

KOROLLAR 4.17 Eigenschaften der stetigen Verteilungsfunktion F einer stetigen Zufallsvariablen X mit der Dichte f

(i) F ist eine monoton steigende Funktion, die von 0 auf 1 anwächst.

(ii) F ist überall stetig. (Begründung: $\forall x \in \mathbb{R}$ ist $\lim\limits_{x\to a} F(x) = \lim\limits_{x\to a} \int_{-\infty}^{x} f(t)\,dt =$

$= \int_{-\infty}^{a} f(t)\,dt = F(a)$, weil $f(t)$ auf ganz $\mathbb{R}$ integrierbar ist.)

(iii) Die Verteilungsfunktion F ist eine Stammfunktion der Dichtefunktion f, es gilt also $f(x) = F'(x)\ \forall x$, wo f stetig ist.

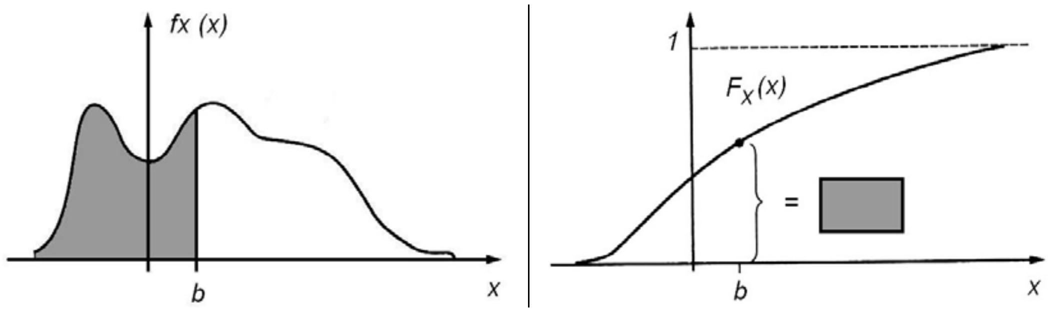


ABBILDUNG 9 Graphische Veranschaulichung und Zusammenhang zwischen Dichtefunktion (links) und Verteilungsfunktion (rechts)[10]

PROPOSITION 4.18 Sei X eine stetige Zufallsvariable. Dann gilt

$$P(X = a) = 0\ \forall a \in \mathbb{R}$$

BEWEIS:

$$F(x-) \underset{4.5}{\overset{\text{Bem.}}{=}} \lim_{\substack{\varepsilon\to x \\ \varepsilon<x}} F(x-\varepsilon) \overset{\text{da } F \text{ stetig ist}}{=} F(x)\ \forall x \in \mathbb{R}$$

$$\Rightarrow P(X = a) \underset{4.6\text{ (iii)}}{\overset{\text{Bem.}}{=}} F(x) - F(x-) = 0$$

□

Eine stetige Zufallsvariable nimmt also jede reelle Zahl nur mit der Wahrscheinlichkeit 0 an. Trotzdem ergeben alle Realisierungen zusammen die Wahrscheinlichkeit 1.

Wegen $P(X = a) = 0$ ist bei der Berechnung von Wahrscheinlichkeiten für Intervalle irrelevant, ob die Intervallgrenzen dazugenommen werden oder ob nicht. Folglich gilt

$$P(a < X < b) = P(a \le X < b) = P(a < X \le b) = P(a \le X \le b) = \int_a^b f(x)\,dx$$

[10] CLAUß, Günter / FINZE, Falk-Rüdiger / PARTZSCH, Lothar, Statistik für Soziologen, Pädagogen, Pychologen und Mediziner. Grundlagen, Frankfurt am Main [4]2002, S.136

4.5 MEHRDIMENSIONALE ZUFALLSVARIABLE

In diesem Kapitel beschäftigen wir uns der Einfachheit halber nur mit zwei Zufallsvariablen X und Y, die sich auf denselben Ereignisraum Ω beziehen. Die Resultate dieses Abschnittes lassen sich aber problemlos auch auf n Zufallsvariable verallgemeinern.

DEFINITION 4.19 *Zweidimensionale Zufallsvariable*

Sei $(\Omega, \mathscr{A}, P)$ ein Wahrscheinlichkeitsraum und seien X, Y zwei Zufallsvariable. Dann ist der Vektor $(X,Y):\Omega \to \mathbb{R}^2$ eine zweidimensionale Zufallsvariable.

DEFINITION 4.20 *Gemeinsame Verteilungsfunktion von X und Y*

Sei $(X,Y):\Omega \to \mathbb{R}^2$ eine zweidimensionale Zufallsvariable. Dann nennt man die Abbildung

$$F_{XY}:\mathbb{R}^2 \to [0,1] \text{ mit}$$

$$F_{XY}(x,y) := P\left(\{\omega : X(\omega) \le x \wedge Y(\omega) \le y\}\right) = P(X \le x, Y \le y)$$

gemeinsame Verteilungsfunktion von X und Y.

Eigenschaften der Verteilungsfunktion F_{XY}

- ♠ $0 \le F_{XY}(x,y) \le 1$
- ♠ $F_{XY}(x,-\infty) = \lim\limits_{y\to-\infty} F(x,y) = 0 \;\forall x \in \mathbb{R}$

  $F_{XY}(-\infty,y) = \lim\limits_{x\to-\infty} F(x,y) = 0 \;\forall y \in \mathbb{R}$
- ♠ $F_{XY}(\infty,\infty) = 1$
- ♠ $F_{XY}(x,\infty) = F_X(x)$ und $F_{XY}(\infty,y) = F_Y(y)$
- ♠ Sei $x_1 < x_2$ und $y_1 < y_2$. Dann gilt

  $F_{XY}(x_1,y_1) \le F_{XY}(x_2,y_1) \le F_{XY}(x_2,y_2)$

  $F_{XY}(x_1,y_1) \le F_{XY}(x_1,y_2) \le F_{XY}(x_2,y_2)$

BEMERKUNG 4.21 *Wahrscheinlichkeit eines Rechtecks:*

Seien $a_1, a_2, b_1, b_2 \in \mathbb{R}$ mit $a_1 < b_1$ und $a_2 < b_2$. Dann beträgt die Wahrscheinlichkeit, dass ein beliebiger Punkt (X,Y) in das achsenparallele Rechteck $(a_1, b_1]x(a_2, b_2]$ fällt

$$P(a_1 < X \le b_1, a_2 < Y \le b_2) =$$

$$= F_{XY}(b_1, b_2) - F_{XY}(b_1, a_2) - F_{XY}(a_1, b_2) + F_{XY}(a_1, a_2)$$

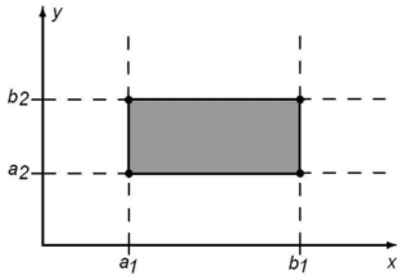


ABBILDUNG 10 Graphische Veranschaulichung für die Wahrscheinlichkeit, dass (X,Y) einen Wert aus $\left\{(x,y)\in\mathbb{R}^2 : a_1 < x \le a_2, b_1 < y \le b_2\right\}$ annimmt

DEFINITION 4.22 *Randverteilungsfunktionen*

Die beiden Verteilungsfunktionen der Zufallsvariablen X und Y

$$F_X(x) = F_{XY}(x,\infty) = \lim_{y\to\infty} F_{XY}(x,y),\ x\in\mathbb{R}$$

$$F_Y(y) = F_{XY}(\infty,y) = \lim_{x\to\infty} F_{XY}(x,y),\ y\in\mathbb{R}$$

nennt man Randverteilungsfunktionen.

DEFINITION 4.23 *Zweidimensionale diskrete Verteilung*

Sei $(X,Y):\Omega\to\mathbb{R}^2$ eine zweidimensionale Zufallsvariable. Dann ist diese Zufallsvariable genau dann diskret, falls X,Y diskrete Zufallsvariable sind.

DEFINITION 4.24 *mehrdimensionale Wahrscheinlichkeitsfunktion*

Sei $(X,Y):\Omega\to\mathbb{R}^2$ eine diskrete zweidimensionale Zufallsvariable. Dann nennt man die Funktion

$$p: X(\Omega)\mapsto\mathbb{R},\ p_{XY}(x,y) = \begin{cases} P\left(X = x_i, Y = y_j\right) \text{ falls } x\in X(\Omega) \\ 0 \text{ sonst} \end{cases}$$

die Wahrscheinlichkeitsfunktion von (X,Y).

Analog zum eindimensionalen Fall (siehe Proposition 4.10) besitzt auch eine zweidimensionale Wahrscheinlichkeitsfunktion p_{XY} die folgenden Eigenschaften

(i) $P\left(X = x_i, Y = y_j\right) \ge 0$

(ii) $\sum_i \sum_j P\left(X = x_i, Y = y_j\right) = 1$

DEFINITION 4.25 Als Randwahrscheinlichkeiten einer diskreten zweidimensionalen Zufallsvariable erhält man die beiden Wahrscheinlichkeiten

$$P(X = x_i) = \sum_j P\left(X = x_i, Y = y_j\right)$$

$$P\left(Y = y_j\right) = \sum_i P\left(X = x_i, Y = y_j\right).$$

Nur falls zwei diskrete Zufallsvariable mit endlichem Wertebereich vorliegen, so lässt sich die gemeinsame Verteilungsfunktion in einer Tabelle darstellen, wobei die Randverteilungen an den Zeilen- und Spaltensummen am Rand abzulesen sind.

BEISPIEL 4.26 Bei einer Telefonbefragung wurden n Personen angerufen, wobei angenommen wird, dass eine Person auch zweimal befragt werden konnte (entspricht einem Ziehen mit Zurücklegen). Dabei gaben k Personen an, dass sie für eine große Koalition seien. Wir interessieren uns jetzt für die ersten beiden Anrufer. Es sei

$$X = \begin{cases} 1 \text{ falls 1.Anrufer eine große Koalition befürwortet hat} \\ 0 \text{ sonst} \end{cases}$$

$$Y = \begin{cases} 1 \text{ falls 2.Anrufer eine große Koalition befürwortet hat} \\ 0 \text{ sonst} \end{cases}.$$

Wie lauten die einzelnen Wahrscheinlichkeiten sowie die Randwahrscheinlichkeiten?

	$Y=0$	$Y=1$	$\sum_i P(X=x_i)$
$X=0$	$\left(\frac{n-k}{n}\right)^2$	$\frac{n-k}{n}\cdot\frac{k}{n}$	$\frac{n-k}{n}$
$X=1$	$\frac{k}{n}\cdot\frac{n-k}{n}$	$\left(\frac{k}{n}\right)^2$	$\frac{k}{n}$
$\sum_j P(Y=y_j)$	$\frac{n-k}{n}$	$\frac{k}{n}$	1

TABELLE 9 Die gemeinsamen Wahrscheinlichkeiten sowie die Randwahrscheinlichkeiten von X und Y beim Ziehen mit Zurücklegen

BEISPIEL 4.27 Bei einer Stichwahl zwischen den beiden KandidatInnen A und B wurden insgesamt n gültige Stimmen abgegeben, wobei die KandidatInnen k bzw. $n-k$ Stimmen erhalten haben. Bei der Auszählung wurden zwei gültige Stimmzettel gezogen. (entspricht einem Ziehen ohne Zurücklegen). Es sei

$$X = \begin{cases} 1 \text{ falls zuerst KandidatIn } A \text{ angekreuz wurde} \\ 0 \text{ falls zuerst KandidatIn } B \text{ angekreuz wurde} \end{cases}$$

$$Y = \begin{cases} 1 \text{ falls danach KandidatIn } A \text{ angekreuz wurde} \\ 0 \text{ falls danach KandidatIn } B \text{ angekreuz wurde} \end{cases}.$$

Wie lauten nun die einzelnen Wahrscheinlichkeiten sowie die Randwahrscheinlichkeiten?

	$Y=0$	$Y=1$	$\sum_i P(X=x_i)$
$X=0$	$\frac{n-k}{n}\cdot\frac{n-k-1}{n-1}$	$\frac{n-k}{n}\cdot\frac{k}{n-1}$	$\frac{n-k}{n}$
$X=1$	$\frac{k}{n}\cdot\frac{n-k}{n-1}$	$\frac{k}{n}\cdot\frac{k-1}{n-1}$	$\frac{k}{n}$
$\sum_j P(Y=y_j)$	$\frac{n-k}{n}$	$\frac{k}{n}$	1

TABELLE 10 Die gemeinsamen Wahrscheinlichkeiten sowie die Randwahrscheinlichkeiten von *X* und *Y* beim Ziehen ohne Zurücklegen

DEFINITION 4.28 *Zweidimensionale stetige Verteilung*

Sei $(X,Y):\Omega\to\mathbb{R}^2$ eine zweidimensionale Zufallsvariable mit der Verteilungsfunktion F_{XY}. Dann ist diese Zufallsvariable stetig, falls es eine integrierbare Funktion $f:\mathbb{R}^2\to\mathbb{R}$ gibt mit

$$F_{XY}(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x} f_{XY}(u,v)\,du\,dv\,.$$

Die Funktion f_{XY} ist die gemeinsame Dichte von *X* und *Y*.
Die Randverteilungsfunktionen sind gegeben durch

$$F_X(x)=\int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{XY}(u,v)\,du\,dv$$

$$F_Y(y)=\int_{-\infty}^{y}\int_{-\infty}^{\infty} f_{XY}(u,v)\,du\,dv.$$

Wie schon im eindimensionalen Fall (siehe Definition 4.13) ist auch eine zweidimensionale Dichtefunktion f_{XY} durch die folgenden beiden Eigenschaften charakterisiert

(i) $f(x,y)\geq 0\ \forall x,y\in\mathbb{R}$

(ii) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\,dy=1$

DEFINITION 4.29 *Randdichten*

Sei $(X,Y):\Omega\to\mathbb{R}^2$ eine zweidimensionale Zufallsvariable mit der Dichtefunktion $f_{XY}(x,y)$. Dann ist

$$f_X(x)=\int_{-\infty}^{\infty} f_{XY}(x,y)\,dy$$

die Randdichte von *X* und analog

$$f_Y(y)=\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx$$

die Randdichte von *Y*.

Die Randdichten sind also die Ableitungen der entsprechenden Randverteilungsfunktionen

$$f_X(x) = \frac{\partial}{\partial x} F_X(x),\, f_Y(y) = \frac{\partial}{\partial y} F_Y(y)$$

Allgemein gilt

$$f_{XY}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x,y)$$

DEFINITION 4.30 *Stochastische Unabhängigkeit*

Sei $(X,Y): \Omega \to \mathbb{R}^2$ eine zweidimensionale Zufallsvariable mit der Verteilungsfunktion F_{XY}. Dann heißt diese Zufallsvariable stochastisch unabhängig, falls

$$F_{XY}(x,y) = F_X(x) \cdot F_Y(y)$$

Für diskrete Zufallsvariable X und Y gilt

$$P\left(X = x_i, Y = y_j\right) = P\left(X = x_i\right) \cdot P\left(Y = y_j\right),$$

für stetige Zufallsvariable X und Y erhält man

$$P\left(X = x_i, Y = y_j\right) = P\left(X \le x_i\right) \cdot P\left(Y \le y_j\right).$$

Die Randverteilungen von X und Y beim Ziehen mit Zurücklegen (Beispiel 4.26) entsprechen den Randwahrscheinlichkeiten ohne Zurücklegen (Beispiel 4.27). Jedoch sind nur die Zufallsvariable X und Y in Beispiel 4.26 unabhängig, da

$$P(X=1) = P(X=1, Y=0) + P(X=1, Y=1) =$$

$$= \frac{k}{n} \cdot \frac{n-k}{n} + \left(\frac{k}{n}\right)^2 = \frac{k}{n} \cdot \left(\frac{n-k+k}{n}\right) = \frac{k}{n}$$

$$P(Y=0) = P(X=0, Y=0) + P(X=1, Y=0) =$$

$$= \left(\frac{n-k}{n}\right)^2 + \frac{k}{n} \cdot \frac{n-k}{n} = \frac{n-k}{n} \cdot \left(\frac{n-k+k}{n}\right) = \frac{n-k}{n}$$

$$P(X=1, Y=0) = \frac{k}{n} \cdot \frac{n-k}{n} = P(X=1) \cdot P(Y=0)$$

In Beispiel 4.27 dagegen gilt für $X = 1, Y = 0$:

$$P(X=1) = P(X=1, Y=0) + P(X=1, Y=1) =$$

$$= \frac{k}{n} \cdot \frac{n-k}{n-1} + \frac{k}{n} \cdot \frac{k-1}{n-1} = \frac{k}{n \cdot (n-1)} \cdot \left(n \cancel{-k} \cancel{+k} - 1\right) = \frac{k}{n}$$

$$P(Y=0) = P(X=0, Y=0) + P(X=1, Y=0) =$$

$$= \frac{n-k}{n} \cdot \frac{n-k-1}{n-1} + \frac{k}{n} \cdot \frac{n-k}{n-1} = \frac{n-k}{n \cdot (n-1)} \cdot \left(n \cancel{-k} - 1 \cancel{+k}\right) = \frac{n-k}{n}$$

$$P(X=1, Y=0) = \frac{k}{n} \cdot \frac{n-k}{n-1} \ne \frac{k}{n} \cdot \frac{n-k}{n} = P(X=1) \cdot P(Y=0)$$

5. Lage- und Streuungsparameter

In diesem Kapitel geht es um den Erwartungswert, die Varianz und die Standardabweichung von Zufallsvariablen.

Der Erwartungswert ist der mittlere Wert der Zufallsvariablen X, um den sich die Werte der Zufallsvariablen anordnen. Der Erwartungswert entspricht somit dem arithmetischen Mittel einer Stichprobe vom Umfang n und ist ein Lageparameter.

Die Varianz und die Standardabweichung dagegen sind Streuungsparameter, da sie die mittlere Abweichung der Zufallsvariablen X vom Erwartungswert μ angeben. Sie entsprechen der empirischen Varianz und der Standardabweichung einer Stichprobe vom Umfang n. Physikalisch betrachtet entspricht der Erwartungswert dem Schwerpunkt, die Varianz dem Trägheitsmoment.

Die Ungleichung von Tschebyschew schließlich liefert eine grobe Abschätzung für die Wahrscheinlichkeit, mit der eine Zufallsvariable um mehr als ε von ihrem Erwartungswert abweicht. Sie ist insofern von Nutzem, wenn keinerlei Informationen über die Verteilung vorliegen. Bessere Ergebnisse erzielt man, man die Verteilung bekannt ist, wie ich in Satz 6.11 anhand der Normalverteilung zeigen werde.

Mathematisch betrachtet hat die Ungleichung von Tschebyschew hauptsächlich theoretische Bedeutung, da sie bei den Beweisen der Gesetze der großen Zahlen in Kapitel 12 benötigt wird. Diese Gesetze besagen, dass der Stichprobenmittelwert bei wachsender Stichprobengröße gegen den Erwartungswert konvergiert.

5.1 Erwartungswert

Definition 5.1 *Erwartungswert einer diskreten Zufallsvariablen*

Sei X eine diskrete Zufallsvariable mit der Wahrscheinlichkeitsverteilung $f(x_i) = P(X = x_i)$. Falls

$$E(X) = \mu = \sum_i x_i \cdot P(X = x_i)$$

existiert, dann heißt $E(X)$ der Erwartungswert von X.

Definition 5.2 *Erwartungswert einer stetigen Zufallsvariablen*

Sei X eine stetige Zufallsvariable mit der Dichte $f(x)$. Falls

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

existiert, dann heißt $E(X)$ der Erwartungswert von X.

Der Erwartungswert $E(X)$ existiert, falls

$$\sum_i |x_i| \cdot P(X = x_i) < \infty \text{ für diskrete Zufallsvariable } X$$

$$\int_{-\infty}^{\infty} |x| \cdot f(x)dx < \infty \text{ für stetige Zufallsvariable } X$$

Bei diskreten Zufallsvariablen muss der Erwartungswert keinen Wert von X annehmen. Im Falle von stetigen Zufallsvariablen gilt aufgrund von Proposition 4.18 sogar $P(X = \mu) = 0$.

Satz 5.3 *Funktionssatz*

Sei $g : \mathbb{R} \to \mathbb{R}$ eine Funktion und X eine Zufallsvariable. Dann ist $Y = g(X)$ ebenfalls eine Zufallsvariable und es gilt

$$E(Y) = E(g(X)) = \begin{cases} \sum_i g \cdot (x_i) \cdot P(X = x_i) \text{ falls } X \text{ diskret ist} \\ \int_{-\infty}^{\infty} g(x) \cdot f(x)dx \text{ falls } X \text{ stetig ist} \end{cases}$$

Beweis:

Für diskrete Zufallsvariable Y gilt

$$y = g(x) \Rightarrow P(Y = y_j) = \sum_{i:g(x_i)=y_j} P(X = x_i)$$

$$E(Y) \underset{5.1}{\overset{\text{Def.}}{=}} \sum_j y_j \cdot P(Y = y_j) = \sum_j y_j \cdot \sum_{i:g(x_i)=y_j} P(X = x_i) =$$

$$= \sum_j \sum_{i:g(x_i)=y_j} g(x_i) \cdot P(X = x_i) = \sum_i g(x_i) \cdot P(X = x_i)$$

Für stetige Zufallsvariable Y gilt

Substitution: $y = g(x), y_1 = g(x_1), y_2 = g(x_2)$

$$x = g^{-1}(y) = h(y), dx = h'(y)dy$$

$$\Rightarrow P(x_1 \le X \le x_2) \underset{4.14}{\overset{\text{Def.}}{=}} \int_{x_1}^{x_2} f(x)dx \underset{y=g(x)}{\overset{\text{Subst.}}{=}} \int_{y_1}^{y_2} f(h(y)) \cdot h'(y)dy$$

Sei $g'(x) > 0 \ \forall x \in [x_1, x_2] \Rightarrow h'(y) > 0$ und $y_2 > y_1$ mit

$$\int_{y_1}^{y_2} f(h(y)) \cdot h'(y)dy = \int_{y_1}^{y_2} f(h(y)) \cdot |h'(y)|dy =$$

$$= P(y_1 \le Y \le y_2)$$

Sei $g'(x) < 0 \ \forall x \in [x_1, x_2] \Rightarrow h'(y) < 0$ und $y_2 < y_1$ mit

$$\int_{y_1}^{y_2} f(h(y)) \cdot h'(y)dy = -\int_{y_2}^{y_1} f(h(y)) \cdot h'(y)dy =$$

$$= -\int_{y_2}^{y_1} f(h(y)) \cdot |h'(y)|dy = P(y_2 \le Y \le y_1)$$

$\Rightarrow f(h(y)) \cdot |h'(y)|$ ist Dichte von Y.

$$\Rightarrow E(Y) = E(g(X)) \underset{5.2}{\overset{\text{Def.}}{=}} \int_{-\infty}^{\infty} y \cdot f(h(y)) \cdot |h'(y)|dy \underset{\text{subst.}}{\overset{\text{Rück=}}{=}} \int_{-\infty}^{\infty} g(x) \cdot f(x)dx$$

□

SATZ 5.4 *Erwartungswert einer linearen Transformation*

Sei Y eine Zufallsvariable mit $Y(y_i) = a + b \cdot X(y_i), y_i \in \Omega$. Falls $E(X) < \infty$, so gilt

$$E(a + b \cdot X) = a + b \cdot E(X) \ \forall a, b \in \mathbb{R}$$

BEWEIS:

Für diskrete Zufallsvariable Y gilt

$$E(Y) = E(a + b \cdot X) \underset{5.3}{\overset{\text{Satz}}{=}} \sum_i (a + b \cdot x_i) \cdot P(X = x_i) =$$

$$= a \cdot \underbrace{\sum_i P(X = x_i)}_{\underset{4.10\,(ii)}{\overset{\text{Prop.}}{=}} 1} + b \cdot \underbrace{\sum_i x_i \cdot P(X = x_i)}_{\underset{5.1}{\overset{\text{Def.}}{=}} E(X)} = a + b \cdot E(X)$$

Für stetige Zufallsvariable Y gilt

$$E(Y) = E(a + b \cdot X) \underset{5.3}{\overset{\text{Satz}}{=}} \int_{-\infty}^{\infty} (a + b \cdot x) \cdot f(x)dx =$$

$$= a \cdot \underbrace{\int_{-\infty}^{\infty} f(x)dx}_{\underset{4.13\,(ii)}{\overset{\text{Def.}}{=}} 1} + b \cdot \underbrace{\int_{-\infty}^{\infty} x \cdot f(x)dx}_{\underset{5.2}{\overset{\text{Def.}}{=}} E(X)} = a + b \cdot E(X)$$

□

Bedeutung der linearen Transformation von $E(X)$:

- ♠ Da eine Verschiebung der Zufallsvariablen X um den Faktor a eine Verschiebung des $E(X)$ ebenfalls um den Faktor a bewirkt, ist $E(X)$ *lageäquivalent*.
- ♠ Da der Skalenfaktor b bei $E(X)$ unverändert bleibt, ist $E(X)$ *skalenäquivalent*.

Der Erwartungswert einer Zufallsvariablen ist ein Lageparameter, da er lage- und skalenäquivalent ist.

SATZ 5.5 *Erwartungswert einer zentrierten Transformation*

Sei $X-\mu$ eine zentrierte Zufallsvariable. Dann gilt für ihren Erwartungswert

$$E(X-\mu)=0$$

BEWEIS:

$$E(X-\mu) \underset{5.4}{\overset{\text{Satz}}{=}} \underbrace{E(X)}_{=\mu}-\mu=0$$

□

DEFINITION 5.6 *symmetrisch verteilte Zufallsvariable*

Sei X Zufallsvariable und $a\in\mathbb{R}$ eine beliebige Zahl. Dann ist X symmetrisch verteilt zu a, falls

$$P(X\le a-b)=P(X\ge a+b)\ \forall b\in\mathbb{R}$$

Wenn X eine symmetrisch verteilte Zufallsvariable ist, dann haben $X-a$ und $-(X-a)$ die gleiche Verteilung

SATZ 5.7 *Erwartungswert symmetrisch verteilter Zufallsvariablen*

Sei X eine zu μ symmetrisch verteilte Zufallsvariable mit dem Erwartungswert $E(X)$. Dann gilt

$$E(X)=\mu$$

BEWEIS:

$$E(X-\mu) \overset{\text{Symmetrie}}{=} E(-X+\mu)$$

$$\underset{5.4}{\overset{\text{Satz}}{\Rightarrow}} E(X)-\mu=-E(X)+\mu$$

$$\Rightarrow 2\cdot E(X)=2\mu \Rightarrow E(X)=\mu$$

□

5.2 VARIANZ UND STANDARDABWEICHUNG

DEFINITION 5.8 *Varianz und Standardabweichung einer Zufallsvariablen*

Sei *X* eine Zufallsvariable mit $\mu = E(X) < \infty$, so nennt man

$$\sigma^2 = V(X) = E\left((X-\mu)^2\right)$$

die Varianz von *X* und

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

die Standardabweichung von *X*.

PROPOSITION 5.9 *Varianz einer diskreten / stetigen Zufallsvariablen*

Sei X eine diskrete Zufallsvariable mit der Wahrscheinlichkeitsverteilung $f(x_i) = P(X = x_i)$ bzw. eine stetige Zufallsvariable mit der Dichte $f(x)$. Falls $\mu = E(X) < \infty$, so ist die Varianz definiert als

$$V(X) = \begin{cases} \sum_i (x_i-\mu)^2 \cdot P(X=x_i) & \text{falls } X \text{ diskret ist} \\ \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)dx & \text{falls } X \text{ stetig ist} \end{cases}$$

SATZ 5.10 *Verschiebungssatz von Steiner*

Die Varianz lässt sich auch schreiben als

$$V(X) = E(X^2) - \mu^2$$

Der Verschiebungssatz ist für Berechnungen besser eignet als Definition 5.8

BEWEIS:

$$V(X) \underset{5.8}{\overset{\text{Def.}}{=}} E\left((X-\mu)^2\right) = E\left(X^2 - 2\mu\cdot X + \mu^2\right) \underset{5.4}{\overset{\text{Satz}}{=}}$$
$$= E(X^2) - 2\mu\cdot \underbrace{E(X)}_{=\mu} + \mu^2 = E(X^2) - \mu^2$$

□

SATZ 5.11 *Varianz einer linearen Transformation*

Sei Y eine Zufallsvariable mit $Y(y_i) = a + b\cdot X(y_i), y_i \in \Omega$. Falls $V(X) < \infty$, so gilt

$$V(a+b\cdot X) = b^2 \cdot V(X) \; \forall a,b \in \mathbb{R}$$

bzw.

$$\sigma(a+b\cdot X) = |b| \cdot \sigma(X) \; \forall a,b \in \mathbb{R}$$

BEWEIS:

$$V(a+b\cdot X) \underset{5.8}{\overset{\text{Def.}}{=}} E\left(\left(\cancel{a}+b\cdot X\cancel{-a}-b\mu\right)^2\right)=$$

$$=E\left(b^2\cdot X^2-2b^2\mu\cdot X+b^2\mu^2\right)\underset{5.4}{\overset{\text{Satz}}{=}}$$

$$=b^2\cdot E\left(X^2\right)-2b^2\mu\cdot\underbrace{E(X)}_{=\mu}+b^2\mu^2=$$

$$=b^2\cdot\left[E\left(X^2\right)-\mu^2\right]\underset{5.10}{\overset{\text{Satz}}{=}}b^2\cdot V(X)$$

□

Bedeutung der linearen Transformation von $V(X)$:

- ♠ Da die Konstante *a* keinen Einfluss auf die Varianz und die Standardabweichung hat, sind beide Parameter *lageinvariant*.
- ♠ Da die Konstante *b* nur bei der Varianz den Faktor b^2 bewirkt, ist die Standardabweichung im Unterschied zur Varianz *skalenäquivalent*.

SATZ 5.12 *Minimumseigenschaft des Erwartungswertes*

Sei *X* eine Zufallsvariable und $a\in\mathbb{R}$. Dann gilt

$$E\left((X-a)^2\right)\geq V(X)$$

Das ist gleichbedeutend mit

$$\min E\left((X-a)^2\right)=V(X)$$

Also ist der Erwartungswert jener Punkt, wo die mittlere quadratische Abweichung von *X* den kleinsten Wert annimmt.

BEWEIS:

$$E\left((X-a)^2\right)=E\left(\left((X-\mu)+(\mu-a)\right)^2\right)\underset{5.4}{\overset{\text{Satz}}{=}}$$

$$=\underbrace{E\left((X-\mu)^2\right)}_{\underset{5.8}{\overset{\text{Def.}}{=}}V(X)}+2\cdot\underbrace{E\left((X-\mu)\cdot(\mu-a)\right)}_{=E\left(\mu\cdot X-a\cdot X-\mu^2+a\mu\right)}+E\left((\mu-a)^2\right)\underset{5.4}{\overset{\text{Satz}}{=}}$$

$$=V(X)\cancel{+2\cdot(\mu-a)\cdot\underbrace{E(X)}_{=\mu}-2\mu\cdot(\mu-a)}+E\left((\mu-a)^2\right)\underset{5.4}{\overset{\text{Satz}}{=}}$$

$$=V(X)+(\mu-a)^2$$

Diese Gleichung wird nun minimal für

$$a=\mu.$$

□

5.3 Parameter von mehrdimensionalen Zufallsvariablen

Satz 5.13 *Erwartungswert einer Funktion mehrerer Zufallsvariablen*

Sei $g:\mathbb{R}^n \to \mathbb{R}$ eine Funktion und $X_1, X_2, ..., X_n$ seien Zufallsvariable. Dann ist $Y = g(X_1, X_2, ..., X_n)$ ebenfalls eine Zufallsvariable und es gilt

$$E(Y) = E\big(g(X_1, X_2, ..., X_n)\big) =$$

$$= \begin{cases} \sum\limits_{(x_1, \ ..., x_n)\in X(\Omega)} g\cdot(x_1, \ ..., x_n)\cdot P(X_1 = x_1, ..., X_n = x_n) & \text{falls } X \text{ diskret ist} \\ \int\limits_{-\infty}^{\infty} g(x_1, \ ..., x_n)\cdot f(x_1, \ ..., x_n)\,dx_1 ... dx_n & \text{falls } X \text{ stetig ist} \end{cases}$$

Satz 5.14 $E(X)$ *und* $V(X)$ *einer Summe von zwei Zufallsvariablen*

Seien X, Y zwei beliebige Zufallsvariable. Dann gilt

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

(ii) $V(X+Y) = V(X) + V(Y) + 2\cdot E\big((X-\mu_X)\cdot(Y-\mu_Y)\big)$

Beweis:

(i) Für diskrete Zufallsvariable gilt

$$E(X+Y) \underset{5.13}{\overset{\text{Satz}}{=}} \sum_i \sum_j (x_i + y_j)\cdot P(X = x_i, Y = y_j) =$$

$$= \sum_i x_i \cdot \underbrace{\sum_j P(X = x_i, Y = y_j)}_{\underset{4.25}{\overset{\text{Def.}}{=}} P(X = x_i)} + \sum_j y_j \cdot \underbrace{\sum_i P(X = x_i, Y = y_j)}_{\underset{4.25}{\overset{\text{Def.}}{=}} P(Y = y_j)} \underset{5.1}{\overset{\text{Def.}}{=}}$$

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

Für stetige Zufallsvariable gilt

$$E(X+Y) \underset{5.13}{\overset{\text{Satz}}{=}} \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} (x+y)\cdot f(x,y)\,dxdy =$$

$$= \int\limits_{-\infty}^{\infty} xdx \cdot \underbrace{\int\limits_{-\infty}^{\infty} f(x,y)\,dy}_{\underset{4.29}{\overset{\text{Def.}}{=}} f_X(x)} + \int\limits_{-\infty}^{\infty} ydy \cdot \underbrace{\int\limits_{-\infty}^{\infty} f(x,y)\,dx}_{\underset{4.29}{\overset{\text{Def.}}{=}} f_Y(y)} \underset{5.2}{\overset{\text{Def.}}{=}}$$

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

(ii) Für diskrete / stetige Zufallsvariable gilt

$$V(X+Y) \overset{\text{Def.}}{\underset{5.8}{=}} E\left(\left(X+Y-E(X+Y)\right)^2\right) \overset{\text{Satz}}{\underset{5.14\text{ (i)}}{=}}$$

$$= E\left(\left(X+Y-\underbrace{E(X)}_{=\mu_X}-\underbrace{E(Y)}_{=\mu_Y}\right)^2\right) =$$

$$= E\left(\left((X-\mu_X)+(Y-X-\mu_Y)\right)^2\right) \overset{\text{Satz}}{\underset{5.14\text{ (i)}}{=}}$$

$$= E\left((X-\mu_X)^2\right) + 2\cdot E\left((X-\mu_X)\cdot(Y-\mu_Y)\right) + E\left((Y-\mu_Y)^2\right) \overset{\text{Def.}}{\underset{5.8}{=}}$$

$$= V(X) + V(Y) + 2\cdot E\left((X-\mu_X)\cdot(Y-\mu_Y)\right)$$

□

DEFINITION 5.15 *Kovarianz und Korrelationskoeffizient*

Der Ausdruck

$$Cov(X,Y) = E\left((X-\mu_X)\cdot(Y-\mu_Y)\right)$$

wird als Kovarianz der beiden Zufallsvariablen X und Y bezeichnet. Der Korrelationskoeffizient

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{V(X)\cdot V(Y)}}$$

ist eine auf die Wurzel des Produktes der beiden Standardabweichungen von X und Y bezogene Zahl zwischen -1 und 1.

SATZ 5.16 Seien X, Y zwei Zufallsvariable. Falls $E(X) < \infty$ und $E(Y) < \infty$, dann gilt

$$X \leq Y \Rightarrow E(X) \leq E(Y)$$

BEWEIS:

Sei $Z \geq 0$ eine Zufallsvariable mit $Z := Y - X$

$$\Rightarrow 0 \leq E(Z) = E(Y-X) \overset{\text{Satz}}{\underset{5.14\text{ (i)}}{=}} E(Y) - E(X)$$

$$E(X) \leq E(Y)$$

□

SATZ 5.17 *Verallgemeinerung von Satz 5.14*

Seien $X_1, X_2, ..., X_n$ beliebige Zufallsvariable. Dann gilt

(i) $E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$

(ii) $V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) + 2\cdot\sum_{i=1}^{n}\sum_{j=i+1}^{n} Cov(X_i, Y_j)$

BEWEIS:

Vollständige Induktion

□

SATZ 5.18 *Eigenschaften von unabhängigen Zufallsvariablen*
Seien X, Y unabhängige Zufallsvariable. Dann gilt

(i) $E(X \cdot Y) = E(X) \cdot E(Y)$

(ii) $Cov(X,Y) = 0$

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

BEWEIS:

(i) Für diskrete Zufallsvariable gilt

$$E(X \cdot Y) \overset{\text{Satz}}{\underset{5.13}{=}} \sum_i \sum_j x_i \cdot y_j \cdot P(X = x_i, Y = y_j) \overset{\text{Def.}}{\underset{4.30}{=}}$$

$$= \sum_i \sum_j x_i \cdot y_j \cdot P(X = x_i) \cdot (Y = y_j) =$$

$$= \sum_i x_i \cdot P(X = x_i) \cdot \sum_j y_j \cdot P(Y = y_j) \overset{\text{Def.}}{\underset{5.1}{=}} E(X) \cdot E(Y)$$

Für stetige Zufallsvariable gilt

$$E(X \cdot Y) \overset{\text{Satz}}{\underset{5.13}{=}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \cdot f(x,y) dxdy \overset{\text{Def.}}{\underset{4.30}{=}}$$

$$= \int_{-\infty}^{\infty} x \cdot f(x) dx \cdot \int_{-\infty}^{\infty} y \cdot f(y) dx \overset{\text{Def.}}{\underset{5.2}{=}} E(X) \cdot E(Y)$$

(ii) Für diskrete / stetige Zufallsvariable gilt

$$Cov(X,Y) \overset{\text{Def.}}{\underset{5.15}{=}} E\big((X - E(X)) \cdot (Y - E(Y))\big) \overset{\text{Satz}}{\underset{5.17\text{ (i)}}{=}}$$

$$= E(X \cdot Y) - E(X \cdot E(Y)) - E(E(X) \cdot Y) +$$

$$+ E(E(X) \cdot E(Y)) \overset{\text{Satz}}{\underset{5.4}{=}}$$

$$= E(X \cdot Y) - E(X) \cdot E(Y) \cancel{- E(X) \cdot E(Y)} \cancel{+ E(X) \cdot E(Y)}$$

$$X,Y \text{ sind unabhängig} \overset{\text{Satz}}{\underset{5.18\text{ (i)}}{\Rightarrow}} E(X \cdot Y) = E(X) \cdot E(Y)$$

$$\Rightarrow Cov(X,Y) = 0$$

(iii) Folgt unmittelbar aus (ii) sowie Satz 5.14 (ii)

□

SATZ 5.19 *Verallgemeinerung von Satz 5.18*
Seien $X_1, X_2, ..., X_n$ beliebige unabhängige Zufallsvariable. Dann gilt

(i) $E\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} E(X_i)$

(ii) $\sum_{i=1}^{n} \sum_{j=i+1}^{n} Cov(X_i, Y_j) = 0$

(iii) $V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i)$

BEWEIS:

Vollständige Induktion

□

5.4 DIE UNGLEICHUNG VON TSCHEBYSCHEW

SATZ 5.20 *Ungleichung von Tschebyschew*

Sei X eine Zufallsvariable mit dem Erwartungswert $E(X)=\mu$ und der Varianz $V(X)=\sigma^2$. Dann gilt $\forall\varepsilon>0$

$$P\left(\left|X-\mu\right|\geq\varepsilon\right)\leq\frac{\sigma^2}{\varepsilon^2}$$

bzw. die Gegenwahrscheinlichkeit

$$P\left(\left|X-\mu\right|<\varepsilon\right)=P\left(\mu-\varepsilon<X<\mu+\varepsilon\right)\geq 1-\frac{\sigma^2}{\varepsilon^2}$$

BEWEIS:

Sei Z eine Zufallsvariable, definiert durch

$$Z(\omega)=\begin{cases}\varepsilon^2 \text{ falls } \left|X(\omega)-E(X)\right|\geq\varepsilon\\ 0 \text{ sonst}\end{cases}.$$

Einerseits gilt

$$Z(\omega)\leq\left|X(\omega)-E(X)\right|^2$$

$$\Rightarrow E(Z)\underset{5.16}{\overset{\text{SATZ}}{\leq}}E\left(\left|X(\omega)-E(X)\right|^2\right)\underset{5.8}{\overset{\text{Def.}}{=}}\sigma^2$$

Andererseits gilt aber auch

$$E(Z)=\varepsilon^2\cdot P\left(\left|X(\omega)-E(X)\right|\geq\varepsilon\right)+0\cdot P\left(\left|X(\omega)-E(X)\right|<\varepsilon\right)=$$
$$\varepsilon^2\cdot P\left(\left|X(\omega)-E(X)\right|\geq\varepsilon\right).$$

Somit folgt

$$E(Z)=\varepsilon^2\cdot P\left(\left|X(\omega)-E(X)\right|\geq\varepsilon\right)\leq\sigma^2$$
$$\Rightarrow P\left(\left|X(\omega)-E(X)\right|\geq\varepsilon\right)\leq\frac{\sigma^2}{\varepsilon^2}.$$

□

Da bekanntlich $0\leq P(X=x_i)\leq 1$ gilt, ist die Ungleichung von Tschebyschew offensichtlich nur dann sinnvoll, wenn

$$\varepsilon^2\geq\sigma^2$$

gilt.

BEISPIEL 5.21 *Abgabe eines Stimmzettels*

Sei X die Uhrzeit der Abgabe eines Stimmzettels am Tag einer Wahl. Die durchschnittliche Uhrzeit sei 11^{30}, die Standardabweichung betrage 2 Stunden. Wie groß ist dann mindestens die Wahrscheinlichkeit, dass ein zufällig ausgewählter Wähler zwischen 8^{00} und 15^{00} Uhr den Stimmzettel in die Urne geworfen hat?

$$P(8<X<15)=P\left(\left|X-11{,}5\right|<3{,}5\right)\geq 1-\frac{2^2}{3{,}5^2}=0{,}6735$$

BEMERKUNG 5.22 *k-Sigma-Regel*

Sei $\varepsilon = k \cdot \sigma$. Dann lässt sich die Ungleichung von Tschebyschew anschreiben als

$$P\left(|X-\mu| \geq k \cdot \sigma\right) \leq \frac{1}{k^2}$$

bzw.

$$P\left(|X-\mu| < k \cdot \sigma\right) = P\left(\mu - k \cdot \sigma < X < \mu + k \cdot \sigma\right) \geq 1 - \frac{\sigma^2}{k^2}$$

Spezielle Werte für k:

- ♠ $P\left(\mu - 2 \cdot \sigma < X < \mu + 2 \cdot \sigma\right) \geq \frac{3}{4} \mathrel{\hat{=}} 75\%$ (*Zwei-Sigma-Regel*)
- ♠ $P\left(\mu - 3 \cdot \sigma < X < \mu + 3 \cdot \sigma\right) \geq \frac{8}{9} \mathrel{\hat{=}} 88{,}89\%$ (*Drei-Sigma-Regel*)
- ♠ $P\left(\mu - 4 \cdot \sigma < X < \mu + 4 \cdot \sigma\right) \geq \frac{15}{16} \mathrel{\hat{=}} 93{,}75\%$ (*Vier-Sigma-Regel*)
- ♠ $P\left(\mu - 5 \cdot \sigma < X < \mu + 5 \cdot \sigma\right) \geq \frac{24}{25} \mathrel{\hat{=}} 96\%$ (*Fünf-Sigma-Regel*)

Die *k-Sigma-Regel* gibt nur eine sehr grobe Abschätzung für die Wahrscheinlichkeit an, dass ein beliebiger Wert *x* in der $k \cdot \sigma -$Umgebung liegt. Dafür sind außer der Kenntnis über den Erwartungswert und die Varianz keine weiteren Informationen über die jeweilige Verteilung notwendig. Der Vorteil der Ungleichung von Tschebyschew liegt also in ihrer allgemeinen Gültigkeit für beliebige Verteilungen.

6. Normalverteilung

Dieses Kapitel steht ganz im Zeichen der Normalverteilung, der wichtigsten stetigen Zufallsvariablen. Die enorme Bedeutung der Normalverteilung wurde bereits im 19.Jahrhundert erkannt. Bei Untersuchungen des menschlichen Körpers ging Quetelet von einem Durchschnittsmenschen aus, also einem Normalmenschen. Deshalb auch die Bezeichnung Normalverteilung. Die enorme Tragweite der Normalverteilung wird sich dann später im zentralen Grenzwertsatz zeigen, den ich in Kapitel 7.3 besprechen werde. Dieser Satz besagt, dass die Summe von n identisch verteilten, unabhängigen Zufallsvariablen für große n annähernd normalverteilt ist.

Ich werde die Dichte- und Verteilungsfunktion der Normalverteilung einführen. Die Dichtefunktion der Normalverteilung geht auf Carl Friedrich Gauß zurück, der diese Kurve bei astronomischen Untersuchungen verwendet hat. Aufgrund ihrer Gestalt wird die Dichte einer Normalfunktion daher auch als *Gaußsche Glockenkurve* bezeichnet.

Ich werde den Erwartungswert und die Varianz der Normalverteilung bestimmen. Außerdem werde ich die Wahrscheinlichkeiten angeben, dass die Zufallsvariable X um höchstens das k-fache der Standardabweichung vom Erwartungswert abweicht. Dies ist ein Analogon zur Ungleichung von Tschebyschew aus Kapitel 5.4.

Ich werde in diesem Kapitel auch die Standardnormalverteilung behandeln, die das Berechnen einzelner Wahrscheinlichkeiten erleichtert. Da die Verteilungsfunktion einer Normalverteilung keine elementare Stammfunktion besitzt, ist nur eine näherungsweise Integration, etwa durch Taylorpolynome, möglich. Diese Berechnung ist händisch sehr mühsam, selbst für Computer ist der Rechenaufwand enorm. Deshalb nimmt man Tabellen zu Hilfe. Da jedoch μ und σ unendlich viele Werte annehmen können, führt man alle Aufgaben auf die so genannte Standardnormalverteilung zurück. Nun findet man alle möglichen Werte in einer einzigen Tabelle aufgelistet.

6.1 DICHTE UND VERTEILUNGSFUNKTION

DEFINITION 6.1 *Normalverteilung*

Eine Zufallsvariable X heißt *normalverteilt* mit den Parametern $\mu \in \mathbb{R}$ und $\sigma^2 > 0$ (kurz: $N(\mu, \sigma^2)$), falls ihre Dichte von der Form

$$f : \mathbb{R} \to \mathbb{R}, x \mapsto f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

ist. Die Verteilungsfunktion einer Normalverteilung ist durch

$$F(x) = P(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-(t-\mu)^2}{2\sigma^2}} dt, x \in \mathbb{R}$$

gegeben.

SATZ 6.2 *Eigenschaften der Dichtefunktion einer Normalverteilung*

(i) $f(x) > 0 \ \forall x \in \mathbb{R}$ und hat die x-Achse als Asymptote

(ii) $f(x) > 0$ hat ein Maximum in $x = \mu$ mit $f(x) = \frac{1}{\sigma\sqrt{2\pi}}$

(iii) $f(x)$ ist symmetrisch bezüglich μ

(iv) $f(x)$ hat zwei Wendepunkte in $x_{1,2} = \mu \pm \sigma$ mit $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{-1}{2}}$

BEWEIS:
Kurvendiskussion

□

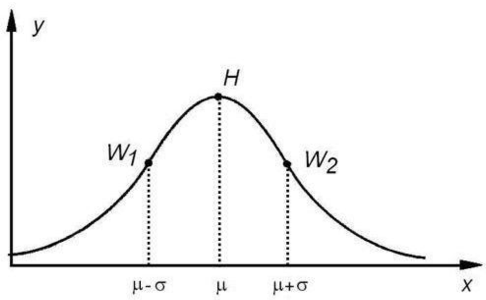


ABBILDUNG 11 Graph der Dichtefunktion mit Hochpunkt $H\left(\mu \middle| \frac{1}{\sigma\sqrt{2\pi}}\right)$ und den beiden Wendepunkten $W_1\left(\mu - \sigma \middle| \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{-1}{2}}\right), W_2\left(\mu + \sigma \middle| \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{-1}{2}}\right)$

ABBILDUNG 12 Deutscher Zehnmarkschein vor der Einführung des Euros im Jahr 2002 mit einem Portrait von Gauß und der Gaußschen Glockenkurve

Auswirkungen der Veränderung der Gaußschen Glockenkurve durch die Parameter μ und σ

- Eine Veränderung von μ (bei festem σ bewirkt ein Verschieben der Kurve entlang der *x*-Achse nach links (für negative μ) bzw. nach rechts (für positive μ).
- Eine Veränderung von σ (bei festem μ bewirkt eine Streckung (bei kleiner werdendem σ) bzw. Stauchung (bei größer werdendem σ) in *y*-Richtung.

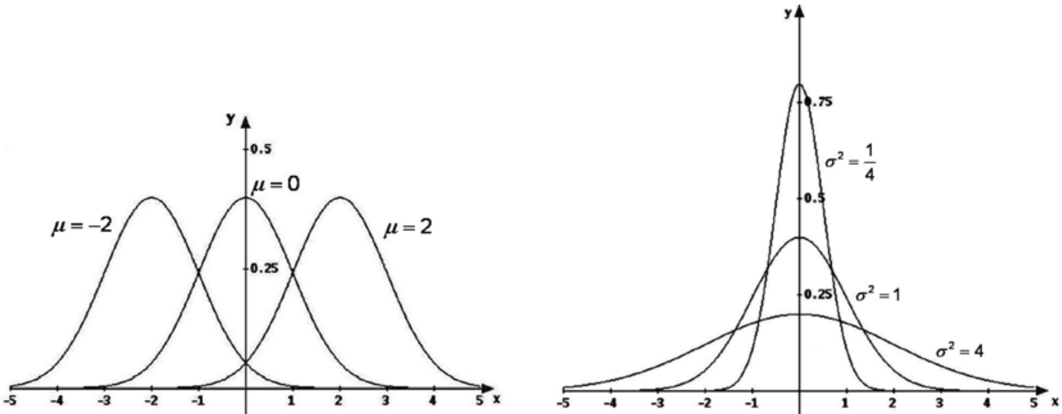


ABBILDUNG 13 Dichten von verschiedenen Normalverteilungen (links: $N(\mu,1)$ – Verteilung, rechts: $N(0,\sigma^2)$ – Verteilung)

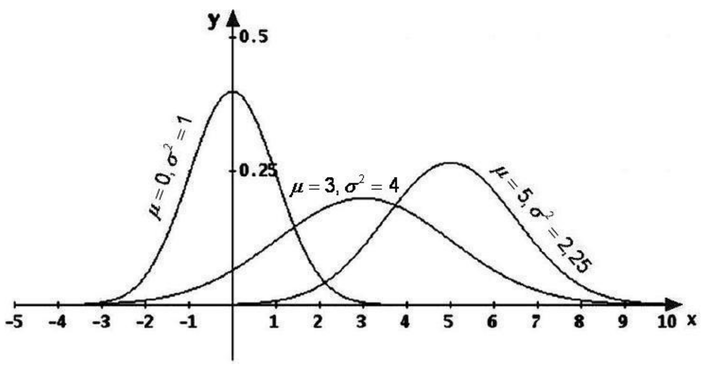


ABBILDUNG 14 Dichten von drei unterschiedlichen Normalverteilungen

6.2 DIE STANDARDNORMALVERTEILUNG

DEFINITION 6.3 *Standardnormalverteilung*

Eine Zufallsvariable Z heißt standardnormalverteilt $\left(\text{kurz: } Z \sim N(0,1)\right)$, falls sie die Dichte

$$\varphi : \mathbb{R} \to \mathbb{R}, z \mapsto \varphi(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}$$

hat. Die *Verteilungsfunktion der Standardnormalverteilung* wird mit

$$\Phi(z) = P(Z \le z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{\frac{-t^2}{2}} dt, z \in \mathbb{R}$$

bezeichnet.

SATZ 6.4 Sei $X \sim N(\mu, \sigma^2)$. Dann gilt

$$Z = \frac{X-\mu}{\sigma} \sim N(0,1)$$

und somit

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right), x \in \mathbb{R}.$$

BEWEIS:

$$F(z) = P\left(\frac{X-\mu}{\sigma} \le z\right) = P(X \le \sigma z + \mu) \underset{6.1}{\overset{\text{Def}}{=}}$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\sigma z+\mu} e^{\frac{-(t-\mu)^2}{2\sigma^2}} dt \underset{\frac{dt}{\sigma}=du}{\overset{\frac{t-\mu}{\sigma}=u}{=}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{\frac{-u^2}{2}} du \underset{6.3}{\overset{\text{Def}}{=}} \varphi(z)$$

□

Geometrisch betrachtet bewirkt die Substitution $\frac{t-\mu}{\sigma} = u$ keine Änderung des Flächeninhaltes der Gaußschen Glockenkurve.

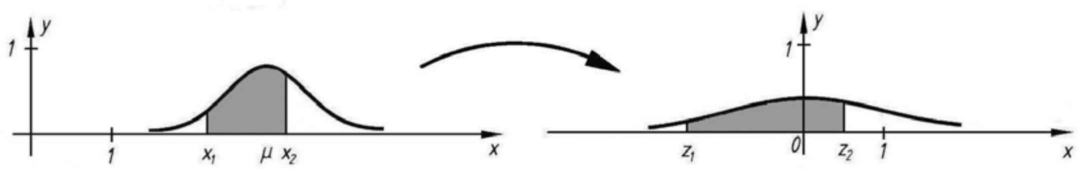


ABBILDUNG 15 Affin-lineare Transformation der Normalverteilung $N(3; 0{,}5)$ (links) in die Standardnormalverteilung $N(0; 1)$ (rechts)[11]

[11] REICHEL, Hans-Christian / MÜLLER, Robert / HANISCH, Günter, Lehrbuch der Mathematik 8, Wien [2]1993, S.111

PROPOSITION 6.5 Der Flächeninhalt unter $\varphi(x)$ (und somit unter jeder Gaußschen Glockenkurve) ist 1. Also ist $\varphi(x)$ tatsächlich eine Dichtefunktion.

BEWEIS:

Existenz des Integrals:

$$e^{\frac{-x^2}{2}} = e^{\frac{-(-x)^2}{2}}$$

$$\Rightarrow \int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx = 2 \cdot \int_{0}^{\infty} e^{\frac{-x^2}{2}} dx = 2 \cdot \left[\int_{0}^{1} e^{\frac{-x^2}{2}} dx + \int_{1}^{\infty} e^{\frac{-x^2}{2}} dx \right]$$

$$e^{\frac{-x^2}{2}} \text{ ist stetig auf } [0,1] \Rightarrow \int_{0}^{1} e^{\frac{-x^2}{2}} dx < \infty$$

$$\text{Für } x \geq 1 \text{ ist } \frac{x^2}{2} \geq \frac{x}{2} \Rightarrow e^{\frac{-x^2}{2}} \leq e^{\frac{-x}{2}}$$

$$\Rightarrow \int_{1}^{\infty} e^{\frac{-x^2}{2}} dx \leq \int_{1}^{\infty} e^{\frac{-x}{2}} dx = \lim_{b\to\infty} \int_{1}^{b} e^{\frac{-x}{2}} dx = \lim_{b\to\infty} \left[2e^{\frac{-1}{2}} - 2e^{\frac{-b}{2}} \right] = 2e^{\frac{-1}{2}} < \infty$$

Wert des Integrals:

$$\left(\int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx \right)^2 = \int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx \cdot \int_{-\infty}^{\infty} e^{\frac{-y^2}{2}} dy \overset{\text{Existenz}}{=} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{\frac{-(x^2+y^2)}{2}} dxdy \underset{\text{koordinaten}}{\overset{\text{Polar-}}{=}}$$

$$= \int_{0}^{2\pi} \int_{0}^{\infty} r \cdot e^{\frac{-r^2}{2}} drd\varphi = \underbrace{\int_{0}^{2\pi} 1 d\varphi}_{=2\pi} \cdot \underbrace{\left(-\int_{0}^{\infty} -r \cdot e^{\frac{-r^2}{2}} dr \right)}_{= \lim_{b\to\infty} \left(e^0 - e^{\frac{-b^2}{2}} \right) = 1} = 2\pi$$

$$\int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} dx = \sqrt{2\pi} \Rightarrow \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} dx = 1$$

□

PROPOSITION 6.6 Sei $Z \sim N(0,1)$. Dann gilt die so genannte *Negativitäts-regel*

$$\Phi(-z) = 1 - \Phi(z),\ z \geq 0$$

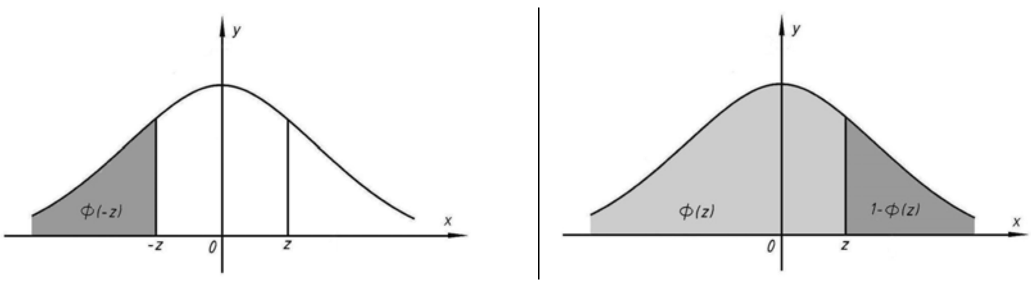


ABBILDUNG 16 Graphische Veranschaulichung der Negativitätsregel

BEWEIS:

$$\Phi(-z) \overset{\text{Def.}}{\underset{6.3}{=}} P(Z \le -z) \overset{\text{Symmetrie}}{=} P(Z \ge z) \overset{\text{Satz}}{\underset{1.16\,(iii)}{=}}$$

$$= 1 - P(Z < z) \overset{\text{Prop.}}{\underset{4.18}{=}} 1 - P(Z \le z) \overset{\text{Def.}}{\underset{6.3}{=}} 1 - \Phi(z)$$

□

PROPOSITION 6.7 Sei $Z \sim N(0,1)$. Dann gilt für beliebige $z \in \mathbb{R}$

(i) $P(Z \le z) = \Phi(z)$

(ii) $P(Z \ge z) = 1 - \Phi(z)$

(iii) $P(|Z| \le z) = 2 \cdot \Phi(z) - 1$

(iv) $P(|Z| \ge z) = 2 \cdot (1 - \Phi(z))$

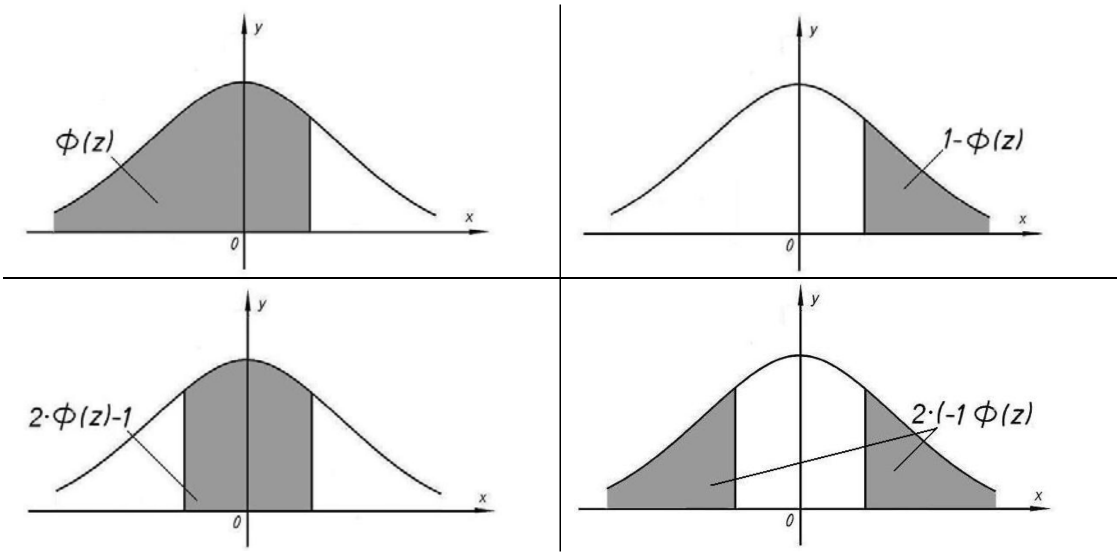


ABBILDUNG 17 Graphische Veranschaulichung der vier Standard- Wahrscheinlichkeitsbereiche, $z > 0$

BEWEIS:

(i) Definition 6.3

(ii) $$P(Z \le z) + P(Z \ge z) \overset{\text{Prop.}}{\underset{6.5}{=}} 1$$

$$\Rightarrow P(Z \ge z) = 1 - P(Z \le z) \overset{\text{Def.}}{\underset{6.3}{=}} 1 - \Phi(z)$$

(iii) $$P(|Z| \le z) = P(-z \le Z \le z) \overset{\text{Bem.}}{\underset{4.6\,(viii)}{=}} \Phi(z) - \Phi(-z) \overset{\text{Prop.}}{\underset{6.6}{=}}$$

$$= \Phi(z) - (1 - \Phi(z)) = 2 \cdot \Phi(z) - 1$$

(iv) $$P(|Z| \ge z) + P(|Z| \le z) \overset{\text{Prop.}}{\underset{6.5}{=}} 1$$

$$\Rightarrow P(|Z| \ge z) = 1 - P(|Z| \le z) \overset{\text{Prop.}}{\underset{6.7\,(iii)}{=}} 1 - (2 \cdot \Phi(z) - 1) = 2 \cdot (1 - \Phi(z))$$

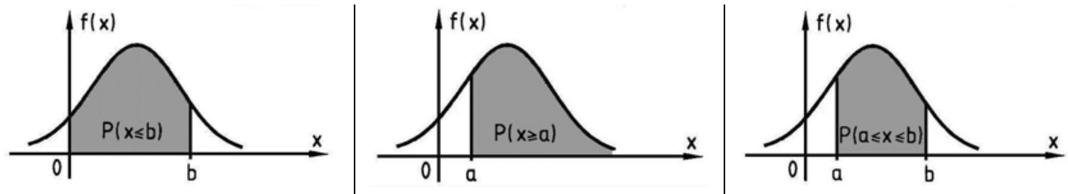
□

KOROLLAR 6.8 Sei $X \sim N\left(\mu, \sigma^2\right)$. Dann gilt für beliebige $a, b \in \mathbb{R}$

(i) $P\left(X \leq b\right) = \Phi\left(\dfrac{b-\mu}{\sigma}\right)$

(ii) $P\left(X \geq a\right) = 1 - \Phi\left(\dfrac{a-\mu}{\sigma}\right)$

(iii) $P\left(a \leq X \leq b\right) = \Phi\left(\dfrac{b-\mu}{\sigma}\right) - \Phi\left(\dfrac{a-\mu}{\sigma}\right)$



$$P\left(X \leq b\right) = \int_{-\infty}^{b} f\left(t\right) dt = F\left(b\right)$$

$$P\left(X \geq a\right) = \int_{a}^{\infty} f\left(t\right) dt = 1 - F\left(a\right)$$

$$P\left(a \leq X \leq b\right) = \int_{a}^{b} f\left(t\right) dt = F\left(b\right) - F\left(a\right)$$

ABBILDUNG 18 Zusammenhang zwischen verschiedenen Wahrscheinlichkeiten und Flächeninhalten

6.3 Erwartungswert und Varianz

Satz 6.9 Sei $Z \sim N(0,1)$. Dann gilt für den Erwartungswert und die Varianz

$$E(Z)=0, V(Z)=1$$

Beweis:

$$E(Z)=\int_{-\infty}^{\infty} x\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{\frac{-x^2}{2}}dx=\frac{-1}{\sqrt{2\pi}}\cdot\int_{-\infty}^{\infty} -x\cdot e^{\frac{-x^2}{2}}dx=$$

$$=\frac{-1}{\sqrt{2\pi}}\cdot\left[\underbrace{\lim_{b\to\infty} e^{\frac{-b^2}{2}}}_{=0}-\underbrace{\lim_{a\to\infty} e^{\frac{-a^2}{2}}}_{=0}\right]=0$$

$$V(Z)\overset{\text{SATZ}}{\underset{5.10}{=}}E(Z^2)-\underbrace{\left(E(Z)\right)^2}_{=0^2=0}=E(Z^2)=$$

$$=\int_{-\infty}^{\infty} x^2\cdot\frac{1}{\sqrt{2\pi}}\cdot e^{\frac{-x^2}{2}}dx=\frac{1}{\sqrt{2\pi}}\cdot\int_{-\infty}^{\infty} x\cdot\left(x\cdot e^{\frac{-x^2}{2}}\right)dx\overset{\text{partielle}}{\underset{\text{Integration}}{=}}$$

$$=\underbrace{\left[-\frac{1}{\sqrt{2\pi}}\cdot x\cdot e^{\frac{-x^2}{2}}\right]_{-\infty}^{\infty}}_{=0}+\frac{1}{\sqrt{2\pi}}\cdot\underbrace{\int_{-\infty}^{\infty} e^{\frac{-x^2}{2}}dx}_{\overset{\text{Prop.}}{\underset{6.5}{=}}\sqrt{2\pi}}=1$$

□

Satz 6.10 Sei $X \sim N(\mu,\sigma^2)$. Dann gilt für den Erwartungswert und die Varianz

$$E(X)=\mu, V(X)=\sigma$$

Beweis:

$$E(X)=E(\mu+\sigma Z)\overset{\text{Satz}}{\underset{5.4}{=}}\mu+\sigma\cdot\underbrace{E(Z)}_{\overset{\text{Satz}}{\underset{6.9}{=}}0}=\mu$$

$$V(X)=V(\mu+\sigma Z)\overset{\text{SATZ}}{\underset{5.11}{=}}\sigma^2\cdot\underbrace{V(Z)}_{\overset{\text{Satz}}{\underset{6.9}{=}}1}=\sigma^2$$

□

SATZ 6.11 Sei $X \sim N\left(\mu, \sigma^2\right)$. Dann erhält man die folgenden Wahrscheinlichkeiten dafür, dass X vom Erwartungswert μ um nicht mehr als das k-fache $(k = 1, 2, 3)$ der Standardabweichung σ abweicht

(i) $P(\mu - \sigma \leq X \leq \mu + \sigma) = 0{,}6826$

(ii) $P(\mu - 2 \cdot \sigma \leq X \leq \mu + 2 \cdot \sigma) = 0{,}9546$

(iii) $P(\mu - 3 \cdot \sigma \leq X \leq \mu + 3 \cdot \sigma) = 0{,}9974$

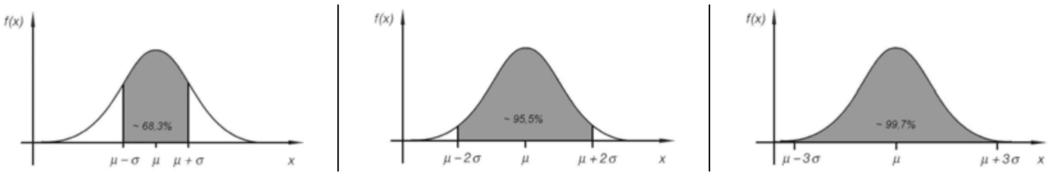


ABBILDUNG 19 Bedeutung der Standardabweichung anhand bestimmter Intervalle (links: $[\mu - \sigma; \mu + \sigma]$, Mitte: $[\mu - 2\sigma; \mu + 2\sigma]$, rechts: $[\mu - 3\sigma; \mu + 3\sigma]$)

BEWEIS:

(i) $P(\mu - \sigma \leq X \leq \mu + \sigma) \underset{6.7\text{ (iii)}}{\overset{\text{Prop.}}{=}} 2 \cdot \Phi(1) - 1 \cong 2 \cdot 0{,}8413 - 1 = 0{,}6826$

(ii) $P(\mu - 2 \cdot \sigma \leq X \leq \mu + 2 \cdot \sigma) \underset{6.7\text{ (iii)}}{\overset{\text{Prop.}}{=}} 2 \cdot \Phi(2) - 1 \cong 2 \cdot 0{,}9773 - 1 = 0{,}9546$

(iii) $P(\mu - 3 \cdot \sigma \leq X \leq \mu + 3 \cdot \sigma) \underset{6.7\text{ (iii)}}{\overset{\text{Prop.}}{=}} 2 \cdot \Phi(3) - 1 \cong 2 \cdot 0{,}9987 - 1 = 0{,}9974$

$\square$

Vergleich mit der *Ungleichung von Tschebyschew* (siehe Bemerkung 5.22)

$$P(\mu - 2 \cdot \sigma < X < \mu + 2 \cdot \sigma) \geq \frac{3}{4} \mathrel{\hat{=}} 75\,\%$$

$$P(\mu - 3 \cdot \sigma < X < \mu + 3 \cdot \sigma) \geq \frac{8}{9} \mathrel{\hat{=}} 88{,}89\,\%$$

Die Abweichungen einer $N\left(\mu, \sigma^2\right)$-verteilten Zufallsvariable vom Erwartungswert sind also erheblich geringer als die Abweichungen einer beliebigen Verteilung mithilfe der Tschebyschewschen Ungleichung.

7. BINOMIALVERTEILUNG

Die Binomialverteilung ist die wichtigste diskrete Verteilung. In diesem Kapitel möchte ich diese Verteilung neben der Bernoulli-Verteilung besprechen, mit deren Hilfe sich die Kenngrößen der Binomialverteilung am Einfachsten beweisen lassen.

Bei einem Bernoulli-Experiment sind nur zwei unterschiedliche Ergebnisse möglich, entweder ein Ereignis A tritt ein oder es tritt nicht ein. Die Bernoulli-Verteilung ist ein Spezialfall der Binomialverteilung für $n = 1$

Mit Hilfe der Binomialverteilung lässt sich bestimmen, wie groß die Wahrscheinlichkeit ist, wie oft ein Bernoulli-Experiment in einer Serie von gleichartigen und unabhängigen Versuchen eintritt.

Der zentrale Grenzwertsatz beschreibt ein Phänomen, das wohl das bedeutendste Resultat der Wahrscheinlichkeitstheorie ist und das auch in der Statistik eine entscheidende Rolle spielt. Dieser Satz ist die Grundlage für die Verwendung der Normalverteilung in vielen praktischen Anwendungen, da die Summe einer großen Anzahl von unabhängigen Zufallsvariablen annähernd normalverteilt ist.

Im Anschluss an die exakte Formulierung dieses Satzes werde ich den Spezialfall für Bernoulli-verteilte Zufallsvariable, den Satz von De Moivre-Laplace, beweisen. Dadurch wird die Approximation der Binomial- durch die Normalverteilung legitimiert.

Ich werde diesen Satz dann später in Kapitel 13 verwenden, wenn es darum geht, ein mathematisches Modell für Wahlumfragen zu erstellen.

7.1 Bernoulli-Verteilung

Definition 7.1 *Bernoulli-Experiment*

Unter einem *Bernoulli-Experiment* versteht man einen Zufallsvorgang, bei dem nur zwei Ausgänge möglich sind, nämlich ein Ereignis *A* tritt ein oder nicht.

Die Zufallsvariable *X* eines *Bernoulli-Experiments* wird folgendermaßen definiert

$$X = \begin{cases} 1 \text{ falls } A \text{ eintritt} \\ 0 \text{ falls } A \text{ nicht eintritt} \end{cases}$$

Beispiel 7.2 *Beispiele für Bernoulli-Experimente*

- ♠ Wähler oder Nichtwähler
- ♠ EU-Befürworter oder EU-Skeptiker
- ♠ Pro oder Contra Volksabstimmung
- ♠ Stamm- oder Wechselwähler
- ♠ Regierungs- oder Oppositionsanhänger
- ♠ Gutheißen eines Parteiprogramms oder Ablehnung desselbigen
- ♠ Beliebige Frage mit ja oder nein beantworten
- ♠ Politikinteressiert oder –verdrossen
- ♠ Wähler ist Mann oder Frau

Sei eine Versuchsserie von *n* Versuchen gegeben. Man betrachte nun ein beliebiges Ereignis *A* und interessiert sich dafür, wie oft *A* bei den *n* Versuchen eingetreten ist.

Definition 7.3 *Bernoulli-Verteilung*

Sei $p \in [0,1]$. Dann heißt eine Zufallsvariable *X Bernoulli-verteilt mit Parameter p*, falls

$$P(X = k) = p^k \cdot (1-p)^{1-k} \quad \forall k \text{ in } \{0,1\}$$

Die *Bernoulli-Verteilung* gibt also die Wahrscheinlichkeit an, dass ein Ereignis *A* genau *k* mal eintritt.

Proposition 7.4 Die Bernoulli-Verteilung ist eine Wahrscheinlichkeitsfunktion.

BEWEIS:

(i) $p(x_i) \geq 0 \; \forall x_i$

Klar, da laut Definition 7.3 $p \in [0,1]$

(ii) $\sum_{i=0}^{\infty} p(x_i) = 1$

$$\sum_{i=0}^{\infty} p(x_i) \overset{\text{Def.}}{\underset{7.3}{=}} P(X=0) + P(X=1) = (1-p) + p = 1$$

□

> **SATZ 7.5** *Kenngrößen einer Bernoulli-Verteilung*
>
> $$E(X) = p, V(X) = p \cdot (1-p)$$

BEWEIS:

$$E(X) \overset{\text{Def.}}{\underset{5.1}{=}} \sum_{k=0}^{1} k \cdot P(X=k) = 1 \cdot p + 0 \cdot (1-p) = p$$

$$V(X) \overset{\text{Satz}}{\underset{5.10}{=}} \underbrace{E(X^2)}_{=0^2 \cdot (1-p) + 1^2 \cdot p} - \underbrace{E(X)^2}_{=p^2} = p - p^2 = p \cdot (1-p)$$

□

7.2 EIGENSCHAFTEN DER BINOMIALVERTEILUNG

DEFINITION 7.6 *Binomialverteilung*

Sei $p \in [0,1]$. Dann heißt eine Zufallsvariable *X Binomialverteilt mit Parameter p*, falls

$$P(X=k) = \binom{n}{k} \cdot p^k \cdot \underbrace{(1-p)}_{=q}^{n-k} \quad \forall k \text{ in } \mathbb{N}_0, \forall n \text{ in } \mathbb{N}$$

PROPOSITION 7.7 Die Binomialverteilung ist eine Wahrscheinlichkeitsfunktion.

BEWEIS:

(i) $p(x_i) \geq 0 \; \forall x_i$

Klar, da laut Definition 7.6 $p \in [0,1]$

(ii) $\sum_{i=0}^{\infty} p(x_i) = 1$

$$\sum_{i=0}^{\infty} P(X=k) \underset{7.6}{\overset{\text{Def.}}{=}} \sum_{i=0}^{\infty} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \underset{2.18}{\overset{\text{Satz}}{=}} \left[p + (1-p)\right]^n = 1$$

□

SATZ 7.8 *Kenngrößen einer Binomialverteilung*

$$E(X) = np, V(X) = np(1-p)$$

BEWEIS:

Seien $X_1, X_2, ..., X_n$ unabhängige Zufallsvariable, definiert durch

$$X_{i,\, i\in\mathbb{N}} = \begin{cases} 1 \text{ falls beim i-ten Experiment } A \text{ eintritt} \\ 0 \text{ sonst} \end{cases}$$

Für den Erwartungswert und die Varianz dieser Bernoulli-verteilten Zufallsvariablen gilt

$$E(X_i) = p, E\left(X_i^2\right) = 0^2 \cdot (1-p) + 1^2 \cdot p = p$$

$$V(X_i) \underset{5.10}{\overset{\text{Satz}}{=}} E\left(X_i^2\right) - E(X_i)^2 = p - p^2 = p \cdot (1-p)$$

Aus der Additivität des Erwartungswertes und der Varianz bei unabhängigen Zufallsvariablen folgt

$$E(X) = E\left(\sum_{i=1}^{n} X_i\right) \underset{5.17\,(i)}{\overset{\text{Satz}}{=}} \sum_{i=1}^{n} E(X_i) = np$$

$$V(X) = V\left(\sum_{i=1}^{n} X_i\right) \underset{5.19\,(iii)}{\overset{\text{Satz}}{=}} \sum_{i=1}^{n} V(X_i) = np(1-p)$$

□

7.3 Approximation der Binomial- durch die Normalverteilung

Ausgangspunkt unserer Überlegung seien n unabhängige, identisch verteilte Zufallsvariable $X_1, X_2, ..., X_n$ mit $E(X_i) = \mu$ und $V(X_i) = \sigma^2$. Sei weiters $Z_n = X_1 + X_2 + ... + X_n$ mit $E(Z_n) = n\mu$ und $V(Z_n) = n\sigma^2$. Somit besitzt die standardisierte Zufallsvariable $Z_n^* := \frac{Z_n - n\mu}{\sigma\sqrt{n}}$ den Erwartungswert

$$E(Z_n^*) = E\left(\frac{Z_n - n\mu}{\sigma\sqrt{n}}\right) \underset{5.4}{\overset{\text{Satz}}{=}} \frac{1}{\sigma\sqrt{n}}\left[E(Z_n) - n\mu\right] = 0$$

und die Varianz

$$V(Z_n^*) = V\left(\frac{Z_n - n\mu}{\sigma\sqrt{n}}\right) \underset{5.11}{\overset{\text{Satz}}{=}} \frac{1}{n\sigma^2} V(Z_n) = 1.$$

Satz 7.9 *Der zentrale Grenzwertsatz*

Seien $X_1, X_2, ...$ unabhängige, identisch verteilte Zufallsvariablen mit endlicher, positiver Varianz und sei Z_n^* wie oben definiert. Dann gilt für $a, b \in \mathbb{R}$ mit $a \le b$

$$\lim_{n\to\infty} P\left(a \le Z_n^* \le b\right) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} dx$$

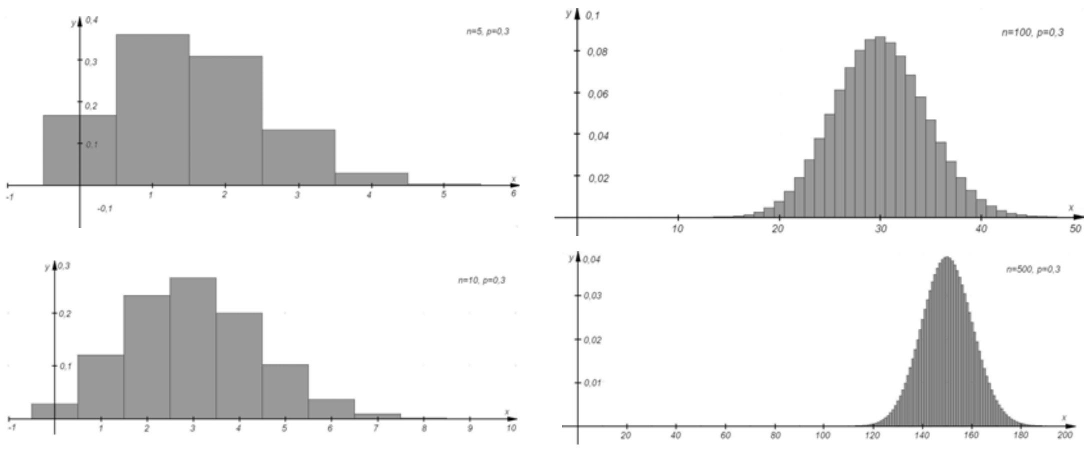


Abbildung 20 Graphische Veranschaulichung des zentralen Grenzwertsatzes anhand der Binomialverteilung

Anschaulich besagt der zentrale Grenzwertsatz, dass die Summe von n unabhängigen Zufallsvariablen annähernd normalverteilt ist. Für die meisten Anwendungen reicht bereits ein n von mindestens 30 aus, um akzeptable Näherungswerte zu erhalten.[12]

[12] BÜCHTER, Andreas / HENN, Hans-Wolfgang, Elementare Stochastik. Eine Einführung in die Mathematik der Daten und des Zufalls, Berlin/Heidelberg ^{2}2006, S.370

Betrachtet man die Summe von n unabhängigen Zufallsvariablen mit gleicher Verteilung und endlicher Varianz, so nähert sich für große n die Wahrscheinlichkeitsfunktion immer mehr der Dichte der Normalverteilung an, egal von welcher Verteilung ausgegangen wird.

DEFINITION 7.10 *Kompakte Konvergenz*

Eine Funktionenfolge $(g_n)_{n\geq 1}:\mathbb{R}\to\mathbb{R}$ heißt kompakt konvergent gegen eine Funktion $h:\mathbb{R}\to\mathbb{R}$, wenn für jedes $K>0$ gilt

$$\lim_{n\to\infty}\sup_{-K\leq x\leq K}\left|g_n(x)-g(x)\right|=0.$$

Für das nächste Lemma müssen wir noch die Wahrscheinlichkeitsfunktion $f_n(x)$ wie folgt definieren

$$f_n:\mathbb{R}\to\mathbb{R}, f_n(x):=p_n(k)=P(Z_n=k), x\in\left]k-\frac{1}{2},k+\frac{1}{2}\right].$$

$$p_n(k)=\int_{k-\frac{1}{2}}^{k+\frac{1}{2}}f_n(y)dy$$

LEMMA 7.11 Seien $(Z_n)_{n\geq 1}$ $\mathbb{Z}$ – wertige Zufallsvariablen mit den Wahrscheinlichkeitsfunktionen $p_n(k)=P(Z_n=k)$ und sei $g_n(x):=\sigma\sqrt{n}f_n\left(x\sigma\sqrt{n}+n\mu\right)$. Seien weiters $(u_n)_{n\geq 1}$ und $(o_n)_{n\geq 1}$ ganzzahlige Folgen mit $u_n\leq o_n$, wobei $\frac{u_n-n\mu}{\sigma\sqrt{n}}$ und $\frac{o_n-n\mu}{\sigma\sqrt{n}}$ beschränkt sind. Falls $(g_n)_{n\geq 1}$ kompakt gegen eine Funktion $g:\mathbb{R}\to\mathbb{R}$ konvergiert, dann gilt

$$\lim_{n\to\infty}\left|P(u_n\leq Z_n\leq o_n)-\int_{\frac{u_n-\frac{1}{2}-n\mu}{\sigma\sqrt{n}}}^{\frac{o_n+\frac{1}{2}-n\mu}{\sigma\sqrt{n}}}g(x)dx\right|=0.$$

BEWEIS:

Für ganze Zahlen $u\leq o$ erhalten wir

$$P(u\leq Z_n\leq o)=\sum_{i=u}^{o}p_n(k)=\sum_{i=u}^{o}\int_{i-\frac{1}{2}}^{i+\frac{1}{2}}f_n(y)dy=\int_{u-\frac{1}{2}}^{o+\frac{1}{2}}f_n(y)dy$$

Die Substitution $y=n\mu+x\sigma\sqrt{n}$ liefert uns

$$P(u\leq Z_n\leq o)=\int_{\frac{u-\frac{1}{2}-n\mu}{\sigma\sqrt{n}}}^{\frac{o+\frac{1}{2}-n\mu}{\sigma\sqrt{n}}}\sigma\sqrt{n}f_n\left(x\sigma\sqrt{n}+n\mu\right)dx=\int_{\frac{u-\frac{1}{2}-n\mu}{\sigma\sqrt{n}}}^{\frac{o+\frac{1}{2}-n\mu}{\sigma\sqrt{n}}}g_n(x)dx$$

Diese Identität benutzen wir nun

$$\left| P\left(u_n \le Z_n \le o_n\right) - \int\limits_{\frac{u_n - \frac{1}{2} - n\mu}{\sigma\sqrt{n}}}^{\frac{o_n + \frac{1}{2} - n\mu}{\sigma\sqrt{n}}} g(x)dx \right| = \left| \int\limits_{\frac{u_n - \frac{1}{2} - n\mu}{\sigma\sqrt{n}}}^{\frac{o_n + \frac{1}{2} - n\mu}{\sigma\sqrt{n}}} g_n(x)dx - \int\limits_{\frac{u_n - \frac{1}{2} - n\mu}{\sigma\sqrt{n}}}^{\frac{o_n + \frac{1}{2} - n\mu}{\sigma\sqrt{n}}} g(x)dx \right| = .$$

$$= \left| \int\limits_{\frac{u_n - \frac{1}{2} - n\mu}{\sigma\sqrt{n}}}^{\frac{o_n + \frac{1}{2} - n\mu}{\sigma\sqrt{n}}} \left(g_n(x) - g(x)\right)dx \right| \underset{\text{für Integrale}}{\overset{\text{Dreiecksungleichung}}{\le}} \int\limits_{\frac{u_n - \frac{1}{2} - n\mu}{\sigma\sqrt{n}}}^{\frac{o_n + \frac{1}{2} - n\mu}{\sigma\sqrt{n}}} \left|g_n(x) - g(x)\right|dx$$

Da $\dfrac{u_n - n\mu}{\sigma\sqrt{n}}$ und $\dfrac{o_n - n\mu}{\sigma\sqrt{n}}$ beschränkt sind, können wir die folgende Ungleichungskette angeben

$$-K \le \frac{u_n - \frac{1}{2} - n\mu}{\sigma\sqrt{n}} \le \frac{o_n + \frac{1}{2} - n\mu}{\sigma\sqrt{n}} \le K$$

mit der wir auch noch das Integral abschätzen können

$$\int\limits_{\frac{u_n - \frac{1}{2} - n\mu}{\sigma\sqrt{n}}}^{\frac{o_n + \frac{1}{2} - n\mu}{\sigma\sqrt{n}}} \left|g_n(x) - g(x)\right|dx \le 2K \cdot \sup_{-K \le x \le K} \left|g_n(x) - g(x)\right| \underset{\text{Konvergenz}}{\overset{\text{kompakte}}{\to}} 0$$

□

Mit denselben Voraussetzungen wie in Lemma 7.11 erhalten wir noch das folgende zweite Resultat:

LEMMA 7.12 Seien $a, b \in \mathbb{R}$ mit $a \le b$. Dann gilt

$$\lim_{n\to\infty} P\left(a < Z_n^* \le b\right) = \int\limits_a^b g(x)dx .$$

BEWEIS:

Da laut Voraussetzung $Z_n \in \mathbb{Z}$ ist, gilt für die Wahrscheinlichkeit

$$P\left(a < Z_n^* \le b\right) = P\left(a < \frac{Z_n - n\mu}{\sigma\sqrt{n}} \le b\right) =$$

$$= P\left(n\mu + a\sigma\sqrt{n} < Z_n \le n\mu + b\sigma\sqrt{n}\right)$$

Wir erhalten

$$\left|P\left(a < Z_n^* \le b\right) - \int_a^b g(x)\,dx\right| =$$

$$= \left|P\left(\left\lfloor n\mu + a\sigma\sqrt{n}\right\rfloor + 1 < Z_n \le \left\lfloor n\mu + b\sigma\sqrt{n}\right\rfloor\right) - \int_a^b g(x)\,dx\right| =$$

$$= \left|P\left(\left\lfloor n\mu + a\sigma\sqrt{n}\right\rfloor + 1 < Z_n \le \left\lfloor n\mu + b\sigma\sqrt{n}\right\rfloor\right) - \int_{\frac{\lfloor n\mu + a\sigma\sqrt{n}\rfloor + \frac{1}{2} - n\mu}{\sigma\sqrt{n}}}^{\frac{\lfloor n\mu + b\sigma\sqrt{n}\rfloor + \frac{1}{2} - n\mu}{\sigma\sqrt{n}}} g(x)\,dx + \right.$$

$$\left. + \int_{\frac{\lfloor n\mu + a\sigma\sqrt{n}\rfloor + \frac{1}{2} - n\mu}{\sigma\sqrt{n}}}^{\frac{\lfloor n\mu + b\sigma\sqrt{n}\rfloor + \frac{1}{2} - n\mu}{\sigma\sqrt{n}}} g(x)\,dx - \int_a^b g(x)\,dx\right| \underset{\text{ungleichung}}{\overset{\text{Dreiecks-}}{\le}}$$

$$\le \left|P\left(\left\lfloor n\mu + a\sigma\sqrt{n}\right\rfloor + 1 < Z_n \le \left\lfloor n\mu + b\sigma\sqrt{n}\right\rfloor\right) - \int_{\frac{\lfloor n\mu + a\sigma\sqrt{n}\rfloor + \frac{1}{2} - n\mu}{\sigma\sqrt{n}}}^{\frac{\lfloor n\mu + b\sigma\sqrt{n}\rfloor + \frac{1}{2} - n\mu}{\sigma\sqrt{n}}} g(x)\,dx\right| +$$

$$+ \left|\int_{\frac{\lfloor n\mu + a\sigma\sqrt{n}\rfloor + \frac{1}{2} - n\mu}{\sigma\sqrt{n}}}^{\frac{\lfloor n\mu + b\sigma\sqrt{n}\rfloor + \frac{1}{2} - n\mu}{\sigma\sqrt{n}}} g(x)\,dx - \int_a^b g(x)\,dx\right|$$

Beide Klammerausdrücke konvergieren gegen 0. Der erste aufgrund von Lemma 7.11, beim zweiten konvergieren $\dfrac{\left\lfloor n\mu + a\sigma\sqrt{n}\right\rfloor + \frac{1}{2} - n\mu}{\sigma\sqrt{n}} \to a$ und $\dfrac{\left\lfloor n\mu + b\sigma\sqrt{n}\right\rfloor + \frac{1}{2} - n\mu}{\sigma\sqrt{n}} \to b$ für $n \to \infty$.

□

Lemma 7.13 Sei K eine positive reelle Zahl und sei C die Konstante aus der Stirlingschen Formel von Lemma 2.5. Dann gilt gleichmäßig für alle $x \in [-K, K]$ mit der Eigenschaft $np + x\sqrt{npq} \in \mathbb{N}_0$, dass

$$P\left(Z_n = np + x\sqrt{npq}\right) \sim \frac{1}{\sqrt{npq}} \cdot \frac{1}{C} \cdot e^{\frac{-x^2}{2}}.$$

BEWEIS:

Aus der Definition des Binomialkoeffizienten folgt

$$\binom{n}{np+x\sqrt{npq}} = \frac{n!}{\left(np+x\sqrt{npq}\right)!\cdot\left(n-np-x\sqrt{npq}\right)!} =$$

$$= \frac{n!}{\left(np+x\sqrt{npq}\right)!\cdot\left(n(1-p)-x\sqrt{npq}\right)!} =$$

$$= \frac{n!}{\left(np+x\sqrt{npq}\right)!\cdot\left(nq-x\sqrt{npq}\right)!}$$

Wendet man nun die Stirlingsche Formel auf diese Binomialkoeffizienten an, so erhält man

$$\binom{n}{np+x\sqrt{npq}} \sim \frac{\cancel{C}\cdot n^{n+\frac{1}{2}}\cdot e^{-n}}{\cancel{C}\cdot\left(np+x\sqrt{npq}\right)^{np+x\sqrt{npq}+\frac{1}{2}}\cdot e^{-np-x\sqrt{npq}}}\cdot$$

$$\cdot\frac{1}{C\cdot\left(nq-x\sqrt{npq}\right)^{nq-x\sqrt{npq}+\frac{1}{2}}\cdot e^{-nq+x\sqrt{npq}}} =$$

$$= \frac{n^n\cdot n^{\frac{1}{2}}\cdot e^{-n}}{\left(np+x\sqrt{npq}\right)^{np+x\sqrt{npq}}\cdot\left(np+x\sqrt{npq}\right)^{\frac{1}{2}}\cdot e^{-np}\cdot e^{-x\sqrt{npq}}}\cdot$$

$$\cdot\frac{1}{C\cdot\left(nq-x\sqrt{npq}\right)^{nq-x\sqrt{npq}}\cdot\left(nq-x\sqrt{npq}\right)^{\frac{1}{2}}\cdot e^{-nq}\cdot e^{x\sqrt{npq}}} =$$

$$= \frac{n^{\frac{1}{2}}}{C\cdot\left(np+x\sqrt{npq}\right)^{\frac{1}{2}}\cdot\left(nq-x\sqrt{npq}\right)^{\frac{1}{2}}}\cdot$$

$$\cdot\frac{n^n}{\left(np+x\sqrt{npq}\right)^{np+x\sqrt{npq}}\cdot\left(nq-x\sqrt{npq}\right)^{nq-x\sqrt{npq}}}$$

$$\cdot\underbrace{\frac{e^{-n}}{e^{-n(p+q)}}}_{=1}\cdot\underbrace{\frac{1}{e^{-x\sqrt{npq}}\cdot e^{x\sqrt{npq}}}}_{=1} \overset{np+x\sqrt{npq}=n\cdot\left(p+x\sqrt{\frac{pq}{n}}\right)}{=}$$

$$= \frac{\cancel{\sqrt{n}}}{C \cdot \cancel{\sqrt{n}} \cdot \left(p + x\sqrt{\frac{pq}{n}}\right)^{\frac{1}{2}} \cdot \sqrt{n} \cdot \left(q - x\sqrt{\frac{pq}{n}}\right)^{\frac{1}{2}}} \cdot$$

$$\cdot \frac{n^n}{n^{np+x\sqrt{npq}} \cdot \left(p + x\sqrt{\frac{pq}{n}}\right)^{np+x\sqrt{npq}}} \cdot$$

$$\cdot \frac{1}{n^{nq-x\sqrt{npq}} \cdot \left(q - x\sqrt{\frac{pq}{n}}\right)^{nq-x\sqrt{npq}}} =$$

$$= \frac{1}{C \cdot \sqrt{n} \cdot \left(p + x\sqrt{\frac{pq}{n}}\right)^{\frac{1}{2}} \cdot \left(q - x\sqrt{\frac{pq}{n}}\right)^{\frac{1}{2}}} \cdot$$

$$\cdot \frac{\cancel{n^n}}{\cancel{n^{n(p+q)}} \cdot \left(p + x\sqrt{\frac{pq}{n}}\right)^{np+x\sqrt{npq}} \cdot \left(q - x\sqrt{\frac{pq}{n}}\right)^{nq-x\sqrt{npq}}} =$$

$$\overset{\text{Heraus-}}{\underset{\text{heben}}{=}} \frac{1}{C \cdot \sqrt{npq} \cdot \left(1 + x\sqrt{\frac{q}{np}}\right)^{\frac{1}{2}} \cdot \left(1 - x\sqrt{\frac{p}{nq}}\right)^{\frac{1}{2}}} \cdot$$

$$\cdot \frac{1}{\left(p + x\sqrt{\frac{pq}{n}}\right)^{np+x\sqrt{npq}} \cdot \left(q - x\sqrt{\frac{pq}{n}}\right)^{nq-x\sqrt{npq}}}$$

Da

$$\lim_{n\to\infty}\left(1 + x\sqrt{\frac{q}{np}}\right) = \lim_{n\to\infty}\left(1 - x\sqrt{\frac{p}{nq}}\right) = 1$$

gilt

$$\binom{n}{np + x\sqrt{npq}} \sim \frac{1}{C} \cdot \frac{1}{\sqrt{npq}} \cdot \frac{1}{\left(p + x\sqrt{\frac{pq}{n}}\right)^{np+x\sqrt{npq}} \cdot \left(q - x\sqrt{\frac{pq}{n}}\right)^{nq-x\sqrt{npq}}}$$

Und wir erhalten für binomialverteilte Wahrscheinlichkeiten

$$P\left(Z_n = np + x\sqrt{npq}\right) = \binom{n}{np + x\sqrt{npq}} \cdot p^{np+x\sqrt{npq}} \cdot q^{nq-x\sqrt{npq}} \sim$$

$$\sim \frac{1}{C} \cdot \frac{1}{\sqrt{npq}} \cdot \frac{p^{np+x\sqrt{npq}} \cdot q^{nq-x\sqrt{npq}}}{\left(p + x\sqrt{\frac{pq}{n}}\right)^{np+x\sqrt{npq}} \cdot \left(q - x\sqrt{\frac{pq}{n}}\right)^{nq-x\sqrt{npq}}} =$$

$$\overset{\text{Heraus-}}{\underset{\text{heben}}{=}} \frac{1}{C} \cdot \frac{1}{\sqrt{npq}} \cdot \frac{1}{\left(1 + x\sqrt{\frac{q}{np}}\right)^{np+x\sqrt{npq}} \cdot \left(1 - x\sqrt{\frac{p}{nq}}\right)^{nq-x\sqrt{npq}}}$$

Zuletzt bleibt noch zu zeigen, dass

$$\frac{1}{\left(1 + x\sqrt{\frac{q}{np}}\right)^{np+x\sqrt{npq}} \cdot \left(1 - x\sqrt{\frac{p}{nq}}\right)^{nq-x\sqrt{npq}}} \sim e^{\frac{-x^2}{2}} .$$

gilt.

Wir verwenden die Taylorreihen-Entwicklung des Logarithmus, die wir nach oben abschätzen

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - + \ldots = x - \frac{x^2}{2} + \sum_{i=3}^{\infty} (-1)^{i+1} \cdot \frac{x^i}{i} \quad \forall x \in \left]-1,1\right[$$

$$\sum_{i=3}^{\infty} (-1)^{i+1} \cdot \frac{x^i}{i} \le \frac{x^3}{3} \le x^3 \Rightarrow \sum_{i=3}^{\infty} (-1)^{i+1} \cdot \frac{x^i}{i} = O\left(x^3\right)$$

$$\Rightarrow \log(1+x) = x - \frac{x^2}{2} + O\left(x^3\right)$$

Dieses Ergebnis wenden wir nun auf den Logarithmus des Nenners an und erhalten

$$\log\left[\left(1 + x\sqrt{\frac{q}{np}}\right)^{np+x\sqrt{npq}} \cdot \left(1 - x\sqrt{\frac{p}{nq}}\right)^{nq-x\sqrt{npq}}\right] =$$

$$= \left(np + x\sqrt{npq}\right) \cdot \log\left(1 + x\sqrt{\frac{q}{np}}\right) + \left(nq - x\sqrt{npq}\right) \cdot \log\left(1 - x\sqrt{\frac{p}{nq}}\right) =$$

$$= \left(np + x\sqrt{npq}\right) \cdot \left(x\sqrt{\frac{q}{np}} - \frac{1}{2}x^2 \cdot \frac{q}{np} + O\left(x^3 \cdot \left(\frac{q}{np}\right)^{3/2}\right)\right) +$$

$$+ \left(nq - x\sqrt{npq}\right) \cdot \left(-x\sqrt{\frac{p}{nq}} - \frac{1}{2}x^2 \cdot \frac{p}{nq} + O\left(x^3 \cdot \left(\frac{p}{nq}\right)^{3/2}\right)\right) =$$

Da

$$O\left(x^3 \cdot \left(\frac{q}{np}\right)^{3/2}\right) = O\left(\frac{1}{n^{3/2}}\right), O\left(x^3 \cdot \left(\frac{p}{nq}\right)^{3/2}\right) = O\left(\frac{1}{n^{3/2}}\right)$$

folgt nach Ausmultiplikation

$$\left(np + x\sqrt{npq}\right) \cdot \left(x\sqrt{\frac{q}{np}} - \frac{1}{2}x^2 \cdot \frac{q}{np} + O\left(\frac{1}{n^{3/2}}\right)\right) =$$

$$= x\sqrt{npq} + x^2 q - \frac{x^2}{2}q - \underbrace{\frac{x^3}{2} \cdot \sqrt{\frac{q^3}{np}}}_{=O\left(\frac{1}{\sqrt{n}}\right)} + \underbrace{O\left(\frac{p}{\sqrt{n}}\right)}_{=O\left(\frac{1}{\sqrt{n}}\right)} + O\left(\frac{x\sqrt{pq}}{n}\right)$$

bzw.

$$\left(nq - x\sqrt{npq}\right) \cdot \left(-x\sqrt{\frac{p}{nq}} - \frac{1}{2}x^2 \cdot \frac{p}{nq} + O\left(\frac{1}{n^{3/2}}\right)\right) =$$

$$= -x\sqrt{npq} + x^2 p - \frac{x^2}{2}p + \underbrace{\frac{x^3}{2} \cdot \sqrt{\frac{p^3}{nq}}}_{=O\left(\frac{1}{\sqrt{n}}\right)} + \underbrace{O\left(\frac{q}{\sqrt{n}}\right)}_{=O\left(\frac{1}{\sqrt{n}}\right)} - O\left(\frac{x\sqrt{pq}}{n}\right)$$

Summiert man diese beiden Ergebnisse, so führt das zu

$$\cancel{x\sqrt{npq}} + x^2 q - \frac{x^2}{2}q + O\left(\frac{1}{\sqrt{n}}\right) \cancel{+O\left(\frac{x\sqrt{pq}}{n}\right)} -$$

$$\cancel{-x\sqrt{npq}} + x^2 p - \frac{x^2}{2}p + O\left(\frac{1}{\sqrt{n}}\right) \cancel{-O\left(\frac{x\sqrt{pq}}{n}\right)} =$$

$$= x^2 q - \frac{x^2}{2}q + x^2 p - \frac{x^2}{2}p + O\left(\frac{1}{\sqrt{n}}\right) =$$

$$x^2 \cdot \left(1 - p - \frac{1}{2} \cdot (1-p) + p - \frac{1}{2}p\right) + O\left(\frac{1}{\sqrt{n}}\right) = \frac{x^2}{2} + O\left(\frac{1}{\sqrt{n}}\right)$$

Aus

$$\log\left[\left(1 + x\sqrt{\frac{q}{np}}\right)^{np + x\sqrt{npq}} \cdot \left(1 - x\sqrt{\frac{p}{nq}}\right)^{nq - x\sqrt{npq}}\right] \sim \frac{x^2}{2}$$

folgt somit

$$\frac{1}{\left(1 + x\sqrt{\frac{q}{np}}\right)^{np + x\sqrt{npq}} \cdot \left(1 - x\sqrt{\frac{p}{nq}}\right)^{nq - x\sqrt{npq}}} \sim e^{\frac{-x^2}{2}}.$$

□

SATZ 7.14 *Satz von De Moivre-Laplace*

Sei Z_n die Summe von Bernoulli-verteilten Zufallsvariablen und sei $f_n : \mathbb{R} \to \mathbb{R}, f_n(x) := P(Z_n = k)$ mit $x \in \left]k - \frac{1}{2}, k + \frac{1}{2}\right]$. Dann konvergiert

$$g_n(x) := \sqrt{npq} \cdot f_n\left(x\sqrt{npq} + np\right).$$

kompakt gegen $\frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-x^2}{2}}$.

BEWEIS:

Aufgrund von Lemma 7.13 konvergiert $g_n(x)$ kompakt gegen $\frac{1}{C} \cdot e^{\frac{-x^2}{2}}$ Als Einziges müssen wir jetzt nur noch zeigen, dass $C = \sqrt{2\pi}$.

Die Ungleichung von Tschebyschew (Satz 5.19) liefert die folgende Abschätzung:

$$P\left(\left|Z_n^*\right| < K\right) = P\left(-K \le Z_n^* \le K\right) \ge 1 - \frac{1}{K^2} \cdot \underbrace{V\left(Z_n^*\right)}_{=1}$$

$$1 - \frac{1}{K^2} \le \underbrace{P\left(-K \le Z_n^* \le K\right)}_{\to \int_{-K}^{K} \frac{1}{C} \cdot e^{\frac{-x^2}{2}} dx \text{ (Lemma 7.12)}} \le 1$$

Für $K \to \infty$ vereinfacht sich dieser Ausdruck zu

$$\int_{-\infty}^{\infty} \frac{1}{C} \cdot e^{\frac{-x^2}{2}} dx = 1.$$

Aus dem Beweis von Proposition 6.5 folgt, dass $C = \sqrt{2\pi}$.

□

BEMERKUNG 7.15 Als Faustregel, dass die Binomialverteilung durch die Normalverteilung approximiert werden kann, muss

$$\sigma^2 = np(1-p) > 9$$

gelten.

TEIL 2

WAHLMANIPULATIONEN,

MANDATSVERTEILUNGEN UND MACHTINDIZES

8. DAS WILL-ROGERS-PHÄNOMEN UND GERRYMANDER

Bei einer Mehrheitswahl erhält der Wahlgewinner alle Stimmen, während die Wahlverlierer komplett leer ausgehen. Beispiel für so ein Wahlsystem ist etwa die Wahl eines Präsidenten. Bei einer Verhältniswahl dagegen werden die Sitze entsprechend der erhaltenen Wählerstimmen vergeben. Auf diesem Wahlsystem beruht der österreichische Nationalrat.

In diesem Kapitel möchte ich kurz darauf eingehen, wie ohne mathematische Mittel das Wahlergebnis beim Mehrheitswahlrecht durch die Einteilung der Wahlkreise entscheidend verändert werden kann. Durch teils abenteuerliche Grenzziehung ist eine geschickte Manipulation möglich, die aus einem vermeintlichen Wahlverlierer plötzlich einen Wahlsieger macht.

Bei dem sogenannten *Will-Rogers-Phänomen*, bekannt auch als *Stage-Migration*, handelt es sich um einen Effekt, der bei der falschen Gruppierung von Daten auftreten kann, wenn der Wechsel eines Elements zu einer anderen Gruppe eine Erhöhung bzw. Erniedrigung des Mittelwerts in beiden Gruppen bewirkt, obwohl sich gesamt betrachtet nichts ändert.

Die Bezeichnung Will-Rogers-Phänomen stammt von Feinstein, der im Jahr 1985 im "*New England Journal of Medicine*" erstmals diesen Begriff verwendete in memoriam Will Rogers, von dem folgender Ausspruch betreffend die Migration zur Zeit der Wirtschaftskrise stammt: „Als die „Okies" Oklahoma verließen und nach Kalifornien zogen, haben sie dadurch den durchschnittlichen Intelligenzquotienten in beiden Bundesstaaten erhöht."[13]

Unter Gerrymander versteht man die absichtliche und willkürliche Festsetzung der Grenzen von Wahlkreisen mit dem Zweck das Wahlergebnis zu seinen Gunsten hin zu beeinflussen.

Der Begriff Gerrymander setzt sich aus Elbridge *Gerry* und aus Sala*mander* zusammen. In seiner Funktion als Gouverneur von Massachusetts unterzeichnete *Gerry* 1812 nämlich ein Gesetz, mit dem die Grenzen zwischen den einzelnen Wahldistrikten neu gezogen und die Opposition de facto entmachtet wurde. Einzelne Wahldistrikte nahmen dabei recht sonderbare Formen an, so etwa der linke Teil von Essex County, dem die Gestalt eines Salamanders nachgesagt wurde.

[13] BECK-BORNHOLDT, Hans-Peter / DUBBEN, Hans-Hermann, Der Hund, der Eier legt. Erkennen von Fehlinformation durch Querdenken, Reinbek bei Hamburg 2001, S.203

Erstmals wurde das Wort Gerrymander in der Boston Gazette vom 26.März 1812 verwendet[14], zusammen mit einer Karikatur eines Teils von Essex County, die Abbildung 21 wiedergibt.

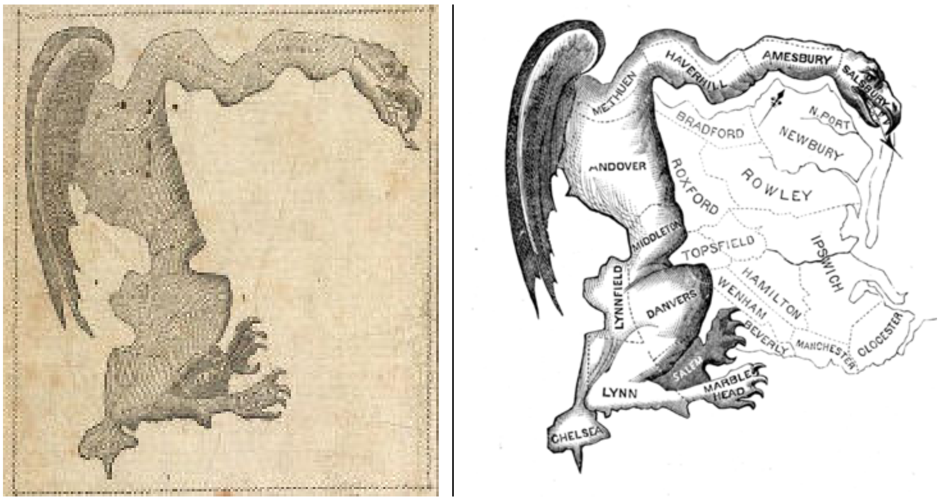


ABBILDUNG 21 Satirische Darstellung der Wahlbezirke Massachusetts aus dem Jahre 1812, links das Original[15], rechts eine spätere Überarbeitung, die die Karikatur mit ganz Essex County zeigt[16]

[14] GRIFFITH, Elmer, The Rise and Development of the Gerrymander, Dissertation, The University of Chicago, Chicago 1907, S.17
[15] http://www.loc.gov/exhibits/treasures/trr113.html [03.02.2009]
[16] DUBBEN, Hans-Hermann / BECK-BORNHOLDT, Hans-Peter, Mit an Wahrscheinlichkeit grenzender Sicherheit. Logisches Denken und Zufall, Reinbek bei Hamburg ^{2}2006, S.97

8.1 EIN FIKTIVES BEISPIEL[17]

Betrachten wir Abbildung 22, so würden wir einen Wahlsieg von weiß erwarten.

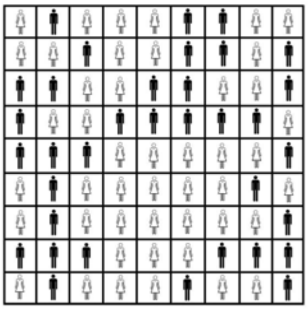


ABBILDUNG 22 Ausgangssituation:
Weiß kann gegenüber schwarz insgesamt eine Mehrheit von 46:35 verzeichnen

Es ist daher auch nicht weiter überraschend, wenn bei einer Einteilung in neun Gebiete weiß in der Mehrheit dieser Gebiete der Wahlsieger ist. Abbildung 23 zeigt mögliche Aufteilungen mit weiß als klarem Wahlsieger.

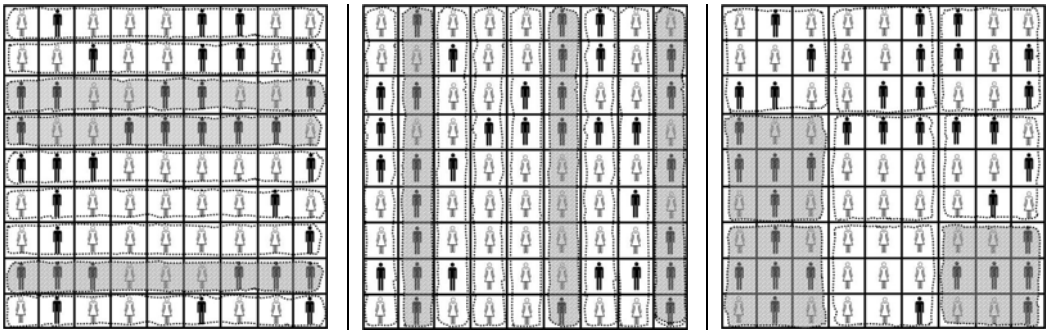


ABBILDUNG 23 6:3-Wahlsiege für Weiß, egal ob die Einteilung der Wahlkreise horizontal (links), vertikal (Mitte) oder in 3x3-Quadraten (recht) erfolgt

Dennoch kann auch schwarz die Mehrheit der Wahlkreise durch geschicktes Einteilen der Wahlkreise erringen, ebenso wie weiß in allen neun Gebieten die Oberhand erhalten kann. Abbildung 24 zeigt diesen Sachverhalt.

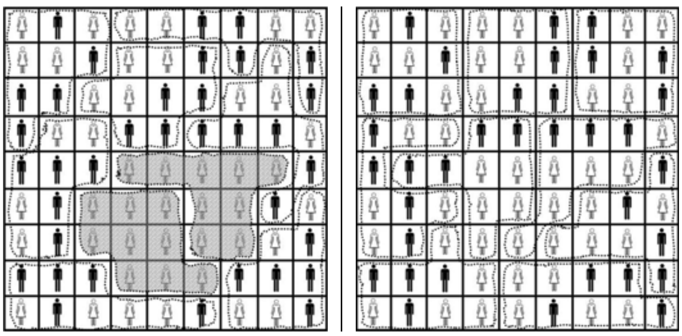


ABBILDUNG 24 Andere mögliche Wahlsiege: Links gewinnt plötzlich überraschenderweise schwarz in sieben von neun Gebieten, recht weiß sogar in allen neun Gebieten

[17] DUBBEN, Wahrscheinlichkeit, S.93ff.

8.2 GERRYMANDER IN DER REALITÄT

BEISPIEL 8.1 *Gerrymander in den USA*

Auch heutzutage ist Gerrymander in den USA eine weitverbreitete Art der Wahlkreismanipulation, wie Abbildung 25 anhand zweier Beispiele verdeutlicht:

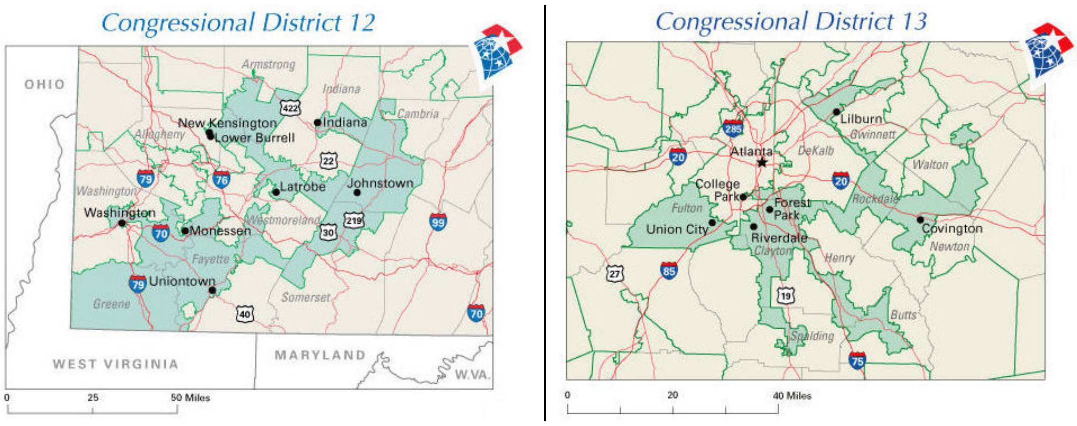


ABBILDUNG 25 Exemplarische Auswahl von Gerrymander in den USA, links der 12.Bundeswahlkreis (= Congressional District) von Pennsylvania, rechts der 13.Bundeswahlkreis von Georgia[18]

Weitere Beispiele für Gerrymander in den USA finden sich auf der Internetseite http://www.westmiller.com/fairvote2k/in_gerry.htm [03.02.2009].

BEISPIEL 8.2 *Gerrymander in Deutschland*

Im deutschen Bundeswahlgesetz (§6 Wahl nach Landeslisten, Absatz 6) ist geregelt, dass nur Parteien in den Bundestag einziehen dürfen, die entweder fünf Prozent der Stimmen erhalten haben oder in drei Wahlkreisen ein Direktmandat errungen haben.[19]

Als für die Bundestagswahl 2002 die Wahlkreise neu eingeteilt worden sind, verlor die PDS in Berlin eines ihrer Direktmandate und konnte in der kommenden Legislaturperiode nur noch zwei Abgeordnete für den Bundestag stellen.[20]

1998 konnte die PDS noch vier Direktmandate erreichen (in den Wahlkreisen 249, 258, 260 und 261), vier Jahre später nur noch zwei (in den Wahlkreisen 86 und 87).

[18] http://www.nationalatlas.gov/printable/congress.html#list [03.03.2009]
[19] http://www.bundestag.de/parlament/funktion/gesetze/bwahlg_pdf.pdf [02.03.2009]
[20] DUBBEN, Wahrscheinlichkeit, S.99

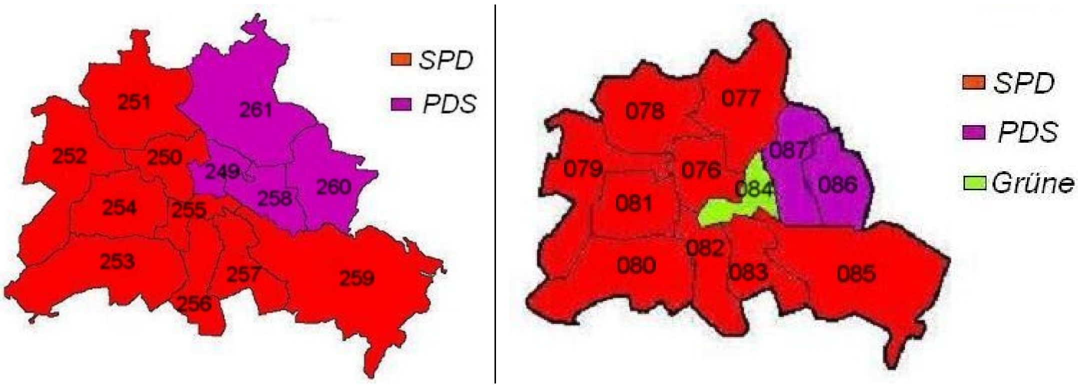


ABBILDUNG 26 Vergleich der Wahlkreiseinteilung und der gewonnenen Direktmandate für die Bundestagswahl in Berlin 1998 (links)[21] mit 2002 (rechts)[22]

BEISPIEL 8.3 *Gerrymander in Österreich*

Auch in Österreich gab es bereits einen offensichtlichen Fall von Gerrymander, als am 1.Jänner 1922 Wien durch die Abspaltung von Niederösterreich ein eigenständiges Bundesland wurde. Dies ermöglichte der ÖVP eine Vormachtstellung in Niederösterreich aufzubauen und der SPÖ eine Hochburg in Wien zu etablieren.[23]

Ein weiterer Fall von Gerrymander ist in Wien der Wahlkreis Innen-West, der aus den Bezirken Neubau, Josefstadt und Alsergrund besteht. Durch die Zusammenfassung genau dieser drei ehemals bürgerlichen Bezirke wird die SPÖ gegenüber der ÖVP bevorzugt.[24]

Neben Gerrymander gibt es auch noch weitere Möglichkeiten der Beeinflussung von Wahlergebnissen. Eine besteht darin, die Vormachtstellung politischer Gegner in einem bestimmten Wahlkreis durch den Bau von Wohnanlagen gezielt zu schwächen. Dies wurde etwa von der SPÖ in Wien praktiziert. Als Beispiel sei der Karl-Marx-Hof in Wien Döbling genannt, wo die ÖPV den Bezirksvorsteher stellt.[25]

---

[21] http://www.statistik-berlin.de/wahlen/vorwahlen/btw-98/ergebnisse/karten/karte3.html [03.02.2009]

[22] http://www.bundeswahlleiter.de/bundestagswahl2002/deutsch/ergebnis2002/bund_land/btw2002/kru11_btw2002.htm [11.06.2007]

[23] http://www.noe.gv.at/service/politik/landtag/Geschichte.htm [08.06.2007]

[24] http://de.wikipedia.org/wiki/Gerrymander [30.07.2010]

[25] PICHLER, Herbert, Wie „macht“ man ein Wahlergebnis? Ein erster Blick in Wahlrecht, Wahlarithmetik und Wahlgeographie, In: Von Wahl zu Wahl. Informationen zur Politischen Bildung Nr. 21, Wien 2004, S.88

9. MANDATSVERTEILUNGEN

Die zwei wichtigsten Verfahren zur Mandatsbestimmung sind das Verfahren nach Hare und das Verfahren nach d'Hondt. Nach einer Vorstellung der beiden Verfahren (inklusive einer stochastischen Analyse der Fairness beim Hare-Verfahren) werde ich die Mandatsvergabe bei österreichischen Nationalratswahlen besprechen, wo beide Verfahren zur Anwendung kommen.

Aufgrund der aktuellen Volkszählung werden in Österreich mit dem Hare-Verfahren für den Nationalrat die einzelnen Mandate den Wahlkreisen zugeordnet. Das Verfahren nach d'Hondt wird in Österreich bei der Vergabe der Mandate des Bundeswahlkreises bei Nationalratswahlen angewendet.

Das Verfahren nach Hare wurde 1792 in Amerika von Alexander Hamilton vorgeschlagen und 1850 von Samuel F. Vinton wieder in Erinnerung gebracht. 1855 wandte Carl Andrae erstmals dieses Verfahren in Europa an. Seinen Namen verdankt dieses Verfahren jedoch Thomas Hare, der in der 2.Hälfte des 19.Jahrhunderts diese Methode zusammen mit dem System der Vorzugswahl, die noch heutzutage in Irland angewendet wird, einführte. Horst F. Niemeyer entwickelte 1970 eine Modifikation, indem er vorschlug Parteien mit absoluter Stimmenmehrheit zuerst bei der Vergabe der Restmandate zu berücksichtigen.

Das Verfahren nach d'Hondt geht auf Thomas Jefferson zurück, der sich 1792 gegen die von Hamilton vorgeschlagene Methode dank eines Vetos des amerikanischen Präsidenten durchsetzen konnte. Knapp hundert Jahre später veröffentlichte der Jurist Victor d'Hondt im Jahr 1882 erneut dieses Verfahren, ehe zu Beginn des zwanzigsten Jahrhunderts der Physiker Eduard Hagenbach-Bischoff einen effektiven Algorithmus entwickelte.

9.1 Einführung

Es soll ein Gremium besetzt werden, das im Normalfall jedoch aus weniger Mitgliedern als die Ausgangsmenge besteht. Gesucht sind nun mathematische Modelle, mit denen die Mandate der im Gremium vertretenen Parteien oder Fraktionen berechnet werden können.

Beispiele für Ausgangsmengen und Gremien:

- ♠ *Parlament:* Die Ausgangsmenge besteht aus den bei einer Nationalratswahl abgegebenen gültigen Stimmen. Diese legen die Anzahl der Mandate der politischen Parteien im Parlament fest.
- ♠ *Parlamentarischer Ausschuss:* Hier ist die Ausgangsmenge die Anzahl der Abgeordneten. Die Mandate einer Partei in diesem Gremium werden bestimmt durch deren jeweilige Parlamentsstärke.

Sei n die Anzahl der Parteien. Jede Partei P_i habe S_i Stimmen erhalten. Für die Gesamtzahl S aller gültigen Stimmen gilt dann

$$S = \sum_{i=1}^{n} S_i \ .$$

Sei M_i die Anzahl der Mandate der i-ten Partei. Weiters sei M die zu vergebenden Mandate. Dann gilt

$$M = \sum_{i=1}^{n} M_i \ .$$

Die Gremiumsanteile der Parteien bezeichnet man als Quote. Sie lassen sich aufgrund einfacher Schlussrechnung nach der Formel

$$Q_i = \frac{S_i}{S} \cdot M \ .$$

berechnen.

Beispiel 9.1 *Anteilsberechnung anhand der Nationalratswahl 2006*

In Österreich werden bei Nationalratswahlen seit 1971 183 Mandate vergeben.[26] Bei der Nationalratswahl 2006 erhielten SPÖ, ÖVP, Grüne, FPÖ und BZÖ zusammen 4513746 gültige Stimmen. Es ergab sich folgende Stimmverteilung mit den dazugehörigen Quoten:

Partei	Stimmverteilung[27]	Quote[28]
SPÖ	1663986	67,463
ÖVP	1616493	65,537
Grüne	520130	21,088
FPÖ	519598	21,066
BZÖ	193539	7,847

Tabelle 11 Nicht ganzzahlige Anteile bei der Mandatsverteilung zur Nationalratswahl 2006

[26] STATISTIK AUSTRIA (Hg.), Statistisches Jahrbuch Österreich 2006, Wien 2005, S.489
[27] http://www.bmi.gv.at/wahlen/ [01.07.2007]
[28] Anteile hier und bei folgenden Beispielen stets auf drei Dezimalen gerundet

Wie in Beispiel 9.1 ersichtlich, entstehen normalerweise bei der Berechnung von Sitzverteilungen keine ganzzahligen Anteile. Die Schwierigkeit liegt nun darin, wie die dabei auftretenden Reste auf die restlichen Mandate aufzuteilen sind. Im Laufe der Zeit wurden mehrere unterschiedliche Verfahren entwickelt, die angeben, wie mit diesen Resten umzugehen ist.

Einige Anforderungen an Verteilungsverfahren:

♠ *Eindeutigkeit:* Für jede Gremiumsgröße soll die Sitzverteilung nicht mehrdeutig sein.

♠ *Ganzzahligkeit:* Die für die Besetzung eines Gremiums zu berechnenden Mandate müssen ganzzahlig sein.

♠ *Homogenität:* Bei fixer Mandatszahl liefern proportionale Stimmanteile dieselben Mandatsverteilungen.

♠ *Mehrheitstreue:* Die Mehrheitsverhältnisse in der Ausgangsmenge sollen sich auch im Gremium widerspiegeln. Einer Partei mit absoluter Stimmmehrheit steht also auch die absolute Mandatsmehrheit zu. Ebenso sollte eine Partei mit absoluter Minderheit der Stimmen keine absolute Mandatsmehrheit erhalten.

$$S_i > \frac{1}{2}S \Rightarrow M_i > \frac{1}{2}M, S_i < \frac{1}{2}S \Rightarrow M_i < \frac{1}{2}M$$

♠ *Mindestvertretung:* Tritt dann auf, wenn jede Partei in einem Gremium vertreten sein soll, egal wie klein sie ist.

♠ *Monotonie:* Eine Partei mit weniger Stimmen in der Ausgangsmenge als eine andere Partei erhält im Gremium dann auch weniger Mandate.

$$S_i \le S_j \Rightarrow M_i \le M_j$$

♠ *Quotenbedingung:* Eine Partei sollte mindestens so viele Mandate erhalten wie die abgerundete Quote und maximal so viele wie die aufgerundete Quote.

$$\lfloor Q_i \rfloor \le M_i \le \lfloor Q_i \rfloor + 1$$

♠ *Summenforderung:* Die festgelegte Größe des Gremiums bestimmt die genaue Anzahl der zu vergebenden Mandate.

$$M = \sum_{i=1}^{n} M_i$$

♠ *Unabhängigkeitsbedingung:* Die Anzahl der Mandate einer Partei P_i hängt nur von ihren eigenen erhaltenen Stimmen s_i ab, nicht jedoch von den Stimmen der anderen Parteien.

♠ *Verhältnistreue:* Das Gremium soll ein Abbild der tatsächlichen Anteile der Ausgangsmenge sein.

Leider lassen sich in den meisten Fällen nicht alle Forderungen gleichzeitig erfüllen. Da die Ganzzahligkeit und die Summenforderung (wenn man von Überhangsmandaten absieht) notwendige Bedingungen sind, müssen bei den restlichen Forderungen Abstriche gemacht werden.

Sogar die Monotonieeigenschaft ist keine notwendige Anforderung an Verteilungsverfahren. Bereits zweimal, 1953 und 1959, wurde die Monotonieeigenschaft bei der Mandatsverteilung des österreichischen Nationalrates verletzt. Jedesmal erhielt die ÖVP ein Mandat mehr als die SPÖ, obwohl sie weniger Stimmen erhalten hatte.[29]

In den nächsten Kapiteln möchte ich auf die beiden Berechnungsverfahren von Hare und d'Hondt eingehen. Es wird sich zeigen, dass diese beiden Verfahren unterschiedliche Mandatszahlen liefern können und dass jedes Verfahren seine spezifischen Vor- und Nachteile hat.

Verfahren, bei denen lediglich die gewünschte Gremiumsgröße berechnet wird, nennt man integrale Verfahren. Verfahren, die der Reihe nach sämtliche Gremiumsgrößen aufbauen, bis die gewünschte Größe erreicht ist, bezeichnet man als inkrementelle Verfahren.

Verfahren, die zunächst die Quote berechnen und diese anschließend runden, nennt man Quotenverfahren. Divisorverfahren dagegen teilen die Stimmen einer Partei durch einen entsprechenden Divisor und runden danach diesen Quotienten.

Das Verfahren nach Hare ist ein integrales Quotenverfahren, das Verfahren nach d'Hondt ist ein Beispiel für ein inkrementelles Divisorverfahren.

[29] Details siehe STATISTIK AUSTRIA, Jahrbuch 2006, S.489

9.2 DAS VERFAHREN NACH HARE

Andere Bezeichnungen für dieses Verfahren:

- ♠ Hare/Niemeyer-Verfahren
- ♠ Hamilton-Methode
- ♠ Verfahren vom größten Rest
- ♠ Methode der stärksten Bruchteile

Algorithmus des Verfahrens nach Hare:

- ♠ 1. Schritt: *Grundmandate*
  Zunächst wird für jede Partei P_i die Quote Q_i berechnet. Jede Partei erhält dann als Grundmandate $\overline{M_i}$ die abgerundete Anzahl der jeweiligen Quote.
  $$\overline{M_i} = \left\lfloor \frac{S_i}{S} \cdot M \right\rfloor$$
- ♠ 2. Schritt: *Restmandate*
  Eventuell noch zu vergebende Mandate werden in Reihenfolge der Größe der auftretenden Reste $R_i = \overline{M_i} - \left\lfloor \overline{M_i} \right\rfloor$ vergeben.

BEISPIEL 9.2 *Fortsetzung von* Beispiel 9.1

Weist man jeder Partei zunächst ihren abgerundeten Gremiumsanteil zu, so erhält die SPÖ 67, die ÖVP 65, die Grünen und die FPÖ jeweils 21 und das BZÖ 7 Mandate. Es bleiben noch 2 Mandate übrig. Diese werden dem BZÖ und der ÖVP zugeteilt, da sie die größten Restanteile besitzen.

Partei	Quote	Rest	Reihenfolge der Reste	Mandate
SPÖ	67,463	0,463	3	$67+0=67$
ÖVP	65,537	0,537	2	$65+1=66$
Grüne	21,088	0,088	4	$21+0=21$
FPÖ	21,066	0,066	5	$21+0=21$
BZÖ	7,847	0,847	1	$7+1=8$

TABELLE 12 Nationalratswahl 2006, Mandatsverteilung nach Hare

Vorteile des Verfahrens nach Hare:

- ♠ Das Verfahren beruht auf einen einfachen Algorithmus, der leicht nachzuvollziehen und anzuwenden ist.
- ♠ Das Verfahren von Hare benachteiligt kleinere Partei nicht so sehr wie das Verfahren nach d'Hondt (siehe Satz 9.4, Beispiel 9.10)
- ♠ Erfüllung der Quotenbedingung

DEFINITION 9.3 *Gleichverteilung*

Eine Zufallsvariable X heißt *gleichverteilt* mit den Parametern $\alpha, \beta \in \mathbb{R}$ mit $\alpha < \beta$, falls ihre Dichte von der Form

$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha} \text{ falls } \alpha \le x \le \beta \\ 0 \text{ sonst} \end{cases}$$

ist.

Die Gleichverteilung wird verwendet, wenn bekannt ist, dass die Zufallsvariable X nur Werte aus dem Intervall $[\alpha, \beta]$ annehmen kann und jeder Wert gleich wahrscheinlich ist. Somit hängt die Wahrscheinlichkeit, dass X in ein beliebiges Teilintervall von $[\alpha, \beta]$ fällt, allein von der Länge dieses Teilintervalls ab.

SATZ 9.4 *Stochastische Analyse der Fairness beim Hare-Verfahren*[30]

Seien P_A und P_B die beiden Parteien in einem Zwei-Parteiensystem mit P_A als die stimmenstärkere Partei. P_A habe aufgrund von S_A Stimmen M_A Mandate erhalten. Weiters sei $\widehat{S_A} := \frac{S_A}{S}$ der Stimmenanteil von P_A eine gleichverteilte Zufallsvariable auf dem kompakten Intervall $[1/2; 1]$. Dann gilt für den Erwartungswert der Differenz $D := M_A - M\widehat{S_A}$

$$E(D) = \begin{cases} 0 \text{ falls } M \in \mathbb{N}_g \\ \dfrac{1}{4M} \text{ falls } M \in \mathbb{N}_u \end{cases}.$$

BEWEIS:

Zunächst fassen wir die Mandate M_A als Zufallsvariable auf, für die gilt

$$M_A = \begin{cases} M\widehat{S_A} & \text{falls } M\widehat{S_A} \in \mathbb{N} \\ \left\lfloor M\widehat{S_A} \right\rfloor & \text{falls } M\widehat{S_A} \notin \mathbb{N} \text{ und } M\widehat{S_A} - \left\lfloor M\widehat{S_A} \right\rfloor \le 1/2 \\ \left\lfloor M\widehat{S_A} \right\rfloor + 1 & \text{falls } M\widehat{S_A} \notin \mathbb{N} \text{ und } M\widehat{S_A} - \left\lfloor M\widehat{S_A} \right\rfloor > 1/2 \end{cases}$$

Sei $X := M\widehat{S_A} - \left\lfloor M\widehat{S_A} \right\rfloor$. Dann gilt für die Differenz D

$$D = \begin{cases} 0 & \text{falls } M\widehat{S_A} \in \mathbb{N} \\ -X & \text{falls } M\widehat{S_A} \notin \mathbb{N} \text{ und } X \le 1/2 \\ 1 - X & \text{falls } M\widehat{S_A} \notin \mathbb{N} \text{ und } X > 1/2 \end{cases}$$

30 siehe HESSE Christian, Angewandte Wahrscheinlichkeitstheorie, Braunschweig / Wiesbaden 2003, S.113-114 bzw. HESSE, Christian / MEISTER, Alexander, Übungsbuch zur angewandten Wahrscheinlichkeitstheorie. Aufgaben und Lösungen, Wiesbaden 2005, S.53-54

Um eine Aussage über die Fairness machen zu können, interessant uns nun der Erwartungswert dieser Differenz. Dazu benötigen wir noch die Dichte von X auf $[0,1[$. Der Fall $D=0$ tritt aufgrund der Annahme einer gleichverteilten Zufallsvariablen nur mit Wahrscheinlichkeit null ein und kann somit vernachlässigt werden.

Zunächst müssen wir die Verteilung von X auf $[0,1[$ darstellen. Dabei sind zwei Fälle zu unterscheiden, je nachdem ob die Mandatszahl M gerade oder ungerade ist. Letztgenannter Fall wird noch einmal zweimal unterteilt, zunächst in das Intervall $]0,1/2]$, danach in $]1/2,1[$.

Sei $M \in \mathbb{N}_g$. Dann gilt für ein beliebiges $x \in [0,1[$

$$F(x)=P(X \le x)=P\left(M\widehat{S_A} \in \left[\left\lfloor M\widehat{S_A}\right\rfloor, \left\lfloor M\widehat{S_A}\right\rfloor + x\right]\right) \overset{M/2 \le \lfloor M\widehat{S_A} \rfloor \le M-1}{=}$$

$$= \sum_{i=M/2}^{M-1} P\left(M\widehat{S_A} \in [i, i+x]\right) = \sum_{i=M/2}^{M-1} P\left(\widehat{S_A} \in \left[\frac{i}{M}, \frac{i+x}{M}\right]\right) \underset{\text{auf } [1/2,1]}{\overset{\text{Gleichverteilung}}{=}}$$

$$= \sum_{i=M/2}^{M-1} \frac{\frac{i+x}{M} - \frac{i}{M}}{1-\frac{1}{2}} = \frac{2x}{M} \sum_{i=M/2}^{M-1} 1 = \frac{2x}{M} \cdot \left(M - 1 - \frac{M}{2} + 1\right) = x$$

Die zugehörige Dichte ist

$$f(x) = F'(x) \ = 1$$

Für den Erwartungswert gilt

$$E(D) = \int_0^{1/2} (-x) \cdot 1\,dx + \int_{1/2}^{1} (1-x) \cdot 1\,dx = \left[-\frac{x^2}{2}\right]_{x=0}^{x=1/2} + \left[x - \frac{x^2}{2}\right]_{x=1/2}^{x=1} =$$

$$= -\frac{1}{8} + 1 - \frac{1}{2} - \frac{1}{2} + \frac{1}{8} = 0$$

Sei $M \in \mathbb{N}_u$. Dann gilt für ein beliebiges $x \in]0,1/2]$

$$F(x) = P(X \le x) = \sum_{i=(M+1)/2}^{M-1} P\left(\widehat{S_A} \in \left[\frac{i}{M}, \frac{i+x}{M}\right]\right) =$$

$$= \frac{2x}{M} \cdot \left(M - 1 - \frac{M+1}{2}\right) = x \cdot \frac{M-1}{M}$$

Ganz ähnlich lässt sich die Verteilungsfunktion für $\left]1/2,1\right[$ berechnen

$$
\begin{aligned}
F(x) &= P(X \leq x) = \\
&= \underbrace{P\left(M\widehat{S_A} \in \left[\frac{M}{2}, \frac{M-1}{2}+x\right]\right)}_{=\frac{\frac{2x-1}{2M}}{\frac{1}{2}}} + \sum_{i=(M+1)/2}^{M-1} P\left(M\widehat{S_A} \in [i, i+x]\right) = \\
&= \frac{2}{M}\cdot\left(x-\frac{1}{2}\right) + x\cdot\frac{M-1}{M} = \frac{x\cdot(1+M)-1}{M}
\end{aligned}
$$

Somit erhält man für eine ungerade Mandatsanzahl die Dichte

$$
f(x) = F'(x) = \begin{cases} \dfrac{M-1}{M} & \text{falls } x \in \left]0,1/2\right[\\ \dfrac{M+1}{M} & \text{falls } x \in \left]1/2,1\right[\end{cases}
$$

Für den Erwartungswert dieser Dichte gilt

$$
\begin{aligned}
E(D) &= \int_0^{1/2} (-x)\cdot\frac{M-1}{M}dx + \int_{1/2}^{1} (1-x)\cdot\frac{M+1}{M}dx = \\
&= \frac{M-1}{M}\cdot\left[-\frac{x^2}{2}\right]_{x=0}^{x=1/2} + \frac{M+1}{M}\cdot\left[x-\frac{x^2}{2}\right]_{x=1/2}^{x=1} = \\
&= \frac{M-1}{M}\cdot\left(-\frac{1}{8}\right) + \frac{M+1}{M}\cdot\left(\frac{1}{2}-\frac{1}{2}+\frac{1}{8}\right) = \frac{\cancel{-M}+1+1\cancel{+M}}{8M} = \frac{1}{4M}
\end{aligned}
$$

Zusammengefasst erhält man folgenden Erwartungswert

$$
E(D) = \begin{cases} 0 & \text{falls } M \in \mathbb{N}_g \\ \dfrac{1}{4M} & \text{falls } M \in \mathbb{N}_u \end{cases}
$$

Offensichtlich ist also das Hare-Verfahren bei gerader Mandatsanzahl zur Gänze fair. Bei ungerader Mandatszahl wird die größere Partei ein wenig bevorzugt. Jedoch wird in diesem Fall das Verfahren umso fairer, je mehr Mandate vergeben werden.

□

Nachteile des Verfahrens nach Hare

- ♠ *Abweichung von der Proportionalität:* Bei kleineren Anteilen kann es zu größeren Abweichungen von der Proportionalität kommen.
- ♠ *Mehrdeutigkeit:* Es kann vorkommen, dass auftretende Reste ident sind, aber nur weniger noch nicht vergebene Sitze zur Verfügung stehen. In der Praxis entscheidet dann das Los, wer dieses Mandat erhält.
- ♠ *Getrennte Berechnung für jede Gremiumsgröße:* Somit lässt sich nicht auf einen Blick feststellen, wie sich das Gremium bei zunehmender Mandatszahl schrittweise verändert.
- ♠ *Alabama-Paradoxon*: Bei gleicher Stimmverteilung kann eine Partei bei Vergrößerung des Gremiums einen Sitz verlieren, den sie in einem kleineren Gremium schon erhalten hat. Man spricht dann von einem sogenannten Rücksprung. (siehe Beispiele 9.5 und 9.6)
- ♠ *Populations-Paradoxon*: Eine Partei *A* kann trotz Stimmenzuwachs ein Mandat verlieren, während gleichzeitig eine andere Partei *B* mit Stimmverlusten ein Mandat dazugewinnen kann. (siehe Beispiel 9.7)
- ♠ *Neue-Partei-Paradoxon*[31]*:* Eine Partei kann die Mandatsvergabe beeinflussen, auch wenn Sie selbst keine Mandate erhält. (siehe Beispiel 9.8)

BEISPIEL 9.5 *Alabama-Paradoxon*

Auch bei der Mandatsberechnung der Nationalratswahl 2006 tritt beim 10.Mandat bei der Verwendung des Verfahrens nach Hare ein Rücksprung auf Das BZÖ verliert dabei einen Sitz im Parlament, den diese Partei bei 9 zu vergebenden Mandaten bereits erhalten hat.

	9		10		11	
Partei	Quote	Mandate	Quote	Mandate	Quote	Mandate
SPÖ	3,318	3	3,686	4	4,055	4
ÖVP	3,223	3	3,581	4	3,939	4
Grüne	1,037	1	1,152	1	1,268	1
FPÖ	1,036	1	1,151	1	1,266	1
BZÖ	0,386	1	0,429	0	0,472	1

TABELLE 13 Das Alabama-Paradoxon bei der Nationalratswahl 2006

Es empfiehlt sich daher bei der Hare-Methode immer auch das um eins kleinere und größere Gremium zu berechnen und danach erst die Mandate zu vergeben. In diesem Beispiel sollte das BZÖ also auch bei 10 Mandaten ein Mandat erhalten und zugleich die ÖVP aufgrund des geringeren Restes im Vergleich zur SPÖ ein Mandat verlieren.

BEISPIEL 9.6 *Sonderfall des Alabama-Paradoxon*

Das Alabama-Paradoxon kann jedoch nicht nur Rücksprünge bewirken. Überraschenderweise wird jedoch in der von mir verwendeten Literatur der Möglichkeit keinerlei Beachtung geschenkt, dass eine Partei nicht nur ein Mandat verlieren sondern auch einen Sitz gewinnen kann, der ihr eigentlich noch nicht zustehen würde.

---
[31] englische Bezeichnung: New States Paradox

Bei der Nationalratswahl 2006 würde etwa dieser Fall bei insgesamt 55 zu vergebenden Mandaten auftreten. Das BZÖ hätte hier ein drittes Mandat erhalten, das jedoch erst bei der Vergabe von 60 Sitzen nach dem Verfahren von Hare fix dem BZÖ zustehen würde. Bei 55 Mandaten sollte also das BZÖ entgegen der Berechnung auch lediglich nur zwei Mandate erhalten, dafür aber die Grünen bereits ein siebentes aufgrund des nächstgrößeren Restes.

	54		55		56	
Partei	Quote	Mandate	Quote	Mandate	Quote	Mandate
SPÖ	19,907	20	20,276	20	20,644	21
ÖVP	19,339	20	19,697	20	20,055	20
Grüne	6,223	6	6,338	6	6,453	7
FPÖ	6,216	6	6,331	6	6,446	6
BZÖ	2,315	2	2,358	3	2,401	2

TABELLE 14 Sonderfall des Alabama-Paradoxons anhand der Nationalratswahl 2006

BEISPIEL 9.7 *Populations-Paradoxon*

Betrachten wir folgendes Beispiel[32], wo 656 Mandate vergeben werden. Trotz eines Gewinns von 26000 Stimmen muss Partei *A* ein Mandat an Partei *E* abgeben, die einen Verlust von 5000 Stimmen zu verzeichnen hat.

	Situation 1			Situation 2		
Partei	Stimmen	Quote	Mandate	Stimmen	Quote	Mandate
A	17140354	259,437	260	17166354	259,23	259
B	16082960	243,433	243	16106960	243,23	243
C	3427196	51,874	52	3452196	52,13	52
D	3424315	51,831	52	3454315	52,16	52
E	3265407	49,425	49	3260407	49,24	50

TABELLE 15 Populations-Paradoxon

BEISPIEL 9.8 *Neue-Partei-Paradoxon*

Im folgenden Beispiel sollen 13 Mandate vergeben werden[33]. Wie sich leicht nachrechnen lässt, bewirkt Partei *D* die Abwanderung eines Mandates von *A* zu *C* ohne selbst ein Mandat zu erhalten.

	Mandatsvergabe ohne D			Mandatsvergabe mit D		
Partei	Stimmen	Anteil	Mandate	Stimmen	Anteil	Mandate
A	4223	5,668	6	4223	5,490	5
B	3539	4,750	5	3539	4,601	5
C	1924	2,582	2	1924	2,501	3
D	---	---	---	314	0,408	0

TABELLE 16 Neue-Partei-Paradoxon

32 PUKELSHEIM, Friedrich, Divisor oder Quote? Zur Mathematik von Mandatszuteilungen bei Verhältniswahlen, Report 392, Institut für Mathematik, Universität Augsburg, Augsburg 1998, S.18

33 KOPFERMANN, Klaus, Mathematische Aspekte der Wahlverfahren. Mandatsverteilung bei Abstimmungen, Mannheim 1991, S.115

9.3 Das Verfahren nach d'Hondt

Andere Bezeichnungen für dieses Verfahren:

- ♠ Jefferson-Verfahren
- ♠ Hagenbach-Bischoff-Methode
- ♠ Höchstzahlverfahren
- ♠ Divisorverfahren mit Abrundung

Algorithmus des Verfahrens nach d'Hondt:

- ♠ 1. Schritt: *Ermittlung der Höchstzahlen*
  Die Höchstzahlen H_k erhält man, indem man die Anzahl der Stimmen der einzelnen Parteien nacheinander jeweils durch 1,2,3,… dividiert.
  $$H_k = \frac{S_i}{k}, k \in \mathbb{N}$$
- ♠ 2. Schritt: *Mandatsvergabe*
  Diese Höchstzahlen werden nun absteigend geordnet. Das erste Mandat geht an die Partei mit der größten Höchstzahl, das zweite Mandat an die Partei mit der zweitgrößten Höchstzahl. Dies wird so lange fortgesetzt, bis alle Mandate vergeben sind.

Vorteile des Verfahrens nach d'Hondt

- ♠ Die beim Hare-Verfahren geschilderten Paradoxa (Alabama-Paradoxon, Populations-Paradoxon und Neue-Partei-Paradoxon) können bei dem Verfahren nach d'Hondt nicht auftreten.
- ♠ Einzelne Gremiumsgrößen müssen nicht getrennt voneinander berechnet werden.

Nachteile des Verfahrens nach d'Hondt

- ♠ Aufwendiger (Höchstzahlverfahren) bzw. nicht so leicht nachvollziehbar (Divisormethode) wie das Verfahren nach Hare.
- ♠ Benachteiligung kleinerer Parteien. (siehe Beispiel 9.10)
- ♠ Mehrdeutigkeit: Falls eine Höchstzahl mehrmals vorkommt, so können die Mandate nur durch Los zugewiesen werden. (siehe Beispiel 9.11)
- ♠ Die Quotenbedingung wird nicht erfüllt. (siehe Beispiel 9.12)

Beispiel 9.9 *Fortsetzung von Beispiel 9.1*

Wendet man das Verfahren nach d'Hondt auf die Nationalratswahl 2006 an, so erhält man zunächst die folgenden ersten Höchstzahlen:

	SPÖ	ÖVP	Grüne	FPÖ	BZÖ
:1	1663986	1616493	520130	519598	193539
:2	831993	808246,5	260065	259799	96769,5
:3	554662	538831	173376,667	173199,333	64513
:4	415996,5	404123,25	130032,5	129899,5	48384,75
:5	332797,2	323298,6	104026	103919,6	38707,8

Tabelle 17 Berechnung der ersten Höchstzahlen anhand der Nationalratswahl 2006

Der Größe nach geordnet weisen diese Höchstzahlen nun den entsprechenden Parteien ihre jeweiligen Mandatssitze zu.

Höchst-zahlen	Partei-zugehörigkeit	Sitz	SPÖ	ÖVP	Grüne	FPÖ	BZÖ
1663986	SPÖ	1	1	0	0	0	0
1616493	ÖVP	2	1	1	0	0	0
831993	SPÖ	3	2	1	0	0	0
808246,5	ÖVP	4	2	2	0	0	0
554662	SPÖ	5	3	2	0	0	0
538831	ÖVP	6	3	3	0	0	0
520130	Grüne	7	3	3	1	0	0
519598	FPÖ	8	3	3	1	1	0
415996,5	SPÖ	9	4	3	1	1	0
404123,25	ÖVP	10	4	4	1	1	0
…	…	…	…	…	…	…	…
24470,382	SPÖ	183	68	66	21	21	7

TABELLE 18 Nationalratswahl 2006, Mandatsverteilung nach d'Hondt

Mandatsberechnung mithilfe eines Divisors:

Alternativ zu dem vorgestellten Höchstzahl-Algorithmus können die Mandate M_i auch berechnet werden, indem man die Stimmen S_i der einzelnen Parteien durch einen passenden Divisor d dividiert und anschließend diesen Bruch abrundet.

$$M_i = \left\lfloor \frac{S_i}{d} \right\rfloor$$

Laut dem deutschen Mathematiker Friedrich Pukelsheim ist die beste Divisorwahl durch

$$d = \frac{S}{M + \frac{\# P_i}{2}}$$

gegeben.[34]

Bei der Nationalratswahl 2006 ergab sich folgender Divisor

$$d = \frac{4513746}{183 + \frac{5}{2}} \cong 24332{,}863$$

BEISPIEL 9.10 *Benachteiligung kleinerer Parteien*

Betrachten wir erneut die Nationalratswahl 2006. Mit dem Hare-Verfahren bekommt das BZÖ bereits bei 36 zu vergebenden Parlamentssitzen ein zweites Mandat, mit dem Verfahren nach d'Hondt erst bei 45.

Die Tatsache, dass das Verfahren nach d'Hondt größere Parteien begünstigt, kann unter Umständen auch ein Vorteil sein, wenn man nämlich an stabilen politischen Verhältnissen interessiert ist.[35]

[34] PUKELSHEIM, Divisor oder Quote, S.5

BEISPIEL 9.11 *Mehrdeutigkeit beim Verfahren nach d'Hondt*

Die Ausgangsmenge betrage 356. Partei *A* habe 203 Stimmen erhalten und Partei *B* 119. Auf Partei *C* entfielen 34 Stimmen. Falls das Gremium eine Größe von 19 haben soll, dann tritt an dieser Stelle eine Mehrdeutigkeit auf.[36]

Höchstzahlen	Parteizugehörigkeit	Sitz	A	B	C
18,455	A	18	11	6	1
17	B oder C	19	11	6(+1?)	1(+1?)
17	B oder C	20	11	7	2
16,917	A	21	12	7	2

TABELLE 19 Mehrdeutigkeit beim Verfahren nach d'Hondt

BEISPIEL 9.12 *Verletzung der Quotenbedingung*

Bei einer Wahl seien 10000 gültige Stimmen abgegeben worden. Partei *A* habe 2000 Stimmen, Partei *B* 390, Partei *C* 380, Partei *D* 370 und Partei *E* 360 Stimmen erhalten, Parteien *F* bis *Y* jeweils 325 Stimmen.[37] Mit Hilfe des Verfahrens nach d'Hondt sollen nun fünf Mandate auf diese Parteien aufgeteilt werden.

Obwohl die Quote für Partei *A*

$$Q_A = \frac{2000}{10000} \cdot 5 = 1$$

beträgt, würde Partei *A* nach dem Verfahren nach d'Hondt alle fünf Mandate erhalten.

[35] MEYER, Jörg, Wahlen: Paradoxa bei der Sitzverteilung, In: Mathematica didactica Vol. 18(1995), Heft 1, S.27

[36] MAUSBERG, Wolfgang (Hg.), Anteile, Zugriffe und Reihenfolgen (AZUR). Berechnungsverfahren für Sitzverteilungen und Reihenfolgen bei der Besetzung von Gremien und für die Zuteilung sonstiger Berechtigungen im parlamentarischen Bereich. ZI 5 – Arbeitspapier 1998/002, Bonn 1998, S.13

[37] SCHICK, Karl, Wahlberechnungsverfahren, In: Praxis der Mathematik Vol.28(1986), Heft 2, S.114

9.4 DIE MANDATSVERTEILUNG NACH DER ÖSTERREICHISCHEN NATIONALRATSWAHL 2006

9.4.1 ÜBERBLICK ÜBER DIE MANDATSVERTEILUNG

Die Mandatsvergabe ist in Österreich im Bundesgesetz über die Wahl des Nationalrates aus dem Jahr 1992 geregelt.[38] Alle Paragraphen aus dem folgenden Abschnitt beziehen sich auf dieses Gesetz.

Österreich ist in neun Landeswahlkreise gegliedert. Jedes Bundesland entspricht dabei einem dieser Landeswahlkreise, welche ihrerseits in 43 Regionalwahlkreise unterteilt sind. (§2,1 und §3,1)

Aufgrund der letzten Volkszählung wird zunächst einmal die sogenannte Verhältniszahl gebildet, die der Quotient aus der Zahl der Staatsbürger (inklusive Auslandsösterreicher) und der Zahl 183 ist.

BEISPIEL 9.13 *Berechnung der Verhältniszahl für die Nationalratswahl 2006*

2001 gab es laut Volkszählung insgesamt 7383035 österreichische Staatsbürger, aufgeteilt in 7321971 Personen, die in Österreich lebten und 61064 Auslandsösterreicher.

Landeswahlkreis	inländische Staatsbürger	Auslands-österreicher	Gesamt
Burgenland	264995	1025	266020
Kärnten	527275	5843	533118
Niederösterreich	1451762	4845	1456607
Oberösterreich	1277030	8216	1285246
Salzburg	454934	4278	459212
Steiermark	1129735	10244	1139979
Tirol	609897	6351	616248
Vorarlberg	304351	3785	308136
Wien	1301992	16477	1318469
Gesamt	7321971	61064	7383035

Tabelle 20 Ergebnis der letzten Volkszählung aus dem Jahr 2001[39]

Wir erhalten die folgende Verhältniszahl

$$\frac{7383035}{183} \cong 40344{,}453$$ [40]

Die Mandate werden anschließend den einzelnen Landeswahlkreisen nach dem Verfahren nach Hare zugewiesen. Die dafür benötigte Quote erhält man als Quotient der Bevölkerung eines Bundeslandes und der soeben berechneten Verhältniszahl. (§4,3-4,4)

[38] downloadbar unter http://www.bmi.gv.at/wahlen/Nationalrats-Wahlordnung_1992.pdf [03.02.2009]

[39] Bundesministerium für Inneres, Amtsblatt vom 23.9.2004

[40] Ergebnis wurde abgerundet, wie auch in Beispiel 9.14

BEISPIEL 9.14 *Mandatszuteilung an die Landeswahlkreise aufgrund der Volkszählung im Jahr 2001*

Aufgrund des Verfahrens nach Hare ergab sich für die Nationalratswahl 2006 folgende Mandatszuteilung an die Landeswahlkreise

Landes-wahlkreis	Quote	Rest-reihenfolge	Mandate
Burgenland	6,593	4	$6+1=7$
Kärnten	13,241	8	$13+0=13$
Niederösterreich	36,104	9	$36+0=36$
Oberösterreich	31,856	1	$31+1=32$
Salzburg	11,382	5	$11+0=11$
Steiermark	28,256	7	$28+0=28$
Tirol	15,274	6	$15+0=15$
Vorarlberg	7,637	3	$7+1=8$
Wien	32,680	2	$32+1=33$

TABELLE 21 Nationalratswahl 2006, Mandatszuteilung an die Landeswahlkreise aufgrund der Volkszählung im Jahr 2001 nach Hare

Zahl der Mandate in den Regionalwahlkreisen
Die Ermittlung der einzelnen Mandate der Regionalwahlkreise erfolgt nach dem selben Prinzip wie die Vergabe Mandate der Landeswahlkreis. (§4,5)

Dreistufiges Verfahren der Mandatsvergabe
In Österreich erfolgt die Mandatsvergabe in 3 Schritten:

- ♠ Vergabe der Mandate der Regionalwahlkreise
- ♠ Vergabe der Mandate der Landeswahlkreise
- ♠ Vergabe der Mandate des Bundeswahlkreises

9.4.2 ERSTER SCHRITT: DIE MANDATE DER REGIONALWAHLKREISE

Algorithmus zur Mandatsberechnung im 1.Schritt:

- ♠ *Ermittlung der Wahlzahl WZ_1*
  Die Wahlzahl eines Landeswahlkreises ist der aufgerundete Quotient aus der Anzahl aller dort abgegebenen gültigen Stimmen und der Anzahl der entsprechenden Mandate. (§ 96,4)
  $$WZ_1(L_i) = \left\lfloor \frac{\text{Anzahl der gültigen Stimmen in } L_i}{\text{Anzahl der Mandate von } L_i} \right\rfloor + 1$$
  Die Anzahl der Mandate von L_i wurde in Beispiel 9.14 berechnet.
- ♠ *Ermittlung der Regionalwahlkreis-Mandate M_R einer bestimmten Partei P_i*
  Jede Partei erhält so viele Mandate, wie die Wahlzahl in der Anzahl der gültigen Regionalwahlkreisstimmen dieser Partei enthalten ist. (§ 97)
  $$M_1(P_j) = \left\lfloor \frac{\text{Anzahl der gültigen Stimmen von } P_j \text{ im RKW}}{\text{Wahlzahl } L_i} \right\rfloor$$

BEISPIEL 9.15 *Berechnung der Wahlzahl von Wien*

$$WZ_1(\text{Wien}) = \left\lfloor \frac{808\,571}{33} \right\rfloor + 1 = 24\,503$$

Landeswahlkreis	gültige Stimmen[41]	Mandate	Wahlzahl
Burgenland (1)	188 823	7	26 975
Kärnten (2)	327 544	13	25 196
Niederösterreich (3)	1 004 901	36	27 914
Oberösterreich (4)	832 452	32	26 015
Salzburg (5)	282 415	11	25 675
Steiermark (6)	727 829	28	25 994
Tirol (7)	361 783	15	24 119
Vorarlberg (8)	173 963	8	21 746
Wien (9)	808 571	33	24 503

TABELLE 22 Übersicht über die unterschiedlichen Wahlzahlen der einzelnen Landeswahlkreise zur Bestimmung der Mandate für die Regionalwahlkreise aufgrund der Nationalratswahl 2006

BEISPIEL 9.16 *Berechnung der Mandate im Regionalwahlkreis Wien Süd-West (9E)*

2006 betrug die Wahlzahl von Wien Süd West $24\,503$ (siehe Beispiel 9.15)
Die Grünen erhielten etwa im Regionalwahlkreis 9E insgesamt 26.480 gültige Stimmen und somit 1 Mandat.

$$M_1(\text{Grüne in 9E}) = \left\lfloor \frac{26\,480}{24\,503} \right\rfloor \cong \lfloor 1{,}08 \rfloor = 1$$

Parteien in 9E	gültige Stimmen in 9E	Mandate in 9E
SPÖ	59 086	2
ÖVP	38 521	1
Grüne	26 480	1
FPÖ	19 686	0
BZÖ	2 747	0

TABELLE 23 Übersicht über die erhaltenen Mandate der Parteien im Regionalwahlkreis 9E aufgrund der Nationalratswahl 2006[42]

Die Mandate, die im ersten Ermittlungsschritt in den Regionalwahlkreisen vergebenen werden, nennt man Grundmandate.[43]

---

[41] http://www.bmi.gv.at/wahlen/NRW_06_gesamtergE.asp [03.02.2009]
[42] http://wahl06.bmi.gv.at/gkz_9.htm [03.02.2009]
[43] Zentrum Polis. Politik Lernen in der Schule (Hg.), Nationalratswahlen und Wahlrecht in Österreich. Polis aktuell Nr. 7/2006, S.7

Landeswahlkreis	SPÖ	ÖVP	Grüne	FPÖ	BZÖ	Σ
Burgenland (1)	2	2	0	0	0	4
Kärnten (2)	3	0	0	0	0	3
Niederösterreich (3)	9	11	0	0	0	20
Oberösterreich (4)	10	7	0	1	0	18
Salzburg (5)	2	2	0	0	0	4
Steiermark (6)	7	6	0	0	0	13
Tirol (7)	1	4	0	0	0	5
Vorarlberg (8)	0	2	0	0	0	2
Wien (9)	10	2	2	1	0	15
Grundmandate	44	36	2	2	0	84

TABELLE 24 Ergebnis der Mandatsverteilung nach dem ersten Schritt aufgrund der Nationalratswahl 2006[44]

Im ersten Schritt wurden also 84 Mandate von 183 möglichen Mandaten vergeben. Somit waren noch 99 Mandate offen.

9.4.3 ZWEITER SCHRITT: DIE MANDATE DER LANDESWAHLKREISE

Um am zweiten Ermittlungsverfahren teilnehmen zu können, muss eine Partei entweder in einem der Regionalwahlkreise ein Mandat erhalten oder bundesweit mehr als 4% der gültigen Stimmen erreicht haben. (§100,1)

Somit nahmen nach der Nationalratswahl 2006 nur noch SPÖ, ÖVP, Grüne, FPÖ und BZÖ am zweiten Ermittlungsverfahren teil.

Algorithmus zur Mandatsberechnung im 2.Schritt

Jede Partei erhält so viele Mandate, wie die Wahlzahl in der Anzahl der gültigen Landeswahlkreisstimmen dieser Partei enthalten ist. Von dieser Zahl werden danach die im ersten Ermittlungsverfahren erzielten Mandate abgezogen. (§ 101)

$$M_2\left(P_j\right)=\left\lfloor\frac{\text{Anzahl der gültigen Stimmen von } P_j \text{ im LKW}}{\text{Wahlzahl } L_i}\right\rfloor-M_R\left(P_j\right)$$

BEISPIEL 9.17 *Mandate des BZÖ in Kärnten*

In Kärnten erhielt das BZÖ bei der Nationalratswahl 2006 insgesamt 81574 gültige Stimmen. Im ersten Durchgang konnte das BZÖ noch keine Mandate in Kärnten erringen. Somit erhielt sie nun drei Mandate.

$$M_2\left(\text{BZÖ in Kärnten}\right)=\left\lfloor\frac{81574}{25196}\right\rfloor-0=\lfloor 3{,}24\rfloor=3$$

[44] http://www.bmi.gv.at/wahlen/wahldownloads/NRW_06/NRW_06_MandatsspiegelV.pdf [03.02.2009]

Kärnten (2)	SPÖ	ÖVP	Grüne	FPÖ	BZÖ
Erster Schritt	3	0	0	0	0
Zweiter Schritt	1	2	0	0	3
Gesamtmandate	4	2	0	0	3

TABELLE 25 Vergleich der Mandatsverteilungen nach den ersten beiden Schritten in Kärnten aufgrund der Nationalratswahl 2006[45]

Landeswahlkreis	SPÖ	ÖVP	Grüne	FPÖ	BZÖ	Σ
Burgenland (1)	3	2	0	0	0	5
Kärnten (2)	4	2	0	0	3	9
Niederösterreich (3)	13	14	3	3	0	33
Oberösterreich (4)	11	11	3	3	0	28
Salzburg (5)	3	4	1	1	0	9
Steiermark (6)	10	10	2	2	0	24
Tirol (7)	3	6	1	1	0	11
Vorarlberg (8)	1	3	1	0	0	5
Wien (9)	13	7	5	4	0	29
Gesamtmandate	61	59	16	14	3	153

TABELLE 26 Ergebnis der Mandatsverteilung nach dem zweiten Schritt aufgrund der Nationalratswahl 2006[46]

Somit waren nach dem zweiten Ermittlungsschritt nur noch 30 der 183 Mandate zu vergeben.

9.4.4 DRITTER SCHRITT: DIE MANDATE DER BUNDESWAHLKREISE

Algorithmus zur Mandatsberechnung im 3.Schritt

- ♠ *Ermittlung der Wahlzahl WZ_3*
  Zur Mandatsberechnung im dritten Ermittlungsverfahren wird das Verfahren nach d'Hondt verwendet. Die benötigte Wahlzahl ist die 183.Höchstzahl. (§ 107,4-5)
- ♠ *Mandatsvergabe*
  Jede Partei erhält jetzt so viele Mandate, wie die Wahlzahl in der Anzahl aller gültigen Stimmen dieser Partei enthalten ist. Von dieser Zahl werden danach die in den ersten beiden Ermittlungsverfahren erhaltenen Mandate abgezogen (§ 107,6; § 107,8)

$$M_3\left(P_j\right)=\left\lfloor \frac{\text{Anzahl aller gültigen Stimmen von } P_j}{WZ_3} \right\rfloor - \sum_{i=1}^{2} M_i\left(P_j\right)$$

[45] http://www.bmi.gv.at/wahlen/wahldownloads/NRW_06/NRW_06_MandatsspiegelV.pdf [03.02.2009]
[46] http://www.bmi.gv.at/wahlen/wahldownloads/NRW_06/NRW_06_MandatsspiegelV.pdf [03.02.2009]

BEISPIEL 9.18 *Berechnung der Wahlzahl WZ_3*

2006 betrug die Wahlzahl im dritten Ermittlungsverfahren

$$WZ_3 = \frac{1\,663\,986}{68}$$

BEISPIEL 9.19 *Berechnung der Mandate im dritten Ermittlungsverfahren*

2006 erhielt etwa die ÖVP im dritten Ermittlungsverfahren

$$M_3(\text{ÖVP}) = \left\lfloor \frac{1\,616\,493}{WZ_3} \right\rfloor - 36 - 23 = 7$$

Mandate.

Mandate österreichweit	SPÖ	ÖVP	Grüne	FPÖ	BZÖ
Erster Schritt	44	36	2	2	0
Zweiter Schritt	17	23	14	12	3
Dritter Schritt	7	7	5	7	4
Gesamtmandate	68	66	21	21	7

TABELLE 27 Schrittweise Mandatsverteilung anhand der Nationalratswahl 2006[47]

BEISPIEL 9.20 *Vergleich der Verfahren nach Hare und d'Hondt*

Wendet man statt dem Verfahren nach d'Hondt das Verfahren nach Hare an, so würde ein Mandat von der SPÖ zum BZÖ abwandern.

Verfahren	SPÖ	ÖVP	Grüne	FPÖ	BZÖ
d'Hondt	68	66	21	21	7
Hare	67	66	21	21	8

TABELLE 28 Vergleich der Verfahren nach d'Hondt und Hare anhand der Nationalratswahl 2006[48]

[47] http://www.bmi.gv.at/wahlen/wahldownloads/NRW_06/NRW_06_MandatsspiegelV.pdf [03.02.2009]

[48] http://www.bmi.gv.at/wahlen/wahldownloads/NRW_06/NRW_06_MandatsspiegelV.pdf [03.02.2009]

Verwendete Literatur zu Mandatsverteilungen:

- ♠ CARNAL, Henri / RIEDWYL, Hans, Wer kommt ins Parlament?, In: Spektrum der Wissenschaft 09/2002, S.80-84
- ♠ HESSE Christian, Angewandte Wahrscheinlichkeitstheorie, Braunschweig / Wiesbaden 2003
- ♠ HESSE, Christian / MEISTER, Alexander, Übungsbuch zur angewandten Wahrscheinlichkeitstheorie. Aufgaben und Lösungen, Wiesbaden 2005
- ♠ KOPFERMANN, Klaus, Mathematische Aspekte der Wahlverfahren. Mandatsverteilung bei Abstimmungen, Mannheim 1991
- ♠ MAUSBERG, Wolfgang (Hg.), Anteile, Zugriffe und Reihenfolgen (AZUR). Berechnungsverfahren für Sitzverteilungen und Reihenfolgen bei der Besetzung von Gremien und für die Zuteilung sonstiger Berechtigungen im parlamentarischen Bereich. ZI 5 – Arbeitspapier 1998/002, Bonn 1998[49]
- ♠ MEYER, Jörg, Wahlen: Paradoxa bei der Sitzverteilung, In: Mathematica didactica Vol. 18(1995), Heft 1, S.21-34
- ♠ PUKELSHEIM, Friedrich, Die Väter der Mandatszuteilungsverfahren, In: Spektrum der Wissenschaft, September 2002, S.83[50]
- ♠ PUKELSHEIM, Friedrich, Divisor oder Quote? Zur Mathematik von Mandatszuteilungen bei Verhältniswahlen, Report 392, Institut für Mathematik, Universität Augsburg, Augsburg 1998[51]
- ♠ RÖTTEL, Karl, Berechnung von Sitzverteilungen in Ausschüssen, In: Arbeitskreis Mathematik und Bildung (Hg.), Mathematik – unsichtbar, doch allgegenwärtig, Eichstätt 2002, S.75-98
- ♠ SCHICK, Karl, Wahlberechnungsverfahren, In: Praxis der Mathematik Vol. 28(1986), Heft 2, S.81-88, 105-118
- ♠ SCHWÄRZLER, Jürgen, Methodische Überlegungen zu Wählerstromanalyse und Wahlhochrechnung, Diplomarbeit, Universität Wien, Wien 2000, S.118-132
- ♠ ZENTRUM POLIS – Politik Lernen in der Schule (Hg.), Nationalratswahlen und Wahlrecht in Österreich. Polis aktuell Nr. 7/2006, Wien 2006[52]
- ♠ http://www.wahlrecht.de

---

[49] downloadbar unter http://webarchiv.bundestag.de/archive/2005/0919/ftp/pdf_arch/azur.pdf [03.02.2009] oder unter http://www.bundestag.eu/ftp/pdf_arch/azur.pdf [03.02.2009]

[50] Online-Fassung unter http://www.math.uni-augsburg.de/stochastik/pukelsheim/2002g.html abrufbar [03.02.2009]

[51] downloadbar unter http://www.wahlrecht.de/doku/download/index.htm [03.02.2009]

[52] downloadbar unter http://www.politik-lernen.at/goto/polis/details/pa_nationalratswahlen/ [03.02.2009]

10. MACHTINDIZES

In diesem Kapitel möchte ich mich mit dem Machtbegriff mathematisch auseinandersetzen und erklären, wie sich Macht mit Hilfe von bestimmten Indizes angeben lässt. Machtindizes sind zwar ein Teilgebiet der Spieltheorie, doch werden hierbei lediglich die kombinatorischen Grundlagen aus Kapitel 2 benötigt. Ich werde den Shaplex-Index, den Banzhaf-Indez, den Johnston Index und den Deegan-Packel-Index einführen.

Erstmals entwickelte der Mathematiker Lloyd S. Shapley einen Machtindex, der die Macht mithilfe von Permutationen angab. Da an der Entwicklung dieses Index auch der Wirtschaftswissenschafter Martin Shubik beteiligt war, wird er Shapley-Index gelegentlich auch als Shapley-Shubik-Index bezeichnet.

Die Wurzeln des Banzhaf-Index gehen auf Penrose zurück, der bereits 1946 diesen Index einführte. Benannt ist dieser Index jedoch nach John F. Banzhaf, der diesen Index 1965 anlässlich eines Rechtsstreits über Stimmgewichte in den USA verwendete. Dieser Index misst die Macht anhand von Koalitionen und nicht anhand von Permutationen, wie dies beim Shapley-Index der Fall ist.

Im Unterschied zum Banzhaf-Index berücksichtigt der Johnston-Index, wie viele Spieler eine Gewinnkoalition in eine Verlustkoalition verwandeln können. Ein Spieler soll eine größere Macht besitzen, wenn er als einziger Spieler einer Koalition entscheidend ist.

Der letzte Machtindez, den ich behandeln möchte, ist der Deegan-Packel-Index, der minimale Gewinnkoalitionen betrachtet. Dieser Indez wurde von den Namensgebern 1978 im „*International Journal of Game Theory*“ vorgeschlagen.

Anhand von drei konkreten Beispielen (Machtverhältnisse in der EWG, Österreichs Einfluss im Weltsicherheitsrat und die Macht des amerikanischen Präsidenten) werde ich zeigen, dass diese vier Indizes zu gänzlich unterschiedlichen Ergebnissen führen können.

10.1 KURZE EINFÜHRUNG IN DIE SPIELTHEORIE

Die *Spieltheorie* ist ein Teilgebiet der Mathematik, das strategische Entscheidungssituationen analysiert. Holler / Illing nennen dabei die folgenden vier Situationen[53]:

- ♠ Das Ergebnis hängt von den Entscheidungen mehrerer Entscheidungsträger ab, so dass ein einzelner das Ergebnis nicht unabhängig von der Wahl der anderen bestimmen kann.
- ♠ Jeder Entscheidungsträger ist sich dieser Interdependenz bewusst.
- ♠ Jeder Entscheidungsträger geht davon aus, dass alle anderen sich ebenfalls der Interdependenz bewusst sind.
- ♠ Jeder berücksichtigt bei seinen Entscheidungen die drei vorhergehenden Punkte.

DEFINITION 10.1 *Spiel*

Sei $A = \{1, 2, ..., n\}$ eine Menge von n Spielern und S_i die nichtleere Menge der Entscheidungsstrategien des i-ten Spielers. Dann versteht man unter einem Spiel das Zahlentripel

$$(A; S_1, ..., S_n).$$

Unter einer Abstimmung bzw. Wahl verstehen wir nun ein Spiel mit n Spielern, wobei dem i-ten Spieler S_i Stimmen bzw. Stimmgewichte zustehen. Jede nichtleere Teilmenge der Spieler nennen wir eine Koalition. Bei einer Gewinnkoalition müssen deren Mitglieder zusammen auf mindestens Q Stimmen (= Quote) kommen. Solch eine Gegebenheit fassen wir als $(Q; S_1, ..., S_n) -$ Spiel auf.

Man spricht von einem gewichteten Wahlsystem, wenn die einzelnen Wähler über eine unterschiedliche Anzahl an Stimmen verfügen. Beispiele hierfür wären etwa Parlamentsabstimmungen, wo jede Fraktion unterschiedlich stark vertreten ist, sowie die Wahl des amerikanischen Präsidenten, wo alle Wahlmänner eines Staates ihre Stimme einem einzigen Kandidaten geben.

BEISPIEL 10.2 *Die Europäische Wirtschaftsgemeinschaft (EWG)*

Die Europäische Wirtschaftsgemeinschaft war anfangs ein Zusammenschluss von sechs europäischen Staaten. Sie wurde am 25.März 1957 von Deutschland, Frankreich, Italien, Belgien, Niederlande und Luxemburg durch die Unterzeichnung der Römischen Verträge gegründet und 1993 in Europäische Gemeinschaft unbenannt. Deutschland, Frankreich und Italien hatten 1958 ein Stimmgewicht von jeweils vier, Belgien und die Niederlande von zwei und Luxemburg von eins. Mindestens zwölf dieser siebzehn Stimmen waren für einen Mehrheitsbeschluss notwendig. Somit war die EWG ein $(12; 4, 4, 4, 2, 2, 1) -$ Spiel.

[53] HOLLER, Manfred / ILLING, Gerhard, Einführung in die Spieltheorie, Berlin [6]2006, S.1

BEISPIEL 10.3 *Der Weltsicherheitsrat*[54]

Vor 1966 bestand der Weltsicherheitsrat aus 5 ständigen Mitgliedern (Vereinigte Staaten, Russland, China, Frankreich und Großbritannien), jedoch nur aus 6 nichtständigen Mitgliedern. Um siegreich zu sein, musste damals eine Koalition aus allen ständigen sowie mindestens 2 nichtständigen Mitgliedern bestehen.

Seit 1966 besteht der Weltsicherheitsrat aus 5 ständigen (Vereinigte Staaten, Russland, China, Frankreich und Großbritannien) und 10 nichtständigen Mitgliedern. Um siegreich zu sein, muss eine Koalition aus allen ständigen sowie mindestens 4 nichtständigen Mitgliedern bestehen.

Die 5 ständigen Mitglieder $\left(=S_i\right)$ haben dasselbe Stimmgewicht. Ebenso sind die 10 nichtständigen Mitglieder $\left(=N_i\right)$ untereinander gleich mächtig. Daher gilt

$$S_1 = S_2 = S_3 = S_4 = S_5 = S$$
$$N_1 = N_2 = N_3 = N_4 = N_5 = N_6 = N_7 = N_8 = N_9 = N_{10} = N$$

Man erhält die beiden Ungleichungen

$$\left.\begin{array}{l} 5S + 4N \geq Q \\ \underline{4S + 10N \leq Q - 1} \end{array}\right\} \begin{array}{l} \text{siegreich} \\ \text{nicht siegreich} \end{array}$$

Als Lösung des zugehörigen linearen Gleichungssystems ergibt sich

$$\left.\begin{array}{l} 5S + 4N = Q \\ \underline{4S + 10N = Q - 1} \end{array}\right\} -$$
$$S - 6N = 1$$

Wählt man

$$N = 1 \Rightarrow S = 7, Q = 39\,,$$

dann handelt es sich bei dieser Abstimmung um ein

$$\left(39; 7, 7, 7, 7, 7, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1\right) - \text{Spiel}\,.$$

> **DEFINITION 10.4** *entscheidende Spieler*
>
> Wenn Spieler S_i von einer Gewinnkoalition entfernt wird und die verbleibende Koalition keine Gewinnkoalition mehr ist, dann nennen wir Spieler *i* entscheidend. Solch eine Koalition ist also nur aufgrund der Beteiligung von Spieler *i* siegreich.

[54] HESSE Christian, Angewandte Wahrscheinlichkeitstheorie, Braunschweig / Wiesbaden 2003, S.134 bzw. HESSE, Christian / MEISTER, Alexander, Übungsbuch zur angewandten Wahrscheinlichkeitstheorie. Aufgaben und Lösungen, Wiesbaden 2005, S.93

BEISPIEL 10.5 *Das Problem des Machtbegriffes*

Betrachten wir die beiden Spiele $(51;49,48,3)$ und $(51;26,26,26,22)$[55]. Im ersten Fall hat jede Koalition aus zwei beliebigen Teilnehmern die Mehrheit inne, im zweiten Fall können je zwei der ersten drei Spieler eine Mehrheit ohne Spieler 4 bilden, während der vierte Spieler für eine Mehrheit mindestens zwei der drei ersten Spieler benötigt. Obwohl dieser Teilnehmer ein mehr als siebenmal so großes Stimmgewicht wie Spieler 3 im ersten Beispiel besitzt, so ist seine Macht erheblich geringer.

Macht steht also keineswegs im Zusammenhang zum prozentuellen Stimmanteil, wie man intuitiv vielleicht vermuten würde. Eine Möglichkeit der Machtbeschreibung liefern die sogenannten Machtindizes. Ich möchte jetzt vier derartige Indizes näher besprechen, die versuchen Macht mit mathematischen Mitteln zu bestimmen. Dabei wird sich zeigen, dass diese Indizes teilweise recht unterschiedliche Ergebnisse liefern können.

[55] LUCAS, William, Gewichtete Wahlsysteme: Wie man Macht messen kann, In: Garfunkel, Salomon / Steen, Lynn (Hg.), Mathematik in der Praxis. Anwendungen in Wirtschaft, Wissenschaft und Politik, Heidelberg 1989, S.141

10.2 DER SHAPLEY-INDEX

DEFINITION 10.6 *Der Shapley-Index* (= *SI*)

Seien $S_1, S_2, ..., S_n$ die Spieler in einem gewichteten Mehrheitssystem. Eine Permutation π heißt Permutation, in der Spieler Si entscheidend ist, falls $i = \pi(j), \{S_{\pi(1)}, S_{\pi(2)}, ..., S_{\pi(j-1)}\}$ keine Gewinnkoalition und $\{S_{\pi(1)}, S_{\pi(2)}, ..., S_{\pi(j-1)}, S_i\}$ Gewinnkoalition ist. Dann ist der Shapley-Index eines Spielers S_i definiert als der Quotient der Anzahl der Permutationen, in denen S_i entscheidend ist, und der Anzahl aller Permutationen.

PROPOSITION 10.7 *Eigenschaften des Shapley-Indizes*

(i) $0 \leq SI(S_i) \leq 1$

(ii) $\sum_{i=1}^{n} SI(S_i) = 1$

BEISPIEL 10.8 *Der Shapley-Index im* $(51; 50, 49, 1)$ – Spiel [56]

Insgesamt gibt es die folgenden sechs möglichen Koalitionen, wobei die Spieler, die die nicht mehrheitsfähigen Koalitionen aus den links befindlichen Spielern zu einer Gewinnkoalition verwandeln, jeweils durch ein Kästchen gekennzeichnet sind

$$\{S_1, \boxed{S_2}\}, \{S_1, \boxed{S_3}\},$$
$$\{S_2, \boxed{S_1}\}, \{S_2, S_3, \boxed{S_1}\},$$
$$\{S_3, \boxed{S_1}\}, \{S_3, S_2, \boxed{S_1}\}$$

Der Shapley-Index lässt sich nun problemlos abzählen

$$SI(S_1) = \frac{4}{6}$$
$$SI(S_2) = SI(S_3) = \frac{1}{6}$$

[56] BRAMS, Steven, Game Theory and Politics, New York 1975, S.158ff.

Beispiel 10.9 *Verletzung der Monotonie bei Veränderung der Stimmgewichte*[57]

Zunächst betrachten wir das $(70;50,25,25)-\text{Spiel}$, mit folgenden Koalitionen, die durch Hinzufügen des entscheidenden Spielers $\boxed{S_i}$ in Gewinnkoalitionen verwandelt werden:

$$\left\{S_1,\boxed{S_2}\right\},\left\{S_1,\boxed{S_3}\right\},\left\{S_2,\boxed{S_1}\right\},\left\{S_2,S_3,\boxed{S_1}\right\},\left\{S_3,\boxed{S_1}\right\},\left\{S_3,S_2,\boxed{S_1}\right\}$$

Wandern nun insgesamt fünf Stimmen von Spieler 3 zu Spieler 1 und zehn Stimmen zu Spieler 2, so erhalten wir ein $(70;55,35,10)-\text{Spiel}$ mit den folgenden Gewinnkoalitionen:

$$\left\{S_1,\boxed{S_2}\right\},\left\{S_1,S_3,\boxed{S_2}\right\},\left\{S_2,\boxed{S_1}\right\},\left\{S_2,S_3,\boxed{S_1}\right\},\left\{S_3,S_1,\boxed{S_2}\right\},\left\{S_3,S_2,\boxed{S_1}\right\}$$

Somit besitzt Spieler 1 im $(70;50,25,25)-\text{Spiel}$ einen Shapley-Index von 2/3, im $(70;55,35,10)-\text{Spiel}$ jedoch einen Shapley-Index von nur noch 1/2. Trotz Stimmenzuwachs hat sich also seine Macht verringert.

57 HOLLER / ILLING, Spieltheorie, S.308

10.3 Der Banzhaf-Index

Definition 10.10 *Die totale Banzhaf-Macht* (= *TBM*)

Sei S_i ein Spieler in einem gewichteten Mehrheitssystem. Dann ist die totale Banzhaf-Macht die Anzahl der Gewinnkoalitionen, die den beiden folgenden Bedingungen genügen:
(i) S_i *ist ein Mitglied einer Gewinnkoalition*
(ii) S_i ist entscheidend

Definition 10.11 *Der Banzhaf-Index* (= *BI*)

Seien $S_1, S_2, ..., S_n$ die Spieler in einem gewichteten Mehrheitssystem. Dann ist der Banzhaf-Index eines Spielers S_i definiert als

$$BI(S_i) = \frac{TBM(S_i)}{\sum_{i=1}^{n} TBM(S_i)}.$$

Proposition 10.12 *Eigenschaften des Banzhaf-Indizes*

(i) $0 \le BI(S_i) \le 1$

(ii) $\sum_{i=1}^{n} BI(S_i) = 1$

Beispiel 10.13 *Der Banzhaf-Index im* $(51; 50, 49, 1)$ – Spiel

Offensichtlich gibt es in diesem Spiel die drei Gewinnkoalitionen

$$\{\boxed{S_1}, \boxed{S_2}\}, \{\boxed{S_1}, \boxed{S_3}\}, \{\boxed{S_1}, S_2, S_3\}$$

wobei S_1 überall entscheidend ist, S_2 nur in der ersten und S_3 nur in der zweiten Gewinnkoalition.

Somit erhält man die totale Banzhaf-Macht

$$TBM(S_1) = 3$$

$$TBM(S_2) = TBM(S_3) = 1$$

und daraus die Banzhaf-Indizes

$$BI(S_1) = \frac{3}{3+1+1} = \frac{3}{5}$$

$$BI(S_2) = BI(S_3) = \frac{1}{3+1+1} = \frac{1}{5}$$

Beispiel 10.14 *Schenkungs-Paradoxon*[58]

Felsenthal und Machover geben das folgende Paradoxon an. Trotz Stimm-Abgabe eines Spielers an einen anderen Spieler kann dieser Spieler seine Macht vergrößern.[59]

Zunächst betrachten wir das $(8;5,3,1,1,1)-\text{Spiel}$, wo es die folgenden neun Gewinnkoalitionen gibt

Stimmgewicht 8: $\{\boxed{S_1},\boxed{S_2}\},\{\boxed{S_1},\boxed{S_3},\boxed{S_4},\boxed{S_5}\}$

Stimmgewicht 9: $\{\boxed{S_1},\boxed{S_2},S_3\},\{\boxed{S_1},\boxed{S_2},S_4\},\{\boxed{S_1},\boxed{S_2},S_5\}$

Stimmgewicht 10: $\{\boxed{S_1},\boxed{S_2},S_3,S_4\},\{\boxed{S_1},\boxed{S_2},S_3,S_5\},\{\boxed{S_1},\boxed{S_2},S_4,S_5\}$

Stimmgewicht 11: $\{\boxed{S_1},S_2,S_3,S_4,S_5\}$

Die totale Banzhaf-Macht lässt sich nun problemlos abzählen

$$TBM(S_1)=9$$
$$TBM(S_2)=7$$
$$TBM(S_3)=TBM(S_4)=TBM(S_5)=1$$

Somit erhalten wir die folgenden Banzhaf-Indizes

$$BI(S_1)=\frac{9}{19}$$
$$BI(S_2)=\frac{7}{19}$$
$$BI(S_3)=BI(S_4)=BI(S_5)=\frac{1}{19}$$

Analysieren wir nun das $(8;4,4,1,1,1)-\text{Spiel}$, wo eine Stimme von Spieler S_1 zu S_2 abgewandert ist, so erkennen wir, dass sich dadurch parodoxerweise die Banzhaf-Macht von S_1 auf 1/2 erhöht, während zugleich die Spieler S_3, S_4, S_5 ihre gesamte Macht verlieren, da sie bei keiner Gewinnkoalition mehr entscheidend sind.

---

[58] Englische Bezeichnung: Donation Paradox

[59] FELSENTHAL, DAN S. / MACHOVER, Moshé, The Measuremnet of Voting Power. Theory and Practice, Problems and Paradoxes, Cornwall 1998, S.253

PROPOSITION 10.15 *Berechnung der totalen Banzhaf-Macht*

Seien K_{S_i} die Gewinnkoalitionen mit Beteiligung von Spieler S_i und K_{gesamt} alle möglichen Gewinnkoalitionen. Dann lässt sich die totale Banzhaf-Macht eines Spielers S_i auf folgende Art berechnen:

$$TBM(S_i) = 2 \cdot \# K_{S_i} - \# K_{gesamt}$$

BEWEIS:

Seien K_1 die Gewinnkoalitionen ohne S_i, K_2 die Gewinnkoalitionen K_1 um S_i erweitert und K_3 die Gewinnkoalitonen, wo S_i entscheidend ist. Offensichtlich gilt

$$\# K_1 + \# K_2 + \# K_3 = \# K_{gesamt}$$

Da

$$\# K_1 = \# K_2$$

berechnet sich die totale Banzhaf-Macht wie folgt

$$TBM(S_i) = \# K_3 = \underbrace{2 \cdot \# K_2 - \# K_1 - \# K_2}_{=0} + 2 \cdot \# K_3 - \# K_3 =$$

$$= \underbrace{\left(2 \cdot \# K_2 + 2 \cdot \# K_3\right)}_{=2 \cdot \# K_{S_i}} - \underbrace{\left(\# K_1 + \# K_2 + \# K_3\right)}_{=\# K_{gesamt}}$$

□

Proposition 10.15 ist in vielen Fällen besser geeignet, die totale Banzhaf-Macht eines Spielers auszurechnen als Definition 10.10. Ich werde diese Proposition bei Beispiel 10.25 (Der Banzhaf-Index der EWG im Jahre 1958), sowie bei Beispiel 10.36 (Der Banzhaf-Indez des Präsidenten) anwenden.

10.4 DER JOHNSTON-INDEX

DEFINITION 10.16 *Die totale Johnston-Macht* (= *TJM*)

Sei S_i ein Spieler in einem gewichteten Mehrheitssystem und $K_1, K_2, ..., K_k$ seien die Gewinnkoalitionen, für die Spieler S entscheidend ist. Weiters bezeichne n_j die Anzahl der Spieler, die entscheidend für die betreffende Gewinnkoalition K_j ist. Dann ist die totale Johnston-Macht gegeben durch

$$TJM(S_i) = \sum_{j=1}^{k} \frac{1}{n_j}.$$

DEFINITION 10.17 *Der Johnston-Index* (= *JI*)

Seien $S_1, S_2, ..., S_n$ die Spieler in einem gewichteten Mehrheitssystem. Dann ist der Johnston-Index eines Spielers S_i definiert als

$$JI(S_i) = \frac{TJM(S_i)}{\sum_{i=1}^{n} TJM(S_i)}.$$

BEISPIEL 10.18 *Der Johnston-Index im* $(51; 50, 49, 1)$ – Spiel

Wie wir bereits wissen, gibt es hier insgesamt die drei Gewinnkoalitionen

$$\{\boxed{S_1}, \boxed{S_2}\}, \{\boxed{S_1}, \boxed{S_3}\}, \{\boxed{S_1}, S_2, S_3\}$$

Die Berechnung zeigt, dass in diesem Beispiel der Johnston-Index zufällig mit dem Shapley-Index übereinstimmt.

Gewinn-koalitionen	S_1	S_2	S_3
$\{S_1, S_2\}$	1/2	1/2	0
$\{S_1, S_3\}$	1/2	0	1/2
$\{S_1, S_2, S_3\}$	1	0	0
TJM	2	1/2	1/2
JI	2/3	1/6	1/6

TABELLE 29 Der Johnston-Index im $(51; 50, 49, 1)$ – Spiel

10.5 Der Deegan-Packel-Index

Definition 10.19 *Minimale Gewinnkoalitionen* (= *MGK*)

Unter minimalen Gewinnkoalitionen versteht man diejenigen Koalitionen, wo jedes Mitglied entscheidend ist und durch seinen Austritt die Gewinnkoalition in eine Verlustkoalition verwandeln würde.

Der Deegan-Packel-Index beruht auf den folgenden drei Annahmen:

- ♠ Nur die minimalen Gewinnkoalitionen werden betrachtet.
- ♠ Alle minimalen Gewinnkoalitionen treten mit gleicher Wahrscheinlichkeit auf. Also wird jede Koalition mit gleichem Gewicht berücksichtigt.
- ♠ Alle Spieler, die an einer minimalen Gewinnkoalition beteiligt sind, erhalten denselben Machtanteil.

Definition 10.20 *Die totale Deegan-Packel-Macht* (= *TDPM*)

Sei S_i ein Spieler in einem gewichteten Mehrheitssystem und $MGK_1, MGK_2, ..., MGK_k$ seien die minimalen Gewinnkoalitionen, an denen Spieler S beteiligt ist. Weiters bezeichne n_j die Anzahl der Spieler an der betreffenden Gewinnkoalition MGK_j. Dann ist die Deegan-Packel-Macht eines Spielers S_i gegeben durch

$$TDPM(S_i) = \sum_{j=1}^{k} \frac{1}{n_j}.$$

Definition 10.21 *Der Deegan-Packel-Index* (= DPI)

Seien $S_1, S_2, ..., S_n$ die Spieler in einem gewichteten Mehrheitssystem. Dann ist der Deegan-Packel-Index eines Spielers S_i definiert als

$$DPI(S_i) = \frac{TDPM(S_i)}{\sum_{i=1}^{n} TDPM(S_i)}.$$

Beispiel 10.22 *Der Deegan-Packel-Index im* $(51; 50, 49, 1)$ – Spiel

Hier gibt es lediglich zwei minimale Gewinnkoalitionen, nämlich $\{S_1, S_2\}$ und $\{S_1, S_3\}$. Nach einer kurzen Rechnung erhält man, dass S_1 genauso viel Macht besitzt wie S_2 und S_3 zusammen.

MGK	S_1	S_2	S_3
$\{S_1, S_2\}$	1/2	1/2	0
$\{S_1, S_3\}$	1/2	0	1/2
TDPM	1	1/2	1/2
DPI	1/2	1/4	1/4

Tabelle 30 Der Deegan-Packel-Index im $(51; 50, 49, 1)$ – Spiel

Beispiel 10.23 *Verletzung der Monotonie für gegebene Stimmanteile*

Betrachten wir das $(51;35,20,15,15,15)-\text{Spiel}$. Berechnet man den Deegan-Packel-Index für dieses Spiel, so stellt sich heraus, dass Spieler 2 trotz eines höheren Stimmgewichts weniger Macht besitzt als die Spieler 3, 4 und 5.[60]

MGK	S_1	S_2	S_3	S_4	S_5
$\{S_1, S_2\}$	1/2	1/2	0	0	0
$\{S_1, S_3, S_4\}$	1/3	0	1/3	1/3	0
$\{S_1, S_3, S_5\}$	1/3	0	1/3	0	1/3
$\{S_1, S_4, S_5\}$	1/3	0	0	1/3	1/3
$\{S_2, S_3, S_4, S_5\}$	0	1/4	1/4	1/4	1/4
TDPM	3/2	3/4	11/12	11/12	11/12
DPI	18/60	9/60	11/60	11/60	11/60

Tabelle 31 Der Deegan-Packel-Index im $(51;35,20,15,15,15)-\text{Spiel}$

[60] HOLLER, Spieltheorie, S.324

10.6 VERGLEICH DER MACHTINDIZES

10.6.1 DIE MACHTVERHÄLTNISSE IN DER EWG

Ausgangspunkt unserer Überlegung ist ein $(12; 4, 4, 4, 2, 2, 1)$ – Spiel (siehe Beispiel 10.2).

BEISPIEL 10.24 *Der Shapley-Index der EWG im Jahre 1958*

Insgesamt gab es damals

$$6! = 720 .$$

Permutationen der europäischen Mächte untereinander.

Zur Berechnung des Shapley-Index von Belgien bzw. den Niederlanden müssen wir zunächst zwei Fallunterscheidungen durchführen, da 10 oder 11 andere Stimmen notwendig waren, damit Belgien bzw. die Niederlande mit ihrem Stimmgewicht eine Gewinnkoalition herbeiführen konnte.

Im einfachsten Fall bildete Belgien mit $\{D, F, N\}, \{D, I, N\}$ oder mit $\{F, I, N\}$ eine Gewinnkoalition.

$$\overbrace{\underbrace{D, F, N}_{\text{3!-Möglichkeiten der Anordnung}}\ , B}^{\text{Gewinnkoalition}} \ \Bigg| \ \underbrace{I, L}_{\text{2!-Möglichkeiten der Anordnung}}$$

Im zweiten Fall verhalf Belgien $\{D, F, N, L\}$ (bzw. gleichbedeutend mit $\{D, I, N, L\}$ oder mit $\{F, I, N, L\}$) zu einer Gewinnkoalition.

$$\overbrace{\underbrace{D, F, N, L}_{\text{4!-Möglichkeiten der Anordnung}}\ , B}^{\text{Gewinnkoalition}} \ \Bigg| \ \underbrace{I}_{\text{1!-Möglichkeiten der Anordnung}}$$

Somit erhält man den Shapley-Index von Belgien / den Niederlanden

$$SI(B) = SI(N) = \frac{3 \cdot (3! \cdot 1! \cdot 2! + 4! \cdot 1! \cdot 1!)}{6!} = \frac{9}{60}$$

Da Luxemburg keiner einzigen Koalition zum Sieg verhelfen konnte, galt

$$SI(L) = 0$$

Gemäß Proposition 10.7 (ii) berechnet sich nun der Shapley-Index der restlichen drei Länder nach der Formel

$$SI(D) = SI(F) = SI(I) = \frac{1 - 2 \cdot S(B)}{3} = \frac{14}{60}$$

BEISPIEL 10.25 *Der Banzhaf-Index der EWG im Jahre 1958*

Wenden wir Proposition 10.15 auf diese Fragestellung an, so ergibt sich folgende Tabelle.

GK	D	F	I	B	N	L
$\{D, F, I\}$	2	2	2	0	0	0
$\{D, F, B, N\}$	2	2	0	2	2	0
$\{D, I, B, N\}$	2	0	2	2	2	0
$\{F, I, B, N\}$	0	2	2	2	2	0
$\{D, F, I, L\}$	2	2	2	0	0	2
$\{D, F, B, N, L\}$	2	2	0	2	2	2
$\{D, I, B, N, L\}$	2	0	2	2	2	2
$\{F, I, B, N, L\}$	0	2	2	2	2	2
$\{D, F, I, B\}$	2	2	2	2	0	0
$\{D, F, I, N\}$	2	2	2	0	2	0
$\{D, F, I, B, L\}$	2	2	2	2	0	2
$\{D, F, I, N, L\}$	2	2	2	0	2	2
$\{D, F, I, B, N\}$	2	2	2	2	2	0
$\{D, F, I, B, N, L\}$	2	2	2	2	2	2
Summe	24	24	24	20	20	14
TBM	10	10	10	6	6	0
BI	5/21	5/21	5/21	3/21	3/21	0

TABELLE 32 Auflistung aller Gewinnkoalitionen und die Banzhaf-Indizes in der EWG

Die totale Banzhaf-Macht erhält man, indem man von der Punktesumme die Gesamtzahl der Gewinnkoalitionen, also 14, subtrahiert.

Deutschland, Frankreich und Italien besitzen also jeweils eine Banzhaf-Macht von 5/21, Belgien und Niederlande von 3/21 und Luxemburg erneut von 0.

BEISPIEL 10.26 *Der Johnston-Index der EWG im Jahre 1958*

Die Berechnung zeigt, dass Deutschland, Frankreich und Italien jeweils doppelt so viel Macht besaßen wie Belgien und die Niederlande, während Luxemburg keinerlei Macht zuerkannt wurde.

GK	D	F	I	B	N	L	Gewicht
$\{\boxed{D},\boxed{F},\boxed{I}\}$	1/3	1/3	1/3	0	0	0	12
$\{\boxed{D},\boxed{F},\boxed{B},\boxed{N}\}$	1/4	1/4	0	1/4	1/4	0	12
$\{\boxed{D},\boxed{I},\boxed{B},\boxed{N}\}$	1/4	0	1/4	1/4	1/4	0	12
$\{\boxed{F},\boxed{I},\boxed{B},\boxed{N}\}$	0	1/4	1/4	1/4	1/4	0	12
$\{\boxed{D},\boxed{F},\boxed{I},L\}$	1/3	1/3	1/3	0	0	0	13
$\{\boxed{D},\boxed{F},\boxed{B},\boxed{N},L\}$	1/4	1/4	0	1/4	1/4	0	13
$\{\boxed{D},\boxed{I},\boxed{B},\boxed{N},L\}$	1/4	0	1/4	1/4	1/4	0	13
$\{\boxed{F},\boxed{I},\boxed{B},\boxed{N},L\}$	0	1/4	1/4	1/4	1/4	0	13
$\{\boxed{D},\boxed{F},\boxed{I},B\}$	1/3	1/3	1/3	0	0	0	14
$\{\boxed{D},\boxed{F},\boxed{I},N\}$	1/3	1/3	1/3	0	0	0	14
$\{\boxed{D},\boxed{F},\boxed{I},B,L\}$	1/3	1/3	1/3	0	0	0	15
$\{\boxed{D},\boxed{F},\boxed{I},N,L\}$	1/3	1/3	1/3	0	0	0	15
$\{D,F,I,B,N\}$	0	0	0	0	0	0	16
$\{D,F,I,B,N,L\}$	0	0	0	0	0	0	17
TJM	3	3	3	3/2	3/2	0	
TJ	1/4	1/4	1/4	1/8	1/8	0	

TABELLE 33 Auflistung aller Gewinnkoalitionen und die Johnston-Indizes in der EWG

BEISPIEL 10.27 *Der Deegan-Packel-Index der EWG im Jahre 1958*

Laut dem Deegan-Packel-Index besaßen Deutschland, Frankreich und Italien jeweils 5/24 der gesamten Macht, Belgien und die Niederlande 3/16 und Luxemburg hatte keine Macht inne.

GK	D	F	I	B	N	L
$\{D,F,I\}$	1/3	1/3	1/3	0	0	0
$\{D,F,B,N\}$	1/4	1/4	0	1/4	1/4	0
$\{D,I,B,N\}$	1/4	0	1/4	1/4	1/4	0
$\{F,I,B,N\}$	0	1/4	1/4	1/4	1/4	0
TDPM	5/6	5/6	5/6	3/4	3/4	0
DPI	5/24	5/24	5/24	3/16	3/16	0

TABELLE 34 Auflistung aller Gewinnkoalitionen und die Deegan-Packel-Indizes in der EWG

BEISPIEL 10.28 *Vergleich der Machtindizes anhand der EWG im Jahre 1958*

Auch wenn die einzelnen Indizes unterschiedliche Resultate liefern, so sind sie sich doch in einigen wesentlichen Punkten einig: Deutschland, Frankreich und Italien besitzen die größte Macht, gefolgt von Belgien und den Niederlanden. Zudem wird Luxemburg keinerlei Einfluss zugesprochen.

Dass Luxemburg von keinem Index Macht erhält, mag zwar anfangs verwunderlich erscheinen, ist aber plausibel, da Luxemburg in keiner einzigen Gewinnkoalition entscheidend war.

BEISPIEL 10.29 *Das Paradoxon der neuen Mitglieder*[61]

1973 traten England, Dänemark und Irland der Europäischen Wirtschaftsgemeinschaft bei. Für einen Beschluss waren nun mindestens 41 Stimmen notwendig. Das Stimmgewicht der einzelnen Staaten wurde wie folgt festgesetzt.

Stimmgewicht 10	Stimmgewicht 5	Stimmgewicht 3	Stimmgewicht 2
Deutschland	Belgien	Dänemark	Luxemburg
England	Niederlande	Irland	
Frankreich			
Italien			

TABELLE 35 Stimmgewicht der Mitgliedsstaaten der EWG im Jahr 1973

Luxemburg war der einzige Staat, dessen Stimmgewicht im Vergleich zu 1958 lediglich verdoppelt wurde. Alle anderen ursprünglichen Mitgliedsstaaten erhielten das 2,5fache ihres ursprünglichen Stimmgewichts. Trotz dieser offensichtlich schlechteren Behandlung erhielt Luxemburg trotzdem Macht, da es nun in etlichen Gewinnkoalitionen entscheidend war. Dieses Phänomen wird in der Literatur als Paradoxon der neuen Mitglieder bezeichnet.

10.6.2 ÖSTERREICHS EINFLUSS IM WELTSICHERHEITSRAT

Wir gehen von einem $(39; 7, 7, 7, 7, 7, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$ – Spiel aus (siehe Beispiel 10.3).

Österreich war bis 2009 dreimal Mitglied im Weltsicherheitsrat, das erste Mal 1973/74, danach 1991/1992 und zuletzt 2009/2010[62].

[61] englische Bezeichnung: Paradox of New Members
[62] http://de.wikipedia.org/wiki/Sicherheitsrat_der_Vereinten_Nationen [18.05.2007]

BEISPIEL 10.30 *Österreichs Macht anhand des Shapley-Index*

Damit Österreich bei einer Wahl im Weltsicherheitsrat entscheidend war, mussten neben den 5 ständigen Mitgliedern S_i noch genau 3 weitere aus den verbleibenden 9 nicht ständigen Mitgliedern N_j dieser Koalition angehören. Dafür gab es $\binom{9}{3}$ Möglichkeiten.

$$\underbrace{\overbrace{S_1, S_2, S_3, S_4, S_5, N_1, N_2, N_3}^{\text{Gewinnkoalition}}}_{\text{8!-Möglichkeiten der Anordnung}}, A \;\Big|\; \underbrace{N_4, N_5, N_6, N_7, N_8, N_9}_{\text{6!-Möglichkeiten der Anordnung}}$$

$\Rightarrow$ Anzahl der Permutationen, in denen Österreich entscheidend war:

$$\binom{9}{3} \cdot 8! \cdot 6! = \frac{9!}{3! \cdot \cancel{6!}} \cdot 8! \cdot \cancel{6!}$$

Anzahl aller Permutationen: 15!

$\Rightarrow$ Shapley-Index von Österreich:

$$SI(A) = \frac{9! \cdot 8!}{3! \cdot 15!} = \frac{4}{2145}$$

$\Rightarrow$ Shapley-Index eines ständigen Mitgliedes:

$$SI(A) = \frac{1}{5} \cdot \left(1 - 10 \cdot SI(A)\right) = \frac{421}{2145}$$

BEISPIEL 10.31 *Österreichs Macht anhand des Banzhaf-Index*

Die Gesamtzahl an Gewinnkoalitionen setzt sich aus allen fünf ständigen und mindestens vier nicht ständigen Mitgliedern zusammen. Da jedes ständige Mitglied in jeder Gewinnkoalition entscheidend ist, ist dies zugleich auch schon die totale Banzhaf-Macht eines ständiges Mitgliedes.

$$TBM(\text{eines ständigen Mitglieds}) = \binom{5}{5} \cdot \sum_{i=4}^{10} \binom{10}{i} = 848$$

Die Anzahl der möglichen Gewinnkoalition mit Österreich als entscheidendem Faktor ergibt die totale Banzhaf-Macht Österreichs.

$$TBM(\text{Österreich}) = \binom{5}{5} \cdot \binom{9}{3} \cdot 1 = 84$$

Mit diesem Wissen beträgt nun der Banzhaf-Index

$$BI(\text{eines ständigen Mitglieds}) = \frac{848}{5 \cdot 848 + 10 \cdot 84} = \frac{212}{1270}$$

$$BI(\text{Österreich}) = \frac{84}{5 \cdot 848 + 10 \cdot 84} = \frac{21}{1270}$$

BEISPIEL 10.32 *Österreichs Macht anhand des Johnston-Index*

Es gibt zwei unterschiedliche Arten von möglichen Gewinnkoalitionen, die in nachfolgender Tabelle aufgelistet sind.

mögliche Gewinnkoalitionen	Anzahl an entscheidenden Spielern	entscheidende Gruppierungen
$\{5S, 4N\}$	9	*S, N*
$\{5S, 5-10N\}$	5	*S*

TABELLE 36 Auflistung aller unterschiedlichen Gewinnkoalitionen im Weltsicherheitsrat

Somit gilt für die totale Johnston-Macht eines ständigen Mitgliedes

$$TJM(\text{eines ständigen Mitglieds}) =$$

$$= \frac{1}{9} \cdot \binom{5}{5} \cdot \binom{10}{4} + \frac{1}{5} \cdot \binom{5}{5} \cdot \sum_{i=5}^{10} \binom{10}{i} = \frac{2264}{15}$$

Um die totale Johnston-Macht Österreichs bestimmen zu können, müssen wir diejenigen Gewinnkoalitionen mit Österreich als entscheidendem Faktor betrachten. Dazu müssen neben Österreich noch genau 3 andere der 9 nicht ständigen Mitgliedern der Koalition angehören.

$$TJM(\text{Österreich}) = \frac{1}{9} \cdot \binom{5}{5} \cdot \binom{9}{3} \cdot 1 = \frac{28}{3}$$

Woraus sofort die Machtaufteilung aufgrund des Johnston-Index folgt

$$JI(\text{eines ständigen Mitglieds}) = \frac{2264/15}{5 \cdot 2264/15 + 10 \cdot 28/3} = \frac{566}{3180}$$

$$JI(\text{Österreich}) = \frac{28/3}{5 \cdot 2264/15 + 10 \cdot 28/3} = \frac{35}{3180}$$

BEISPIEL 10.33 *Österreichs Macht anhand des Deegan-Packel-Index*

Insgesamt gibt es

$$\underbrace{\binom{5}{5}}_{\substack{\text{ständige}\\ \text{Miglieder}}} \cdot \underbrace{\binom{10}{4}}_{\substack{\text{nicht ständige}\\ \text{Miglieder}}} = 210$$

Minimalgewinnkoalitionen mit jeweils 9 Mitgliedern.

Davon sind

$$\binom{5}{5} \cdot \binom{9}{3} \cdot 1 = 84$$

Koalitionen mit österreichischer Beteiligung.

Die TDPM-Werte sind somit

$$TDPM\left(\text{eines ständigen Mitglieds}\right)=\frac{1}{9}\cdot 210=\frac{70}{3}$$

$$TDPM\left(\text{Österreich}\right)=\frac{1}{9}\cdot 84=\frac{28}{3}$$

Folglich erhält man

$$DPI\left(\text{eines ständigen Mitglieds}\right)=\frac{70/3}{5\cdot 70/3+10\cdot 28/3}=\frac{5}{45}$$

$$DPI\left(\text{Österreich}\right)=\frac{28/3}{5\cdot 70/3+10\cdot 28/3}=\frac{2}{45}$$

BEISPIEL 10.34 *Vergleich der Machtindizes anhand der EWG im Jahre 1958*

Bei der Berechnung zeigt sich, dass der Deegan-Packel-Index einem ständigen Mitglied lediglich die 2,5 fache Macht von Österreich einräumt. Laut dem Banzhaf-Index wäre dieser Einfluss mehr als 10mal so groß, gemäß dem Johnston-Index mehr als 16mal so groß. Das geringste Mitspracherecht besitzt Österreich laut dem Shapley-Index. Hier wäre ein ständiges Mitglied mehr als 105mal mächtiger als Österreich.

10.6.3 DIE MACHT DES AMERIKANISCHEN PRÄSIDENTEN

Vereinfacht dargestellt gibt es 536 Stimmberechtigte, bestehend aus 435 Mitgliedern des Abgeordnetenhauses, 100 Senatsmitgliedern und dem Präsident. Der Präsident hat zwar Vetomacht, die jedoch von einer Zweidrittelmehrheit sowohl des Abgeordnetenhauses (=290 Abgeordnete) als auch des Senates (=67 Senatoren) außer Kraft gesetzt werden kann.

Damit ein Gesetz verabschiedet wird, muss also einer der beiden folgenden Fälle eintreten:

- ♠ Mindestens 218 Abgeordnete, mindestens 51 Senatoren und der Präsident stimmen für das Gesetz.
- ♠ Mindestens 290 Abgeordnete und mindestens 67 Senatoren stimmen für das Gesetz.

Genau genommen gibt es in Amerika 537 Stimmberechtigte, da in Wirklichkeit auch der Vizepräsident berücksichtigt werden müsste. Dieser kann nämlich mit seiner Stimme die Abstimmung entscheiden, wenn mindestens 218 Abgeordnete, lediglich 50 Senatoren und der Präsident gleiche Meinung haben. Um die Berechnung nicht noch komplizierter zu machen, werde ich von diesem Sonderfall jedoch absehen.

BEISPIEL 10.35 *Der Shapley-Index des Präsidenten*

Um den Shapley-Index des Präsidenten berechnen zu können, müssen wir folgende Fallunterscheidungen vornehmen:

1.Fall: 218-289 Abgeordnete und 51-66 Senatoren stimmen mit dem Präsidenten für den Antrag

$$\underbrace{\frac{1}{536!}}_{\substack{\text{Zahl aller}\\\text{möglichen}\\\text{Permutationen}}} \cdot \underbrace{\sum_{i=218}^{289}\sum_{j=51}^{66}\binom{435}{i}\cdot\binom{100}{j}}_{\substack{\text{Auswahl der betreffenden}\\\text{Abgeordneten und Senatoren}}} \cdot \underbrace{(i+j)!}_{\substack{\text{Permutation}\\\text{der Gesetzes-}\\\text{befürworter}}} \cdot \underbrace{(535-i-j)!}_{\substack{\text{Permutation der}\\\text{Gesetzesablehner}}}$$

2.Fall: Zwei Drittel der Senatoren sprechen sich mit 218-289 und dem Präsidenten für den Antrag aus

$$\frac{1}{536!}\cdot\sum_{i=218}^{289}\sum_{j=67}^{100}\binom{435}{i}\cdot\binom{100}{j}\cdot(i+j)!\cdot(535-i-j)!$$

3. Fall: Eine Zweidrittelmehrheit der Abgeordneten unterstützt mit 51-66 Senatoren den Präsidenten

$$\frac{1}{536!}\cdot\sum_{i=290}^{435}\sum_{j=51}^{66}\binom{435}{i}\cdot\binom{100}{j}\cdot(i+j)!\cdot(535-i-j)!$$

Somit erhalten wir folgenden Shapley-Index des Präsidenten (Fall 1 und 2 lassen sich zusammenfassen):

$$\begin{aligned}SI(P)=&\frac{1}{536!}\cdot\sum_{i=218}^{289}\sum_{j=51}^{100}\binom{435}{i}\cdot\binom{100}{j}\cdot(i+j)!\cdot(535-i-j)!+\\&+\frac{1}{536!}\cdot\sum_{i=290}^{435}\sum_{j=51}^{66}\binom{435}{i}\cdot\binom{100}{j}\cdot(i+j)!\cdot(535-i-j)!=\\&\cong 0{,}16047\end{aligned}$$

Mit einer analogen Vorgangsweise lässt sich auch der Shapley-Index des Abgeordnetenhauses und des Senates angeben.

Ein Abgeordneter verwandelt eine Koalition genau dann in eine Gewinnkoalition um, falls bereits entweder 217 andere Abgeordnete, 51-100 Senatoren und der Präsident vorhanden sind oder 289 andere Abgeordnete zusammen mit 67-100 Senatoren. Somit erhalten wir den folgenden Shapley-Index.

$$\begin{aligned}SI(A)=&\frac{1}{536!}\cdot\sum_{j=51}^{100}\binom{434}{217}\cdot\binom{100}{j}\cdot1\cdot(217+j+1)!\cdot(535-217-j-1)!+\\&+\frac{1}{536!}\cdot\sum_{j=67}^{100}\binom{434}{289}\cdot\binom{100}{j}\cdot(289+j)!\cdot(535-289-j)!=\\&\cong 0{,}00092\end{aligned}$$

Ebenso kann ein Senator eine Gewinnkoalition erzeugen. Es müssen also entweder 218-435 Abgeordnete, 50 andere Senatoren und der Präsident bereits für etwas stimmen oder 290-435 Abgeordnete und 66 andere Senatoren. Dies ergibt den folgenden Shapley-Index

$$\begin{aligned} SI(S) = \frac{1}{536!} \cdot \sum_{i=218}^{435} \binom{435}{i} \cdot \binom{99}{50} \cdot 1 \cdot (i+50+1)! \cdot (535-i-50-1)! + \\ + \frac{1}{536!} \cdot \sum_{i=290}^{435} \binom{435}{i} \cdot \binom{99}{66} \cdot (i+66)! \cdot (535-i-66)! = \\ \cong 0{,}00437 \end{aligned}$$

Somit ergibt sich folgende Machtverteilung:

Shapley-Macht des Präsidenten: 16,05 %
Shapley-Macht des Abgeordnetenhauses: 40,21 %
Shapley-Macht des Senats: 43,74 %

BEISPIEL 10.36 *Der Banzhaf-Index des Präsidenten*

Am einfachsten lässt sich die Gesamtzahl an Gewinnkoalitionen bestimmen, wo der Präsident beteiligt ist.

$$\#K(P) = \sum_{i=218}^{435} \binom{435}{i} \cdot \sum_{j=51}^{100} \binom{100}{j} \cdot 1$$

Die Gesamtzahl an Gewinnkoalitionen setzt sich aus zwei Gruppen zusammen, die Koalitionen mit Beteiligung des Präsidenten und die ohne.

$$\#K(\text{gesamt}) = \underbrace{\sum_{i=218}^{435} \binom{435}{i} \cdot \sum_{j=51}^{100} \binom{100}{j} \cdot 1}_{\text{Koalitionen mit Präsident}} + \underbrace{\sum_{i=290}^{435} \binom{435}{i} \cdot \sum_{j=67}^{100} \binom{100}{j}}_{\text{Koalitionen ohne Präsident}}$$

Um die Gesamtzahl an Gewinnkoalitionen zu bestimmen, zu denen ein bestimmter Abgeordneter gehört, teilen wir abermals diese Koalitionen in die Koalitionen mit dem Präsidenten und die ohne auf. Dabei ist zu beachten, dass jeweils nur noch 434 andere Abgeordnete benötigt werden.

$$\#K(A) = \sum_{i=217}^{434} \binom{434}{i} \cdot 1 \cdot \sum_{j=51}^{100} \binom{100}{j} \cdot 1 + \sum_{i=289}^{434} \binom{434}{i} \cdot 1 \cdot \sum_{j=67}^{100} \binom{100}{j}$$

Analog erhält man die Koalitionen mit einem bestimmten Senator.

$$\#K(S) = \sum_{i=218}^{435} \binom{435}{i} \cdot \sum_{j=50}^{99} \binom{99}{j} \cdot 1 \cdot 1 + \sum_{i=290}^{435} \binom{435}{i} \cdot \sum_{j=66}^{100} \binom{99}{j} \cdot 1$$

Aufgrund von Proposition 10.15 gilt

$$TBM(\text{Präsident}) = 2 \cdot \#K(P) - \#K(\text{gesamt})$$

$$TBM(1\ \text{Abgeordneter}) = 2 \cdot \#K(A) - \#K(\text{gesamt})$$

$$TBM(1\ \text{Senator}) = 2 \cdot \#K(S) - \#K(\text{gesamt})$$

Dies führt unmittelbar zu den entsprechenden Banzhaf-Indizes

$$BI(\text{Präsident}) = \frac{TBM(P)}{TBM(P)+435\cdot TBM(A)+100\cdot TBM(S)} \cong 0{,}0380$$

$$BI(1\ \text{Abgeordneter}) =$$

$$= \frac{TBM(A)}{TBM(P)+435\cdot TBM(A)+100\cdot TBM(S)} \cong 0{,}00146$$

$$BI(1\ \text{Senator}) = \frac{TBM(S)}{TBM(P)+435\cdot TBM(A)+100\cdot TBM(S)} \cong 0{,}00329$$

Die Macht teilt sich also laut dem Banzhaf-Index wie folgt auf:

Banzhaf-Macht des Präsidenten: 3,80 %
Banzhaf-Macht des Abgeordnetenhauses: 63,32 %
Banzhaf-Macht des Senats: 32,88 %

BEISPIEL 10.37 *Der Johnston-Index des Präsidenten*

Zunächst widmen wir uns den unterschiedlichen Arten von möglichen Gewinnkoalitionen und überlegen uns wie viele Spieler jeweils entscheidend sind.

mögliche Gewinnkoalitionen	Anzahl an entscheidenden Spielern	entscheidende Gruppierungen
$\{218A, 51S, P\}$	270	*A, S, P*
$\{219-435A, 51S, P\}$	52	*S, P*
$\{218A, 52-100S, P\}$	219	*A, P*
$\{219-289A, 52-100S, P\}$	1	*P*
$\{219-435A, 52-66S, P\}$	1	*P*
$\{290A, 67S\}$	357	*A, S*
$\{291-435A, 67S\}$	67	*S*
$\{290A, 68-100S\}$	290	*A*

TABELLE 37 Auflistung aller unterschiedlichen Gewinnkoalitionen im amerikanischen Rechtssystem

Mit dieser Information lässt sich sogleich die totale Johnston-Macht des Präsidenten ermitteln.

$$TJM(\text{Präsident}) =$$

$$= \frac{1}{270}\cdot\binom{435}{218}\cdot\binom{100}{51}+\frac{1}{52}\cdot\left[\sum_{i=219}^{435}\binom{435}{i}\right]\cdot\binom{100}{51}+$$

$$+\frac{1}{219}\cdot\binom{435}{218}\cdot\left[\sum_{j=52}^{100}\binom{100}{j}\right]+\left[\sum_{i=219}^{289}\binom{435}{i}\right]\cdot\left[\sum_{j=52}^{100}\binom{100}{j}\right]+$$

$$+\left[\sum_{i=219}^{435}\binom{435}{i}\right]\cdot\left[\sum_{j=52}^{66}\binom{100}{j}\right]$$

Durch eine ähnliche Vorgangsweise erhält man die totale Johnston-Macht eines Abgeordneten sowie eines Senators. Allerdings muss bei den entsprechenden Binomialkoeffizienten dieses Mal berücksichtigt werden, dass der betrachtete Abgeordnete bzw. Senator ausgenommen wird um sicherzugehen, dass der betreffende Abgeordnete bzw. Senator an den jeweiligen Gewinnkoalitionen auch tatsächlich beteiligt ist.

$$\begin{aligned}
&TJM(1\text{ Abgeordneter}) = \\
&= \frac{1}{270}\cdot\binom{434}{217}\cdot 1\cdot\binom{100}{51} + \frac{1}{219}\cdot\binom{434}{217}\cdot 1\cdot\left[\sum_{j=52}^{100}\binom{100}{j}\right] + \\
&= \frac{1}{357}\cdot\binom{434}{289}\cdot 1\cdot\binom{100}{51} + \frac{1}{290}\cdot\binom{434}{289}\cdot 1\cdot\left[\sum_{j=68}^{100}\binom{100}{j}\right]
\end{aligned}$$

$$\begin{aligned}
&TJM(1\text{ Senator}) = \\
&= \frac{1}{270}\cdot\binom{435}{218}\cdot\binom{99}{50}\cdot 1 + \frac{1}{52}\cdot\left[\sum_{i=219}^{435}\binom{435}{i}\right]\cdot\binom{99}{50}\cdot 1 + \\
&+ \frac{1}{357}\cdot\binom{435}{290}\cdot\binom{99}{66}\cdot 1 + \frac{1}{67}\cdot\left[\sum_{i=291}^{435}\binom{435}{i}\right]\cdot\binom{99}{66}\cdot 1
\end{aligned}$$

Wir erhalten also die folgenden Johnston-Indizes

$$JI(\text{Präsident}) = \frac{TJM(P)}{TJM(P) + 435\cdot TJM(A) + 100\cdot TJM(S)} \cong 0{,}869981$$

$$\begin{aligned}
&JI(1\text{ Abgeordneter}) = \\
&= \frac{TJM(A)}{TJM(P) + 435\cdot TJM(A) + 100\cdot TJM(S)} \cong 0{,}000096
\end{aligned}$$

$$JI(1\text{ Senator}) = \frac{TJM(S)}{TJM(P) + 435\cdot TJM(A) + 100\cdot TJM(S)} \cong 0{,}000883$$

Die Machtaufteilung aufgrund des Johnston-Indizes lautet somit:

Johnston-Macht des Präsidenten: 87 %
Johnston-Macht des Abgeordnetenhauses: 4,17 %
Johnston-Macht des Senats: 8,83 %

BEISPIEL 10.38 *Der Deegan-Packel-Index des Präsidenten*

Für die Anzahl der minimalen Gewinnkoalitionen inklusive der Stimme des Präsidenten gilt

$$MGK(\text{mit Präsident}) = \underbrace{\binom{435}{218}}_{\text{Abgeordnete}}\cdot\underbrace{\binom{100}{51}}_{\text{Senat}}\cdot\underbrace{1}_{\text{Präsident}}$$

All diese minimalen Gewinnkoalitionen setzen sich jeweils aus

$$218 + 51 + 1 = 270$$

entscheidenden Spielern zusammen.

Analog erhält man die Anzahl der minimalen Gewinnkoalitionen, an denen der Präsident nicht beteiligt ist.

$$MGK\left(\text{ohne Präsident}\right)=\underbrace{\binom{435}{290}}_{\text{Abgeordnete}}\cdot\underbrace{\binom{100}{67}}_{\text{Senat}}$$

Jede dieser minimalen Gewinnkoalitionen besteht aus

$$290+67=357$$

entscheidenden Spielern.

Somit erhält man die folgenden TDPM-Werte

$$TDPM\left(\text{Präsident}\right)=\frac{1}{270}\cdot\binom{435}{218}\cdot\binom{100}{51}$$

$$TDPM\left(1\text{ Abgeordneter}\right)=\frac{1}{270}\cdot\binom{434}{217}\cdot\binom{100}{51}+\frac{1}{357}\cdot\binom{434}{289}\cdot\binom{100}{51}$$

$$TDPM\left(1\text{ Senator}\right)=\frac{1}{270}\cdot\binom{435}{218}\cdot\binom{99}{50}+\frac{1}{357}\cdot\binom{435}{290}\cdot\binom{99}{50}$$

Mit deren Hilfe lassen sich nun die entsprechenden Deegan-Packel-Indizes berechnen.

$$DPI\left(\text{Präsident}\right)=$$

$$=\frac{TDPM(P)}{TDPM(P)+435\cdot TDPM(A)+100\cdot TDPM(S)}\cong 0{,}00370$$

$$DPI\left(1\text{ Abgeordneter}\right)=$$

$$=\frac{TDPM(A)}{TDPM(P)+435\cdot TDPM(A)+100\cdot TDPM(S)}\cong 0{,}00186$$

$$DPI\left(1\text{ Senator}\right)=$$

$$=\frac{TDPM(S)}{TDPM(P)+435\cdot TDPM(A)+100\cdot TDPM(S)}\cong 0{,}00189$$

Somit ergibt sich die folgende Machtaufteilung:

Deegan-Packel-Macht des Präsidenten: 0,37 %
Deegan-Packel-Macht des Abgeordnetenhauses: 80,74%
Deegan-Packel-Macht des Senats: 18,89 %

Beispiel 10.39 *Vergleich der Machtindizes anhand der Macht des amerikanischen Präsidenten*

Je nachdem welchen Machtindex man der Berechnung zugrunde legt, erhält man grundlegend andere Resultate. Als mächtigste Institution stuft der Shapley-Index den Senat ein, der Banzhaf-Index und der Deegan-Packel-Index das Abgeordnetenhaus und der Johnston-Index den Präsidenten. Die Macht des Präsidenten schwankt dabei zwischen 0,37% (Deegan-Packel-Index) und 87% (Johnston-Index).

Verwendete Literatur zu den Machtindizes:

- ♠ BRAMS, Steven, Game Theory and Politics, New York 1975
- ♠ FELSENTHAL, DAN S. / MACHOVER, Moshé, The Measuremnet of Voting Power. Theory and Practice, Problems and Paradoxes, Cornwall 1998
- ♠ HESSE Christian, Angewandte Wahrscheinlichkeitstheorie, Braunschweig / Wiesbaden 2003
- ♠ HESSE, Christian / MEISTER, Alexander, Übungsbuch zur angewandten Wahrscheinlichkeitstheorie. Aufgaben und Lösungen, Wiesbaden 2005
- ♠ HOLLER, Manfred / ILLING, Gerhard, Einführung in die Spieltheorie, Berlin ^{6}2006
- ♠ LUCAS, William, Gewichtete Wahlsysteme: Wie man Macht messen kann, In: Garfunkel, Salomon / Steen, Lynn (Hg.), Mathematik in der Praxis. Anwendungen in Wirtschaft, Wissenschaft und Politik, Heidelberg, 1989, S.138-152
- ♠ TAYLOR, Alan, Mathematics and Politics. Strategy, Voting, Power and Proof, New York 1995

TEIL 3

WAHLPROGNOSEN UND UMFRAGEN

11. MATHEMATISCHE MODELLE UND WIRKLICHKEIT

Wahlprognosen haftet seit jeher ein recht negatives Image an. Auch Politiker streuen den Meinungsforschern nicht gerade Rosen, vor allem dann, wenn sie eine Wahl verloren haben und die Schuld bei falschen Prognosen suchen.

Der Grund für das schlechte Bild von Prognosen in der Öffentlichkeit liegt darin, dass aus Umfragen gewonnene Taten nicht unbedingt mit den tatsächlichen Wahlergebnissen übereinstimmen müssen, obwohl das mathematische Modell absolut fehlerlos ist.

Ich möchte nun die vier wesentlichsten Gründe angeben, warum Wahlprognosen mit der Wirklichkeit in Widerspruch stehen können und warum sie sich irren können.

Bei jeder Umfrage gibt es Schwankungsbreiten, da Wahrscheinlichkeiten berechnet werden. Eine exakte Vorhersage ist daher von vornherein ein Ding der Unmöglichkeit, sofern nicht alle Personen befragt werden, was sich jedoch nicht realisieren lässt.

Ein erhebliches Problem stellt die Gruppe der unentschlossenen Wähler und Wechselwähler dar, die von Jahr zu Jahr größer wird und die es immer schwieriger macht, den Wahlausgang richtig vorhersagen zu können.

Große Schwierigkeiten bereitet auch die Tatsache, dass viele befragte Personen falsche Angaben machen bzw. dann nicht einmal zur Wahl gehen.

Umfragen kurz vor einer Wahl wären zwar genauer und würden eine bessere Vorhersage liefern, bergen aber die Gefahr, dass dadurch der Ausgang der Wahl beeinflusst werden könnte. Daher werden die letzten Umfragen meistens nicht mehr veröffentlicht.

11.1 SCHWANKUNGSBREITEN

Viele Menschen halten Wahlprognosen für Unfug, weil diese teils erheblich vom tatsächlichen Wahlergebnis abweichen können. In Wirklichkeit wäre es jedoch Scharlatanerie, wenn man das Ergebnis punktgenau vorhersagen würde.

Überall, wo Wahrscheinlichkeiten eine Rolle spielen, gibt es auch einen gewissen Unsicherheitsfaktor. Eine exakte Prognose ist daher von vornherein ausgeschlossen. Auch mit einer Brückenwaage kann man gewogene Mengen nicht auf Gramm genau angeben. Ähnlich verhält es sich bei Wahlprognosen.

Reichweite in %

	50	45	40	35	30	25	20	15	10	8	5	2
Fälle	50	55	60	65	70	75	80	85	90	92	95	98
100	9.8	9.8	9.5	9.3	9.0	8.5	7.8	7.0	5.9	5.3	4.3	2.7
200	6.9	6.9	6.8	6.6	6.4	6.0	5.5	4.9	4.2	3.8	3.0	1.9
300	5.7	5.6	5.5	5.4	5.2	4.9	4.5	4.0	3.4	3.1	2.5	1.6
400	4.9	4.9	4.8	4.7	4.5	4.2	3.9	3.5	2.9	2.7	2.1	1.4
500	4.4	4.4	4.3	4.2	4.0	3.8	3.5	3.1	2.6	2.4	1.9	1.2
600	4.0	4.0	3.9	3.8	3.7	3.5	3.2	2.9	2.4	2.2	1.7	1.1
700	3.7	3.7	3.6	3.5	3.4	3.2	3.0	2.6	2.2	2.0	1.6	1.0
800	3.5	3.4	3.4	3.3	3.2	3.0	2.8	2.5	2.1	1.9	1.5	1.0
900	3.3	3.3	3.2	3.1	3.0	2.8	2.6	2.3	2.0	1.8	1.4	0.9
1000	3.1	3.1	3.0	3.0	2.8	2.7	2.5	2.2	1.9	1.7	1.4	0.9
2000	2.2	2.2	2.1	2.1	2.0	1.9	1.8	1.6	1.3	1.2	1.0	0.6
5000	1.4	1.4	1.4	1.3	1.3	1.2	1.1	1.0	0.8	0.8	0.6	0.4
10000	1.0	1.0	1.0	0.9	0.9	0.8	0.8	0.7	0.6	0.5	0.4	0.3

TABELLE 38 Die theoretischen Schwankungsbreiten (Signifikanzniveau = 95%)[63]

Tabelle 38 gibt einen Überblick über die theoretischen Schwankungsbreiten. Werden etwa einer Partei 25 % vorausgesagt, so bedeutet eine Schwankungsbreite von 3%, dass der tatsächliche Stimmanteil dieser Partei zwischen 22 und 28 % liegt.

Auf das mathematische Modell, auf dem diese Zahlen beruhen, sowie auf den Begriff Signifikanzniveau möchte ich dann später in Kapitel 13 (Beispiel 13.9) näher eingehen.

Kopf-an-Kopf-Duelle lassen sich umso schwieriger vorhersagen, je näher die Kandidaten / die Parteien bei 50% liegen. Die Prognosen sind umso genauer, je näher die Werte bei 0 und bei 100% liegen. Deshalb lassen sich noch am Besten kleinere Parteien vorhersagen, da hier die Schwankungsbreite kleiner ist und somit eine genauere Prognose möglich ist.

Bei 500 Personen gibt es eine maximale Schwankungsbreite von 4,4%. Möchte man diese Zahl halbieren, würde man bereits 2000 Personen

[63] ARBEITSGEMEINSCHAFT MEDIA-ANALYSEN, Media-Analyse 2005 Jahresbericht, Wien 2005, S.33

benötigen. Bei 10000 Befragungen wäre zwar die Schwankungsbreite maximal 1%, aber der dabei entstandene Aufwand (Befragungen, Auswertungen und Kosten) stünde in keinster Relation zu der erhaltenen Genauigkeit.

In der Praxis werden im Schnitt etwa 500-1000 Personen befragt, wen sie bei der nächsten Wahl wählen werden. Dabei müsste man eine maximale Schwankungsbreite zwischen 3,1 und 4,4% in Kauf nehmen. Bei 800 befragten Personen hätte man etwa eine Schwankungsbreite von bis zu 3,5%.

11.2 Wechselwähler und neue Parteien

Früher gab es weniger Parteien als heute. Außerdem war die Zahl der Stammwähler wesentlich größer als heutzutage. Da es kaum Wechselwähler gab, war eine richtige Einschätzung leichter möglich, da man sich bei einer Wahlprognose am Ergebnis früherer Wahlen orientieren konnte.

Bei den Wechselwählern lassen sich zwei Gruppen unterscheiden. Zunächst gibt es diejenigen Wähler, die keine Stammwähler sind. Andererseits ändert sich aber auch auf natürlichem Weg ständig die Wählerschaft, da alte Wähler sterben und dafür neue dazukommen.

Bei den Nationalratswahlen 2006 hatte die SPÖ die größte Zahl der Stammwähler. Insgesamt 85 % der SPÖ-Wähler von 2002 gaben erneut den Sozialdemokraten ihre Stimme. Im Vergleich dazu verlor die ÖVP knapp ein Viertel Ihrer Wähler von 2002. Den geringsten Prozentsatz der Stammwähler hatte 2006 die FPÖ.

in Prozent der Wähler 2002 haben 2006 gewählt	ÖVP	SPÖ	FPÖ	Grüne	BZÖ	HPM
der ÖVP-Wähler 2002	77	7	7	3	2	4
der SPÖ-Wähler 2002	2	85	4	4	0	2
der FPÖ-Wähler 2002	16	8	49	0	24	2
der Grünen-Wähler 2002	9	21	0	63	0	3

Tabelle 39 Wählerbewegung 2006[64]

Tabelle 39 gibt Aufschluss darüber, wie viel Prozent eine Partei 2006 gegenüber der letzten Wahl 2002 verloren hat und zu welcher Partei die Wähler gewechselt sind.

Die große Zahl an unentschlossenen Wählern ist einer der Hauptgründe für ungenaue Prognosen, da diese Wählergruppe unberechenbar ist, zumal sie von Jahr zu Jahr größer wird.

Jahr	erst während des Intensivwahlkampfes	schon länger vorher
1979	9	91
1983	8	92
1986	16	84
1990	14	86
1994	18	82
1995	21	79
1999	20	80
2002	23	77
2006	24	76

Tabelle 40 Zeitpunkt der Wahlentscheidung bei Nationalratswahlen 1979-2006 (Angaben in Prozent)[65]

[64] PLASSER, Fritz / ULRAM, Peter, Die Wahlanalyse 2006. Wer hat wen warum gewählt?, Presseunterlage Fessel-GfK, Wien 2006, S.8

[65] PLASSER, Fritz / ULRAM, Peter, Die Wahlanalyse 2006, S.4

Interessant ist der zu beobachtende Zusammenhang zwischen der Zunahme der Spätentscheider und der Zuname der Wechselwähler. Beide Gruppen lagen vor 1980 noch unter der 10% Grenze, machten aber bei der letzten Nationalratswahl 2006 bereits rund ein Viertel aller Stimmen aus.

Jahr	Wechselwähler	Spätentscheider
1979	7	9
1983	10	8
1986	16	16
1990	17	14
1994	19	18
1995	22	21
1999	18	20
2002	24	23
2006	26	24

TABELLE 41 Wählermobilität bei Nationalratswahlen 1979-2006 (Angaben in Prozent)[66]

Die Meinungsforscher versuchen das Problem der Wechselwähler durch einen Vergleich mit alten Wahlen in den Griff zu bekommen. Eine Analyse der unentschlossenen Wähler und wem diese dann schlussendlich bei den vorhergegangenen Wahlen ihre Stimmen gegeben haben, erlaubt Rückschlüsse auf die aktuelle Situation.

In Prozent der Wechselwähler haben gewählt	ÖVP	SPÖ	FPÖ	Grüne	BZÖ	HPM
1986	24	10	39	22	---	---
1990	11	15	51	9	---	---
1994	10	9	40	15	---	---
1995	21	25	34	5	---	---
1999	16	15	37	17	---	---
2002	48	30	6	13	---	---
2006	15	24	17	11	16	11

TABELLE 42 Wahlverhalten der Wechselwähler 1986-2006
(Differenz auf 100% = Sonstige Parteien bzw. Rundungsfehler)[67]

Voraussetzung für einen Vergleich ist jedoch, dass alle Parteien auch schon bei der letzten Wahl vertreten waren, womit wir auch schon beim nächsten Problem wären.

Parteien, die bei einer Wahl erstmals antreten (z.B. Liste Martin, erstmaliges Antreten bei der Europawahl 2004[68]), mehrere Parteien, die sich von einer einzigen Partei abgespaltet haben (z.B. Abspaltung und Gründung des Liberalen Forums am 4.Februar 1993[69]) oder eine Partei, die in mehrere kleinere Parteien zerfallen ist (z.B. FPÖ teilt sich in BZÖ und FPÖ auf durch

[66] PLASSER, Fritz / ULRAM, Peter, Die Wahlanalyse 2006, S.6
[67] PLASSER, Fritz / ULRAM, Peter, Die Wahlanalyse 2006, S.7
[68] http://de.wikipedia.org/wiki/Liste_Martin [05.02.2009]
[69] http://de.wikipedia.org/wiki/Liberales_Forum [05.02.2009]

die Gründung des BZÖ am 4.April 2005[70]), bereiten den Meinungsforschern erhebliche Schwierigkeiten.

Grund dafür ist, dass keinerlei Erfahrungswerte von früheren Wahlen vorhanden sind. Somit kann man sich nicht auf frühere Umfragen und Wahlausgänge bei der Erstellung einer Wahlprognose stützen.

Man kann keinen Vergleich mit früheren Daten anstellen, was eine Prognose schwierig macht. Somit kommt eine Komponente hinzu, die unberechenbar ist und keine genauen Vorhersagen möglich macht. Daher muss man sich auf empirische Befragungen verlassen, wobei der Wahrheitsgehalt dabei stets zu hinterfragen ist, da viele Personen bewusst falsche Angaben machen, was sogleich zum nächsten Punkt führt.

[70] http://de.wikipedia.org/wiki/BZÖ [05.02.2009]

11.3 FALSCHE ANGABEN DER BEFRAGTEN PERSONEN

Eine Prognose ist immer nur dann etwas wert, wenn die befragten Personen die Wahrheit angeben. Das muss aber gerade bei Wahlbefragungen keineswegs sein. Deshalb ist es notwendig, die erhaltenen Daten auch kritisch zu hinterfragen und ihnen nicht blind zu vertrauen.

Nicht zu unterschätzen ist die Anzahl der falschen Antworten, die man bei einer Wahl-Umfrage erhält. Oft prägt einem sein soziales Umfeld und erzeugt einen gewissen Druck. Aus Angst vor einer Ächtung in der Gesellschaft, sinkt die Bereitschaft vor allem in kleineren Dorfgemeinden sich für eine andere Partei zu deklarieren als die Mehrheit.

Erfahrungsgemäß schneiden die Grünen bei Umfragen stets besser ab als bei den Wahlen selbst. Das liegt darin begründet, da es als positiv angesehen wird, wenn man für Umwelt und gegen Ausländerfeindlichkeit eintritt, zentrale Themen der Grünen.

Dagegen werden rechtsgerichtete Parteien bei Umfragen oft nicht angegeben, da es den Menschen peinlich ist zuzugeben, dass sie selbst eine rechte Gesinnung besitzen.

Um die Dunkelziffer erfassen zu können, muss man bei Befragungen einen einfachen Trick anwenden. Statt der Sonntagsfrage „Wenn würden Sie wählen, wenn kommenden Sonntag Wahlen wären“ kann die Formulierung „Kennen Sie jemanden aus Ihrer Umgebung, der FPÖ/BZÖ wählt?“ Aufschluss über den tatsächlichen Anteil der Wähler dieser Partei geben, da sich Personen bei letztgenannter Fragestellung auch selbst angeben können, ohne Gefahr zu laufen sich selbst outen zu müssen.

Eine weitere Möglichkeit um wahrheitsgemäße Antworten zu erhalten liefert die Statistik mit einem speziellen Stichprobenverfahren, das als Randomized Response bezeichnet wird. Darauf werde ich dann später genauer in Kapitel 14 eingehen.

Bei Befragungen muss auch berücksichtigt werden, dass ein gewisser Prozentteil der befragten Personen schlussendlich dann nicht einmal zur Wahl geht, obwohl sie das bei der Umfrage angegeben haben. Somit können die Umfrageergebnisse damit ebenfalls verfälscht werden.

Es kann wahlentscheidend sein, wenn falsche Schlussfolgerungen aus veröffentlichten Umfragen gezogen werden. Durch veröffentlichte Umfragewerte liegt die Partei in Führung, der man seine Stimme geben würde.

Daher kann es vorkommen, dass jemand aus Bequemheit und im Vertrauen auf diese Umfragen nicht zur Wahl geht, da er seine Partei ohnehin schon als den vermeintlichen Wahlsieger ansieht und es daher nicht für notwendig erachtet, auch seine Stimme abzugeben.

Wenn viele Wähler aus eben genannter Überlegung nicht zur Wahl gehen, kann dadurch dann natürlich das Wahlergebnis beeinträchtigt werden und die Wahlprognosen sich als falsch erweisen.

Man spricht von dem sogenannten Bandwagon-Effekt, wenn man mit dem vermeintlichen Wahlsieger mitläuft. Im Gegensatz dazu steht der Underdog-Effekt, der auftritt, wenn man sich aus Frust oder aus Protest an eine Minderheit anschließt.[71]

[71] http://de.wikipedia.org/wiki/Wahlforschung [05.02.2009]

11.4 UMFRAGEN KURZ VOR EINER WAHL

Immer mehr Wähler, vor allem Wechselwähler, entscheiden sich erst im Laufe des Wahlkampfes, ein hoher Prozentsatz sogar erst am Wahltag selbst. Die Umfragen werden umso präziser, je näher die Wahl rückt, da sich der Anteil der unentschlossenen Wähler von Tag zu Tag verringert.

In Prozent haben ihre definitive Wahl-entscheidung getroffen –	wenige Tage vor der Wahl	1-2 Wochen vor der Wahl	stand schon länger fest
WählerInnen insgesamt	15	9	76
Stammwähler	7	5	89
Wechselwähler	35	24	40

TABELLE 43 Zeitpunkt der Wahlentscheidung nach Wählerschaften 2006[72]

Meinungsforschungsinstitute machen auch noch kurz vor Wahlen Umfragen. Jedoch dürfen solche Umfragen etwa in Portugal, Spanien, Frankreich und Ungarn nicht mehr veröffentlicht werden um die Wähler nicht zu beeinflussen und somit den Wahlausgang zu manipulieren und zu verfälschen. In Österreich gibt es nur die gesetzliche Verordnung, dass Exit Polls (Befragungen am Wahltag) und Wahlhochrechnungen erst nach dem Schließen der Wahllokale veröffentlicht werden dürfen.[73]

Trotzdem ist es auch in Österreich üblich, Umfragen kurz vor einer Wahl nicht mehr zu publizieren. Allerdings handelt es sich dabei nur um ein freiwilliges Agreement.

Eine Woche vor einer Wahl kann man eine Wahl nicht mehr gewinnen sondern nur noch verlieren. Politische Bomben, die eine Woche vor der Wahl hochgehen, können eine Wahl entscheiden, aber nicht mehr von (veröffentlichten) Umfragen erfasst werden.

[72] PLASSER, Fritz / ULRAM, Peter, Die Wahlanalyse 2006, S.5
[73] http://de.wikipedia.org/wiki/Wahlforschung [05.02.2009]

12. Die Gesetze der großen Zahlen

In diesem Kapitel möchte ich nun die Gesetze der großen Zahlen behandeln, die Aussagen über die relative Häufigkeit ermöglichen. Ziel dieses Abschnittes ist es, einen Zusammenhang zwischen der relativen Häufigkeit und der Wahrscheinlichkeit eines bestimmten Ereignisses herzustellen. Ich werde das schwache und das starke sowie das Bernoullische Gesetz der großen Zahlen vorstellen und diese mit Ausnahme des starken Gesetzes auch beweisen.

Beim schwachen Gesetz der großen Zahlen konvergiert die Wahrscheinlichkeit gegen null, dass das arithmetische Mittel um mehr als ε vom Erwartungswert μ abweicht. Dieses Gesetz besagt also, dass sich die Folge des arithmetischen Mittels $\bar{X}$ immer mehr dem Erwartungswert μ nähert, je größer n ist. Daher kann das arithmetische Mittel als guter Schätzwert für den unbekannten Erwartungswert aufgefasst werden.

Mit Hilfe des Bernoullischen Gesetzes der großen Zahlen werde ich Wahrscheinlichkeiten abschätzen und Angaben über den notwendigen Stichprobenumfang machen.

Anschaulich besagt das starke Gesetz der großen Zahlen, dass die relative Häufigkeit mit Wahrscheinlichkeit 1 ein Schätzer für die unbekannte Wahrscheinlichkeit p ist, wenn ein Ereignis A hinreichend oft wiederholt wird.

Mit diesen Gesetzen schließt sich der Kreis zu Kapitel 1, wo ich die Wahrscheinlichkeit intuitiv als Grenzwert der relativen Häufigkeiten eingeführt habe. In Definition 1.7 habe ich Mises' Begriff der naiven Wahrscheinlichkeit als Grenzwert der relativen Häufigkeiten vorgestellt. Mises hatte damit zwar schon die richtige Idee gehabt, jedoch verwendete er die normale Konvergenz. Ersetzt man diese durch die stochastische Konvergenz (Definition 12.5), so erhält man die mathematisch korrekte Formulierung.

12.1 DAS SCHWACHE GESETZ DER GROßEN ZAHLEN

SATZ 12.1 *Das schwache Gesetz der großen Zahlen*

Seien $X_1, X_2, ..., X_n$ unabhängige Zufallsvariable mit dem Erwartungswert $E(X_i) = \mu$ und der Varianz $V(X_i) = \sigma^2$. Dann gilt $\forall \varepsilon > 0$

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n \cdot \varepsilon^2}$$

bzw.

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right) = 0$$

BEWEIS:

Sei

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

das arithmetische Mittel aller Zufallsvariablen $X_1, X_2, ..., X_n$.

Für den Erwartungswert und die Varianz von $\overline{X}$ gilt

$$E\left(\overline{X}\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \underset{5.4}{\overset{\text{Satz}}{=}} \frac{1}{n} \cdot \sum_{i=1}^{n} E(X_i) =$$

$$= \frac{1}{n} \cdot [\underbrace{E(X_1)}_{=\mu} + \underbrace{E(X_2)}_{=\mu} + ... + \underbrace{E(X_n)}_{=\mu}] = \frac{1}{\cancel{n}} \cdot \cancel{n} \cdot \mu = \mu$$

$$V\left(\overline{X}\right) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \underset{5.11}{\overset{\text{Satz}}{=}} \frac{1}{n^2} \cdot \sum_{i=1}^{n} V(X_i) =$$

$$= \frac{1}{n^2} \cdot [\underbrace{V(X_1)}_{=\sigma^2} + \underbrace{V(X_2)}_{=\sigma^2} + ... + \underbrace{V(X_n)}_{=\sigma^2}] = \frac{1}{n^{\cancel{2}}} \cdot \cancel{n} \cdot \sigma^2 = \frac{\sigma^2}{n}$$

Für große *n* strebt dabei die Varianz immer mehr gegen Null

$$\lim_{n\to\infty} V\left(\overline{X}\right) = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0$$

Aus der Ungleichung von Tschebyschew (Satz 5.19) folgt nun

$$P\left(\left|\overline{X} - E\left(\overline{X}\right)\right| \geq \varepsilon\right) \leq \frac{V\left(\overline{X}\right)}{\varepsilon^2}$$

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n \cdot \varepsilon^2}$$

$$\Rightarrow \lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right) = 0$$

□

12.2 Das Bernoullische Gesetz der großen Zahlen

Satz 12.2 *Das Bernoullische Gesetz der großen Zahlen*

Das Ereignis A mit der Wahrscheinlichkeit p werde n-mal unabhängig voneinander wiederholt. Dann gilt für die Zufallsvariable $R_n(A)$ der relativen Häufigkeiten von A

$$E\left(R_n(A)\right) = p, V\left(R_n(A)\right) = \frac{p\cdot(1-p)}{n} \le \frac{1}{4n}$$

$$P\left(\left|R_n(A)-p\right| \ge \varepsilon\right) \le \frac{p\cdot(1-p)}{n\cdot\varepsilon^2} \le \frac{1}{4n\cdot\varepsilon^2}$$

$$\lim_{n\to\infty} P\left(\left|R_n(A)-p\right| \ge \varepsilon\right) = 0 \ \forall\varepsilon > 0$$

Beweis:

Sei $X = n\cdot R_n(A)$ die Zufallsvariable der absoluten Häufigkeiten mit

$$X = \sum_{i=1}^{n} X_i,\ X_i = \begin{cases} 1 \text{ falls } A \text{ beim i-ten Versuch eintritt} \\ 0 \text{ sonst} \end{cases}$$

Die X_i sind somit unabhängige Bernoulli-verteilte Zufallsvariable mit den Kenngrößen

$$E(X_i) = p, V(X_i) = p\cdot(1-p)$$

Folglich besitzt die Zufallsvariable $R_n(A) = \frac{1}{n}\sum_{i=1}^{n} X_i$ den Erwartungswert

$$E\left(R_n(A)\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \overset{\text{Satz}}{\underset{5.4}{=}} \frac{1}{n}\cdot\sum_{i=1}^{n} E(X_i) =$$

$$= \frac{1}{n}\cdot[\underbrace{E(X_1)}_{=p} + \underbrace{E(X_2)}_{=p} + \ldots + \underbrace{E(X_n)}_{=p}] = \frac{1}{\cancel{n}}\cdot\cancel{n}\cdot p = p$$

und die Varianz

$$V\left(R_n(A)\right) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \overset{\text{Satz}}{\underset{5.11}{=}} \frac{1}{n^2}\cdot\sum_{i=1}^{n} V(X_i) =$$

$$= \frac{1}{n^2}\cdot[\underbrace{V(X_1)}_{=p\cdot(1-p)} + \underbrace{V(X_2)}_{=p\cdot(1-p)} + \ldots + \underbrace{V(X_n)}_{=p\cdot(1-p)}] = \frac{1}{n^{\cancel{2}}}\cdot\cancel{n}\cdot p\cdot(1-p) =$$

$$= \frac{p\cdot(1-p)}{n}$$

Betrachten wir nun die Funktion

$$f:[0,1]\to\mathbb{R}, f(p)=\frac{p\cdot(1-p)}{n},$$

die wir mithilfe des maximalen Werts nach oben abschätzen können

$$f(p)=\frac{p\cdot(1-p)}{n}=\frac{1}{n}\cdot\left(p-p^2\right)$$

$$f'(p)=\frac{1}{n}\cdot(1-2p)=0$$

$$p=\frac{1}{2}$$

Vergleichen wir nun diesen Wert mit den Randwerten, so ergibt sich

$$f(0)=f(1)=0$$

$$f\left(\frac{1}{2}\right)=\frac{1}{4n}$$

Somit nimmt das globale Maximum an der Stelle $p=\frac{1}{2}$ den Wert $\frac{1}{4n}$ an.

$$\Rightarrow V\left(R_n(A)\right)=\frac{p\cdot(1-p)}{n}\leq\frac{1}{4n}$$

Der letzte Teil des Bernoullischen Gesetzes der großen Zahlen folgt durch Anwendung des schwachen Gesetzes der großen Zahlen auf die Zufallsvariable $R_n(A)$.

□

BEISPIEL 12.3 *Schätzen des Stichprobenumfangs*

Aufgrund der Befragung von n Personen kennen wir die relative Häufigkeit der Befürworter einer Partei. Wir interessieren und nun für den minimalen Stichprobenumfang n, wenn die relative Häufigkeit mit einer Wahrscheinlichkeit von höchstens 3% von der unbekannten Wahrscheinlichkeit um 2% abweichen darf.

Wir schätzen den Stichprobenumfang mithilfe des Bernoullischen Gesetz der großen Zahlen ab

$$P\left(\left|R_n(A)-p\right|\geq\varepsilon\right)\leq\frac{1}{4n\cdot\varepsilon^2}$$

$$P\left(\left|R_n(A)-p\right|\geq 0{,}02\right)\leq\frac{1}{4\cdot n\cdot 0{,}02^2}\leq 0{,}03$$

$$\Rightarrow n\geq\frac{1}{4\cdot 0{,}02^2\cdot 0{,}03}\cong\underline{\underline{20\,834}}$$

Man müsste also rund 21000 Personen befragen, wenn die relative Häufigkeit mit einer Wahrscheinlichkeit von höchstens 3% von der unbekannten Wahrscheinlichkeit um 2% abweichen darf.

BEISPIEL 12.4 *Schätzen des Stichprobenumfangs bei bekanntem p*

Im Vergleich zu Beispiel 12.3 sei nun etwa $p \leq 0{,}08$ bekannt. Dies sei die Wahrscheinlichkeit dafür, dass ein Wähler eine kleine Partei wählt.

Aufgrund des schwachen Gesetzes der großen Zahlen gilt

$$P\left(\left|R_n(A)-p\right| \geq 0{,}02\right) \leq \frac{p\cdot(1-p)}{n\cdot 0{,}02^2} \leq \max_{0\leq p\leq 0{,}08}\left\{\frac{p\cdot(1-p)}{n\cdot 0{,}02^2}\right\} \leq 0{,}03$$

Wegen der Monotonie nimmt die Funktion $f(p)=p\cdot(1-p)$ im Intervall $[0;0{,}08]$ ihr Maximum in 0,08 an.

Wir erhalten die folgende Abschätzung für den Stichprobenumfang

$$\frac{0{,}08\cdot 0{,}92}{n\cdot 0{,}02^2} \leq 0{,}03$$

$$\Rightarrow n \geq \frac{0{,}08\cdot 0{,}92}{0{,}02^2\cdot 0{,}03} \cong 6134$$

Somit kann mit der Zusatzinformation $p \leq 0{,}08$ der benötigte Stichprobenumfang auf knapp ein Drittel verringert werden. Den größten Stichprobenumfang erhält man, wenn für *p* auch Werte in der Nähe von 50% möglich sind.

12.3 Das starke Gesetz der großen Zahlen

Definition 12.5 *Stochastische und fast sichere Konvergenz*

Sei eine Folge $X_1, X_2, X_3, \ldots$ von Zufallsvariablen und η eine beliebige Zahl.

Die Folge X_n konvergiert stochastisch gegen η, falls

$$\lim_{n\to\infty} P\left(\left|X_n - \eta\right| < \varepsilon\right) = 1 \,\forall \varepsilon > 0$$

Die Folge X_n konvergiert fast sicher gegen η, falls

$$P\left(\lim_{n\to\infty} X_n = \eta\right) = 1$$

Satz 12.6 *Das starke Gesetz der großen Zahlen*

Seien $X_1, X_2, \ldots, X_n$ unabhängige Zufallsvariable mit dem Erwartungswert $E(X) = \mu$. Dann konvergiert die Folge $\frac{1}{n}\sum_{i=1}^{n} X_i$ fast sicher gegen μ, d.h.

$$P\left(\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = \mu\right) = 1$$

Eine allgemeinere Version dieses Satzes stammt von Kolmogoroff. Dessen Beweis findet sich etwa bei Albrecht Irle, Wahrscheinlichkeitstheorie und Statistik, S.181ff.

Die Begriffe „schwach“ bzw. „stark“ beziehen sich auf die Art der Konvergenz. Im ersten Fall ist die stochastische Konvergenz (= Konvergenz in Wahrscheinlichkeiten) gemeint, im zweiten die fast sichere Konvergenz.

Die fast sichere Konvergenz impliziert die stochastische Konvergenz. Die Umkehrung gilt jedoch nicht. Deshalb ist die stochastische Konvergenz der schwächere Konvergenzbegriff.

13. MEINUNGSUMFRAGEN

In diesem Kapitel gehe ich der Frage nach, wie man Rückschlüsse aus einer Wahlumfrage auf die Gesamtheit aller Wähler ziehen kann. Grundlagen für dieses Kapitel sind die Binomialverteilung (Kapitel 7) und die Normalverteilung (Kapitel 6).

Im Normalfall werden bei einer Meinungsumfrage 500-1000 Personen befragt, welche Partei sie wählen würden. Wichtige Fragestellungen in diesem Zusammenhang sind der ungefähre Wähleranteil einer Partei in der Gesamtbevölkerung bzw. ob sich der Wähleranteil seit der letzten Umfrage vergrößert oder verringert hat.

Die Qualität und die Exaktheit einer Umfrage hängen sehr von der verwendeten Stichprobe ab. Deshalb werde ich zunächst einmal den Begriff einer Stichprobe und die Anforderungen erläutern, die an eine Stichprobe gestellt werden.

Als nächstes werde ich erklären, wie man einen unbekannten Anteil schätzen kann. Dazu gibt es zwei verschiedene Methoden, auf die ich näher eingehen möchte, nämlich die Punktschätzung und die Intervallschätzung.

Bei einer Punktschätzung spielt der sogenannte Maximum-Likelihood-Schätzer eine wesentliche Rolle, mit dem sich der unbekannte Wähleranteil bestimmen lässt. Zwar kann der tatsächliche Wert getroffen werden, aber vermutlich wird das nur sehr selten der Fall sein. Vielmehr wird der genaue Anteil im Normalfall nicht mit dem geschätzten Wert übereinstimmen.

Die Punktschätzung liefert somit keinerlei zusätzliche Information über die mögliche Abweichung des geschätzten Wertes vom tatsächlichen Anteil. Diesem Problem wird bei der so genannten Intervallschätzung Rechnung getragen. Hierbei wird der Begriff eines Konfidenzintervalls auftreten.

Zu guter Letzt werde ich auf Hypothesentests eingehen. Hypothesen (altgriechisch υποθεσις = Annahme) sind Rückschlüsse auf die Wählergesamtheit, die man aufgrund der erhaltenen Daten der Stichproben macht. Diese können richtig oder auch falsch sein. Mit Hilfe von Hypothesentests können Aussagen über die Gültigkeit bzw. Ungültigkeit der aufgestellten Hypothesen getätigt werden.

13.1 STICHPROBEN

Politische Meinungsumfragen versuchen stets den Prozentsatz der Bevölkerung mit einer bestimmten Eigenschaft in Erfahrung zu bringen, etwa wie viel Prozent einen Politiker sympathisch finden oder wie viel Prozent der Wahlberechtigten eine bestimmte Partei wählen würden.

Die Statistik liefert nun Modelle, die es ermöglichen aufgrund der Befragung eines kleinen respektiven Personenkreises Rückschlüsse auf die gesamte Bevölkerung zu ziehen. Die befragten Personen bezeichnet man als Stichprobe. Dabei können die Aussagen nur mit einer bestimmten Wahrscheinlichkeit getätigt werden. Eine absolute Sicherheit liegt nicht im Bereich des Machbaren. Dazu müsste schließlich die gesamte Bevölkerung befragt werden, was nicht nur aus Kostengründen ein Ding der Unmöglichkeit ist.

Die Stichprobe ist unabhängig von der Größe der Grundgesamtheit, sie liefert also dieselbe Genauigkeit, egal ob es sich um eine Umfrage in Österreich handelt oder in Deutschland, wo etwa zehnmal so viele wahlberechtigte Personen leben.

Der Begriff „Stichprobe“ ist bereits sehr alt und geht auf die mittelalterlichen Kaufleute zurück. Um die Qualität der Handelswaren feststellen und um über den Preis verhandeln zu können, wurde ein zufälliger Teil der Säcke angestochen und die Ware überprüft.[74]

Eine Stichprobe muss ein repräsentativer Querschnitt der Gesamtbevölkerung sein. Nur wenn sie ein exaktes verkleinertes Abbild der Wirklichkeit ist, ist gewährleistet, dass sie auch richtige Aufschlüsse über die Gesamtheit liefern kann. Dazu gehört das genaue Verhältnis von weiblichen und männlichen Befragten, von verschiedenen Alters- und Berufsgruppen sowie die Berücksichtigung des unterschiedlichen Bildungsniveaus. So wird etwa eine Stichprobe, die nur aus Akademikern besteht, in keinster Weise den tatsächlichen Prozentsatz einer Arbeiterpartei wiedergeben können.

Individualdaten sind die Grundlage für Meinungsumfragen. Dabei erhält man mit Hilfe von einzelnen Personen die gewünschten Informationen. Aggregatdaten dagegen werden bei Wahlhochrechnungen verwendet. Mit Hilfe einer kleinen Anzahl von bereits ausgezählten Stimmen wird versucht das Endergebnis zu prognostizieren.

[74] Roth, Dieter, Empirische Wahlforschung. Ursprung, Theorien, Instrumente und Methoden, Opladen 1998, S.57f.

13.2 ANTEILSSCHÄTZUNG

13.2.1 PUNKTSCHÄTZUNG UND DER MAXIMUM-LIKELIHOOD-SCHÄTZER

Eine Punktschätzung gilt als gut, wenn die beiden folgenden Kriterien zutreffen:

- ♠ Der geschätzte Wert soll möglichst nahe beim tatsächlichen Wert liegen.
- ♠ Wird der Stichprobenumfang größer, so wird auch der geschätzte Wert genauer.

KOROLLAR 13.1 *Maximum-Likelihood-Schätzer*

Um den unbekannten Anteil der Wähler einer Partei p in der Bevölkerung zu bestimmen, werden n Personen befragt. Wenn k Personen angeben, diese Partei zu wählen, dann ist der so genannte Maximum-Likelihood-Schätzer durch

$$\hat{p} = \frac{k}{n}$$

gegeben.

BEWEIS:

Jede befragte Person gebe ihre Antwort unabhängig von der Antwort anderer Personen. Sei X die Anzahl der Personen unter n befragten Personen, die angeben, eine bestimmte Partei zu wählen, und sei $\hat{p}$ der geschätzte Wähleranteil dieser Partei. Dann ist $X \sim B(n, \hat{p})$. Somit gilt

$$L(\hat{p}) = \binom{n}{k} \cdot \hat{p}^k \cdot (1-\hat{p})^{n-k}$$

Logarithmiert man diese Gleichung, so verändert sich das Maximum aufgrund der Monotonie des Logarithmus nicht und man erhält

$$\log L(\hat{p}) = \log\left[\binom{n}{k} \cdot \hat{p}^k \cdot (1-\hat{p})^{n-k}\right]$$

$$\log L(\hat{p}) = \log\binom{n}{k} + k \cdot \log\hat{p} + (n-k) \cdot \log(1-\hat{p})$$

Will man diesen Ausdruck maximieren, so muss man die erste Ableitung nullsetzen:

$$\frac{dL(\hat{p})}{d\hat{p}} = \frac{k}{\hat{p}} - \frac{n-k}{1-\hat{p}} = 0$$

$$k \cdot (1-\hat{p}) = \hat{p} \cdot (n-k)$$

$$k \cancel{- k \cdot \hat{p}} = n \cdot \hat{p} \cancel{- k \cdot \hat{p}} \Rightarrow \underline{\hat{p} = \frac{k}{n}}$$

Da

$$L(0) = L(1) = 0, L(\hat{p}) > 0 \;\forall \hat{p} \in \left]0, 1\right[$$

ist $\hat{p} = \frac{k}{n}$ ein globales Maximum und somit der gesuchte Maximum-Likelihood-Schätzer.

□

BEISPIEL 13.2 *Anteilsschätzung der Wähler einer bestimmten Partei*

Bei einer Stichprobe von 1000 Wahlberechtigten gaben 337 Personen an, dass sie eine bestimmte Partei wählen würden. Mit Hilfe des Maximum-Likelihood-Schätzers schätze man den Anteil der Wähler dieser Partei unter allen Wahlberechtigten.

$$\hat{p} = \frac{337}{1000} = 0{,}337$$

Der geschätzte Anteil der Wähler dieser Partei unter allen Wahlberechtigten beträgt somit 33,7%.

13.2.2 INTERVALLSCHÄTZUNG

Eine Intervallschätzung sollte die beiden folgenden Eigenschaften erfüllen:

♠ Der geschätzte Wert soll im Intervall liegen.
♠ Je sicherer eine Schätzung sein soll, desto größer muss das Intervall gewählt werden.

Wir suchen einen Bereich, in dem der genaue Wert p mit einer bestimmten Wahrscheinlichkeit γ liegt. Diesen Bereich nennt man Konfidenzintervall.

DEFINITION 13.3 *Konfidenzintervalle*

Seien

$$\Theta_u = g_u(X_1, X_2, ..., X_n) \text{ und } \Theta_o = g_o(X_1, X_2, ..., X_n)$$

zwei Stichprobenfunktionen mit den Wahrscheinlichkeiten

$$P(\Theta_u \le \vartheta \le \Theta_o) = \gamma = 1 - \alpha \text{ bzw.}$$

$$P(\vartheta \le \Theta_o) = \gamma = 1 - \alpha; P(\vartheta \ge \Theta_u) = \gamma = 1 - \alpha$$

Dann heißen die Zufallsintervalle

$$[\Theta_u; \Theta_o] \text{ zweiseitige bzw.}$$

$$]-\infty; \Theta_o] \text{ und } [\Theta_u; \infty[\text{ einseitige (links- und rechtsseitige)}$$

Konfidenzintervalle (Vertrauensintervalle) für ϑ zum Konfidenzniveau (Signifikanzniveau) $\gamma = 1 - \alpha$.

Die Zahl γ bezeichnet man als Konfidenzniveau (Signifikanzniveau, Sicherheitswahrscheinlichkeit), die Zahl α als Irrtumswahrscheinlichkeit. Die Summe aus γ und α ist stets 1.

Ein Konfidenzniveau von 0,95 besagt anschaulich, dass in 95% der Fälle der wahre Anteil p in dem betreffenden Konfidenzintervall liegt und in 5% der Fälle nicht. Analog irrt man sich bei einer Irrtumswahrscheinlichkeit von 1% lediglich nur mehr in einem von hundert Fällen.

Offensichtlich wird das Konfidenzintervall umso größer, je größer das Konfidenzniveau und desto kleiner die Irrtumswahrscheinlichkeit ist.

Üblicherweise verwendet man ein 95%- bzw. ein 99%-Konfidenzintervall. Der Grund dafür ist, dass ein 95%-Konfidenzintervall nahezu dem symmetrischen Intervall $[\mu - 2\sigma;\ \mu + 2\sigma]$ einer Normalverteilung entspricht, während das 99%-Konfidenzintervall eine grobe Analogie zu dem $[\mu - 3\sigma;\ \mu + 3\sigma]$ – Intervall einer Normalverteilung ist.

BEISPIEL 13.4 *Fortsetzung von Beispiel 13.2*

Wir suchen nun ein 95%-Konfidenzintervall mit der unteren Grenze p_1 und der oberen Grenze p_2 für unseren vorher berechneten Wähleranteil von 33,7%. Das gesuchte Intervall soll also mit 95%iger Wahrscheinlichkeit den tatsächlichen Wähleranteil p beinhalten, wobei die 5% derart aufzuteilen sind, dass $p < p_1$ und $p > p_2$ jeweils mit einer Wahrscheinlichkeit von lediglich 2,5% eintreten dürfen.

Zunächst gehen wir der Frage nach, wie weit wir den Wähleranteil verkleinern müssen, damit wir mit 97,5 %iger Wahrscheinlichkeit höchstens 337 Wähler von 1000 befragten Personen antreffen, sprich es muss

$$P(X \le 337) = \sum_{i=0}^{337} \binom{1000}{i} \cdot p^i \cdot (1-p)^{1000-i} \ge 0{,}975 \text{ und}$$

$$P(X \le 336) = \sum_{i=0}^{336} \binom{1000}{i} \cdot p^i \cdot (1-p)^{1000-i} < 0{,}975$$

gelten.

Die untere Grenze p_1, erhalten wir somit mit Hilfe eines Computers aus folgender Gleichung

$$\sum_{i=0}^{336} \binom{1000}{i} \cdot p_1^i \cdot (1-p_1)^{1000-i} = 0{,}975$$

$$\Rightarrow \underline{\underline{p_1 \cong 0{,}308}}$$

Analog bestimmen wir nun die obere Grenze p_2. Dazu wollen wir den Fehleranteil so weit vergrößern, dass die Wahrscheinlichkeit, mindestens 337 Wähler von 1000 befragten Personen anzutreffen, bei 97,5% liegt.

$$P(X \ge 337) \ge 0{,}975$$

$$1 - P(X \le 336) \ge 0{,}975$$

$$P(X \le 336) \le 0{,}025 \text{ und } P(X \le 337) > 0{,}025$$

Die obere Grenze p_2 ist somit Lösung von

$$\sum_{i=0}^{337} \binom{1000}{i} \cdot p_2^i \cdot (1-p_2)^{1000-i} = 0{,}025$$

$$\Rightarrow \underline{\underline{p_2 \cong 0{,}367}}$$

Somit liegt der tatsächliche Wähleranteil mit 95%iger Sicherheit zwischen 30,8 und 36,7%.

BEISPIEL 13.5 *Fortsetzung von Beispiel 13.2*

Wir interessieren uns nun für das 99%-Konfidenzintervall. Analog zu Beispiel 13.4 ergeben sich nun die beiden Grenzen p_1 und p_2 wie folgt

$$\sum_{i=0}^{336} \binom{1000}{i} \cdot p_1^i \cdot (1-p_1)^{1000-i} = 0{,}995 \Rightarrow \underline{\underline{p_1 \cong 0{,}298}}$$

$$\sum_{i=0}^{337} \binom{1000}{i} \cdot p_2^i \cdot (1-p_2)^{1000-i} = 0{,}005 \Rightarrow \underline{\underline{p_2 \cong 0{,}377}}$$

Der tatsächliche Wähleranteil befindet sich also mit 99%iger Sicherheit im Bereich $[0{,}298; 0{,}377]$.

Interessiert man sich für die Anzahl der Wähler einer Partei, so verwendet man ein zweiseitiges Konfidenzintervall. Meistens betrachtet man in diesem Fall symmetrische Intervalle um den geschätzten Wert, aber auch oft asymmetrische Intervalle.

Ein einseitiges Konfidenzintervall wird dagegen etwa bei einer Prognose verwendet, ob eine kleine Partei den Einzug in den Nationalrat schafft und die Vier-Prozent-Hürde überwinden kann oder wenn man sich interessiert, ob eine Partei Stimmen verloren oder dazugewonnen hat.

13.2.3 NORMALVERTEILTE KONFIDENZINTERVALLE FÜR P

Die obigen Berechnungen sind sehr mühsam und trotz Verwendung eines Computers sehr zeitaufwendig. Eine einfachere Methode der Auswertung liefert die Normalverteilung, die bekanntlich (Kapitel 7.3, speziell Bemerkung 7.15) als Approximation der Binomialverteilung verwendet werden kann, wenn

$$\sigma^2 = np(1-p) > 9$$

BEISPIEL 13.6 *Benötigter Stichprobenumfang*

Angenommen, eine Partei habe bei der letzten Wahl 2% der Stimmen erhalten. So muss die Stichprobe aus 460 Befragungen bestehen, damit die Berechnung mit Hilfe der Normalverteilung erfolgen darf. Dies erklärt auch die Tatsache, dass für Meinungsumfragen im Normalfall mindestens 500 Personen befragt werden.

$$n > \frac{9}{p(1-p)} = \frac{9}{0{,}02 \cdot 0{,}98} \cong 459{,}18$$

Der kleinste Stichprobenumfang ist notwendig, wenn eine Partei 50% der Stimmen auf sich vereinigen kann. In diesem Fall benötigt man lediglich 36 Personen um die Normalverteilung anwenden zu können.

$$n > \frac{9}{p(1-p)} = \frac{9}{0{,}50 \cdot 0{,}50} = 36$$

SATZ 13.7 *Asymptotische Konfidenzintervalle für p zum Niveau* $1-\alpha$ *bei großem Stichprobenumfang n mit* $np(1-p) > 9$

Sei $\hat{p}$ der Maximum-Likelihood-Schätzer eines Ereignisses A in einer unabhängigen Versuchsserie mit dem Stichprobenumfang n. Betrachtet man die $N(0;1)$ – Verteilung mit $\gamma = 2\cdot\Phi(z)-1$, so erhält man für die Wahrscheinlichkeit p ein asymptotisches zweiseitiges Konfidenzintervall $[p_1; p_2]$ mit den Grenzen

$$p_{1,2} = \frac{n}{n+z^2}\cdot\left(\hat{p} + \frac{z^2}{2n} \pm z\cdot\sqrt{\frac{\hat{p}\cdot(1-\hat{p})}{n} + \frac{z^2}{4n^2}}\right).$$

Durch Weglassen einer Grenze erhält man einseitige Konfidenzintervalle für p zum Konfidenzniveau $1-\frac{\alpha}{2}$.

BEWEIS:

Sei $X = H_n(A)$ die binomialverteilte Zufallsvariable der absoluten Häufigkeit des Ereignissen A in einer unabhängigen Versuchsserie vom Umfang n mit den Parametern n und p. Ist n genügend groß, dann ist X näherungsweise normalverteilt mit $E(X) = np$ und $V(X) = np(1-p)$ (siehe Kapitel 7.3). Somit gilt für die Standardisierung

$$\frac{X-np}{\sqrt{np(1-p)}} \sim N(0;1).$$

Für $z > 0$ erhalten wir

$$2\Phi(z)-1 \approx P\left(\left|\frac{X-np}{\sqrt{np(1-p)}}\right| \le z\right) = P\left(\frac{(X-np)^2}{np(1-p)} \le z^2\right) =$$

$$= P\left((X-np)^2 \le np(1-p)\cdot z^2\right)$$

$$(X-np_{1,2})^2 \le np_{1,2}(1-p_{1,2})z^2 \qquad \left|-np_{1,2}z^2 + np_{1,2}^2 z^2\right.$$

$$X^2 - 2np_{1,2}X + n^2p_{1,2}^2 - np_{1,2}z^2 + np_{1,2}^2z^2 \le 0 \qquad \left|-X^2\right.$$

$$np_{1,2}^2(n+z^2) - np_{1,2}(2X+z^2) \le -X^2 \qquad \left|:n, :(n+z^2)\right.$$

$$p_{1,2}^2 - p_{1,2}\cdot\frac{2X+z^2}{n+z^2} \le -\frac{X^2}{n(n+z^2)} \qquad \left|+\left(\frac{2X+z^2}{2(n+z^2)}\right)^2\right.$$

$$p_{1,2}^2 - p_{1,2} \cdot \frac{2X+z^2}{n+z^2} + \left(\frac{2X+z^2}{2\left(n+z^2\right)}\right)^2 \le -\frac{X^2}{n\left(n+z^2\right)} + \left(\frac{2X+z^2}{2\left(n+z^2\right)}\right)^2$$

$$\left(p_{1,2} - \frac{2X+z^2}{2\left(n+z^2\right)}\right)^2 \le -\frac{4X^2\left(n+z^2\right)}{4n\left(n+z^2\right)^2} + \frac{n\left(4X^2+4Xz^2+z^4\right)}{4n\left(n+z^2\right)^2} =$$

$$= \frac{\cancel{-4nX^2} - 4X^2z^2 \cancel{+4nX^2} + 4nXz^2 + nz^4}{4n\left(n+z^2\right)^2} =$$

$$= \frac{1}{\left(n+z^2\right)^2} \cdot z^2 \cdot \left[\frac{-4X^2+4nX+nz^2}{4n}\right] =$$

$$= \frac{1}{\left(n+z^2\right)^2} \cdot z^2 \cdot \left[\frac{X\left(n-X\right)}{4n} + \frac{z^2}{4}\right] \quad \left|\sqrt{}\right.$$

$$\left|p_{1,2} - \frac{2X+z^2}{2\left(n+z^2\right)}\right| \le \frac{1}{n+z^2} \cdot z \cdot \sqrt{\frac{X\left(n-X\right)}{n} + \frac{z^2}{4}}$$

$$\left|p_{1,2} - \frac{n}{n+z^2} \cdot \left(\frac{X}{n} + \frac{z^2}{2n}\right)\right| \le \frac{1}{n+z^2} \cdot z \cdot \sqrt{n^2 \cdot \left(\frac{1}{n} \cdot \frac{X}{n} \cdot \frac{n-X}{n} + \frac{z^2}{4n^2}\right)}$$

Setzt man nun

$$\frac{X}{n} = \hat{p}$$

so erhält man das gewünschte Resultat

$$\left|p_{1,2} - \frac{n}{n+z^2} \cdot \left(\hat{p} + \frac{z^2}{2n}\right)\right| \le \frac{n}{n+z^2} \cdot z \cdot \sqrt{\frac{\hat{p}\left(1-\hat{p}\right)}{n} + \frac{z^2}{4n^2}}$$

$$\cancel{\frac{n}{n+z^2}} \cdot \left|p_{1,2} \cdot \frac{n+z^2}{n} - \left(\hat{p} + \frac{z^2}{2n}\right)\right| \le \cancel{\frac{n}{n+z^2}} \cdot z \cdot \sqrt{\frac{\hat{p}\left(1-\hat{p}\right)}{n} + \frac{z^2}{4n^2}}$$

$$-z \cdot \sqrt{\frac{\hat{p}\left(1-\hat{p}\right)}{n} + \frac{z^2}{4n^2}} \le p_{1,2} \cdot \frac{n+z^2}{n} - \left(\hat{p} + \frac{z^2}{2n}\right) \le z \cdot \sqrt{\frac{\hat{p}\left(1-\hat{p}\right)}{n} + \frac{z^2}{4n^2}}$$

Als Lösung der quadratischen Gleichung erhält man:

$$p_{1,2} \cdot \frac{n+z^2}{n} = \hat{p} + \frac{z^2}{2n} \pm z \cdot \sqrt{\frac{\hat{p}\left(1-\hat{p}\right)}{n} + \frac{z^2}{4n^2}}$$

$$p_{1,2} = \frac{n}{n+z^2} \cdot \left(\hat{p} + \frac{z^2}{2n} \pm z \cdot \sqrt{\frac{\hat{p}\left(1-\hat{p}\right)}{n} + \frac{z^2}{4n^2}}\right)$$

$\square$

KOROLLAR 13.8 *Vereinfachung des Konfidenzintervalls von Satz 13.7*

Für große Stichprobenumfänge *n* gilt

$$p_{1,2} \approx \hat{p} \pm z \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

BEWEIS:

$$\lim_{n\to\infty} \underbrace{\frac{n}{n+z^2}}_{\to} \cdot \left(\hat{p} + \underbrace{\frac{z^2}{2n}}_{\to 0} \pm z \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \underbrace{\frac{z^2}{4n^2}}_{\to 0}} \right) = \hat{p} \pm z \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\Rightarrow p_{1,2} \approx \hat{p} \pm z \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

□

BEISPIEL 13.9 *Vergleich mit Beispiel 13.4 und 13.5*

Wir berechnen nun die Konfidenzintervalle mit Hilfe von Korollar 13.8

Für das 95%-Konfidenzintervall gilt

$$\gamma = 2 \cdot \Phi(z) - 1 \Rightarrow 0{,}95 = 2 \cdot \Phi(z) - 1 \Rightarrow \Phi(z) = 0{,}975 \Rightarrow z = 1{,}96$$

$$p_{1,2} \approx \hat{p} \pm z \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0{,}337 \pm 1{,}96 \cdot \sqrt{\frac{0{,}337 \cdot 0{,}663}{1000}}$$

$$\Rightarrow \underline{\underline{[30{,}8\%; 36{,}7\%]}}$$

Und ebenso erhalten wir das 99%-Konfidenzintervall:

$$\gamma = 2 \cdot \Phi(z) - 1 \Rightarrow 0{,}99 = 2 \cdot \Phi(z) - 1 \Rightarrow \Phi(z) = 0{,}995 \Rightarrow z = 2{,}575$$

$$p_{1,2} \approx \hat{p} \pm z \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0{,}337 \pm 2{,}575 \cdot \sqrt{\frac{0{,}337 \cdot 0{,}663}{1000}}$$

$$\Rightarrow \underline{\underline{[29{,}9\%; 37{,}6\%]}}$$

Diese Resultate sind denen von früher sehr ähnlich, nur mit dem Unterschied, dass der Rechenaufwand dieses Mal erheblich geringer war und bereits mit gewöhnlichen Taschenrechnern problemlos zu bewältigen ist.

BEISPIEL 13.10 *maximale Schwankungsbreiten bei einem Signifikanzniveau von 95%*

Die maximale Schwankungsbreite befindet sich bei $\hat{p} = \frac{1}{2}$, da

$$f(\hat{p}) = \frac{\hat{p}(1-\hat{p})}{n} = \frac{1}{n} \cdot (\hat{p} - \hat{p}^2)$$

$$f'(\hat{p}) = \frac{1}{n} \cdot (1 - 2\hat{p}) = 0 \Rightarrow \hat{p} = \frac{1}{2}$$

Die maximale Schwankungsbreite bei 100 befragten Personen beträgt

$$\pm 1{,}96 \cdot \sqrt{\frac{0{,}5 \cdot 0{,}5}{100}} = \pm 0{,}098 \mathrel{\hat{=}} \pm 9{,}8\,\%.$$

Analog ergeben sich die maximalen Schwankungsbreiten bei 500

$$\pm 1{,}96 \cdot \sqrt{\frac{0{,}5 \cdot 0{,}5}{500}} \cong \pm 0{,}044 \mathrel{\hat{=}} \pm 4{,}4\,\%$$

bzw. bei 1000 befragten Personen

$$\pm 1{,}96 \cdot \sqrt{\frac{0{,}5 \cdot 0{,}5}{1000}} \cong \pm 0{,}031 \mathrel{\hat{=}} \pm 3{,}1\,\%.$$

BEISPIEL 13.11 *Stichprobenumfang bei einem „Kopf-an-Kopf“-Rennen*

Besonders schwierig gestalten sich „Kopf-an-Kopf“-Duelle für Meinungsforscher, etwa bei der Wahl des Bundespräsidenten. Wie viele Wähler müssen mindestens befragt werden um mit 95%iger Sicherheit das Wahlergebnis richtig vorhersagen zu können, wenn der Unterschied zwischen den beiden Kandidaten mindestens 1 Prozent der Wählerstimmen betragen soll?

Zunächst setzen wir

$$\varepsilon = z \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Da der Unterschied mindestens 1 Prozent betragen soll, suchen wir ein Konfidenzintervall um p

$$[\hat{p} - \varepsilon; \hat{p} + \varepsilon]$$

mit der Länge 0,01.

Somit erhalten wir für unser ε

$$2\varepsilon = 0{,}01 \Rightarrow \varepsilon = 0{,}005$$

Ein „Kopf-an-Kopf“-Rennen bedeutet einen Maximum-Likelihood-Schätzer von

$$\hat{p} = 0{,}5$$

Das 95%-Konfidenzintervall liefert

$z = 1{,}96$ (siehe Beispiel 13.9)

Der benötigte Stichprobenumfang lässt sich nun problemlos berechnen

$$\varepsilon = z \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$0{,}005 = 1{,}96 \cdot \sqrt{\frac{0{,}5 \cdot 0{,}5}{n}} \quad \Big|^2$$

$$0{,}005^2 = 1{,}96^2 \cdot \frac{0{,}25}{n} \Rightarrow n = \frac{1{,}96^2 \cdot 0{,}25}{0{,}005^2} = \underline{\underline{38\,416}}$$

Es müssten also rund 40000 Wähler befragt werden, eine Zahl die in keinster Weise realisierbar ist. Deshalb müssen Abstriche gemacht werden, entweder bei der Sicherheit oder bei dem Unterschied zwischen den Wählerstimmen.

BEISPIEL 13.12 *Ermittlung des Stichprobenumfangs*

Angenommen, eine Partei habe bei der letzten Wahl 23,5% der gültigen Stimmen erhalten. Wie viele Personen müssen befragt werden, wenn ein Marktforschungsinstitut den aktuellen Wähleranteil auf dem 95%-Konfidenzintervall auf $\pm 2\%$ genau schätzen lassen soll?

Wir kennen die folgenden Werte

$$z = 1{,}96$$

$$\hat{p} = 0{,}235$$

$$\varepsilon = 0{,}02$$

In die schon bekannte Formel eingesetzt, ergibt sich der benötigte Stichprobenumfang

$$\varepsilon = z \cdot \sqrt{\frac{\hat{p}\left(1-\hat{p}\right)}{n}}$$

$$0{,}02 = 1{,}96 \cdot \sqrt{\frac{0{,}235 \cdot 0{,}765}{n}} \quad \Big|^2$$

$$0{,}02^2 = 1{,}96^2 \cdot \frac{0{,}235 \cdot 0{,}765}{n} \Rightarrow n = \frac{1{,}96^2 \cdot 0{,}235 \cdot 0{,}765}{0{,}02^2} \cong \underline{\underline{1726{,}56}}$$

Es müssten also 1727 Personen befragt werden um mit 95%iger Sicherheit den aktuellen Wähleranteil auf 2 Prozent genau schätzen zu können.

13.3 HYPOTHESENTESTS

Mit Hilfe von Stichproben versucht man Rückschlüsse auf die Grundgesamtheit zu ziehen, die sowohl richtig als auch falsch sein können. Derartige Annahmen nennt man Hypothesen. Mit Hilfe eines statistischen Tests wird festgestellt, ob die Informationen, die man aus einer konkreten Stichprobe erhalten hat, die aufgestellte Hypothese befürworten oder ob sie dazu im Widerspruch stehen.

Man unterscheidet dabei zwischen Nullhypothesen H_0 und Alternativhypothesen H_1. Nullhypothesen sind die Ausgangshypothesen, mit Alternativhypothesen stellt man dagegen eine andere Hypothese im Vergleich zu H_0 auf.

Hierbei können zwei mögliche Fehler auftreten. Ein Fehler 1.Art liegt vor, falls eine richtige Hypothese abgelehnt wird. Einen Fehler 2.Art begeht man, wenn man eine falsche Hypothese annimmt.

	H_0 ist wahr	H_0 ist falsch
H_0 wird abgelehnt	Entscheidung ist falsch Fehler 1.Art	Entscheidung ist richtig
H_0 wird angenommen	Entscheidung ist richtig	Entscheidung ist falsch Fehler 2.Art

TABELLE 44 Die vier möglichen Ausgänge bei Hypothesentests

Einen Fehler 1.Art begeht man mit der Irrtumswahrscheinlichkeit α, einen Fehler 2.Art mit einer Wahrscheinlichkeit β. Je kleiner α ist, desto größer ist β. Um das Ergebnis nicht im Nachhinein manipulieren zu können, muss die Irrtumswahrscheinlichkeit bereits vor dem Test angegeben werden. Wie schon bei Konfidenzintervallen bezeichnet man mit $\gamma = 1 - \alpha$ das Signifikanzniveau.

Üblicherweise verwendet man eine Irrtumswahrscheinlichkeit von $\alpha = 0{,}05$, was bedeutet, dass statistisch gesehen in 5 von 100 Fälle eine richtige Hypothese fälschlicherweise abgelehnt wird. Man spricht dann von einer von einer signifikanten Abweichung der Prüfgröße X vom erwarteten Wert. Mit $\alpha = 0{,}003$ wollen wir im weiteren Verlauf eine hochsignifikanten Abweichung der Prüfgröße X vom erwarteten Wert bezeichnen.[75]

Die Irrtumswahrscheinlichkeit muss zu Beginn eines Tests festgelegt werden. Dies soll verhindern, dass man das Testergebnis nicht im Nachhinein verfälschen und die Aussage beeinflussen kann.

[75] Dies ist in der Literatur nicht eindeutig. Gelegentlich meint auch schon $\alpha = 0{,}01$ eine hochsignifikante Abweichung.

BEISPIEL 13.13 *Gerichtsverfahren*[76]

Als Nullhypothese wollen wir von der Unschuldsvermutung des Angeklagten ausgehen. Die Alternativhypothese ist daher offensichtlich, dass der Angeklagte für schuldig befunden wird. Einen Fehler 1.Art begeht man, wenn ein Unschuldiger verurteilt wird, bei einem Fehler 2.Art wird ein Schuldiger freigesprochen.

BEMERKUNG 13.14 *Annahme und Ablehnung der Nullhypothese*

Als Ergebnis eines Hypothesentests wird die Nullhypothese H_0 entweder angenommen oder abgelehnt (bzw. verworfen). Die Nullhypothese dabei wird mit der Irrtumswahrscheinlichkeit α abgelehnt. Eine Nullhypothese ist jedoch keinesfalls bewiesen, falls sie nicht abgelehnt werden kann. Sie ist in diesem Fall lediglich statistisch nicht widerlegbar.

Im Fall, dass H_0 abgelehnt wird, hat man H_1 mit Sicherheit γ bewiesen. Wenn aber H_0 nicht abgelehnt wird, heißt das nur, dass man mit Sicherheit γ nichts gegen H_0 einwenden kann. Falls die Nullhypothese durch den Test verworfen wird, können wir eher davon ausgehen, keinen Fehler begangen zu haben, als in dem Fall, dass die Nullhypothese durch den Test angenommen wird.

BEMERKUNG 13.15 *Unterschied zwischen einseitigen und zweiseitigen Anteilstests*

Bei einem einseitigen Anteilstest interessiert man sich dafür, ob sich der ursprüngliche Anteil vergrößert bzw. ob er sich verkleinert hat. Bei einem zweiseitigen Anteilstest dagegen ist nur die Frage interessant, ob sich der ursprüngliche Anteil überhaupt verändert hat.

BEMERKUNG 13.16 *Vorgangsweise bei einseitigen Anteilstests*

♠ Aufstellung einer Nullhypothese und einer Alternativhypothese $H_0 : p \leq p_0$ bzw. $p \geq p_0, H_1 : p > p_0$ bzw. $p < p_0$
♠ Festlegung eines Signifikanzniveaus γ_0 bzw. einer Irrtumswahrscheinlichkeit α_0
♠ Bestimmung der absoluten Häufigkeit *k* mit der gewünschten Eigenschaft anhand einer Stichprobe vom Umfang *n*
♠ Berechnung der Irrtumswahrscheinlichkeit $P(X \geq k) = \alpha$ (rechtsseitiger Anteilstest) bzw. $P(X \leq k) = \alpha$ (linksseitiger Anteilstest) unter der Annahme, dass *X* binomialverteilt ist mit den Parametern *n* und p_0
♠ Ablehnung der Nullhypothese, falls $\alpha \leq \alpha_0$

[76] REICHEL, Hans-Christian u.a., Lehrbuch der Mathematik 7, Wien [2]1992, S.268

BEISPIEL 13.17 *Kanzlerfrage 1*

Bei einer Umfrage unter 500 Personen gaben 27% an, sie würden einen bestimmten Politiker zum Kanzler wählen, wenn sie es könnten. Ein Monat später waren es bereits 31%. Weist die zweite Umfrage auf einen signifikanten bzw. auf einen hochsignifikanten Anstieg der Befürworter hin?

Wir stellen die beiden folgenden Hypothesen auf:

$$H_0 : p \leq 0{,}27, H_1 : p > 0{,}27$$

Zunächst rechnen wir die Prozent in Stimmen um

$$0{,}27 \cdot 500 = 135; 0{,}31 \cdot 500 = 155$$

Die Anzahl der Wähler sei binomialverteilt. Wir legen einen einseitigen Anteilstest zugrunde, da uns nur interessiert, ob dieser Politiker in der Wählergunst gestiegen ist oder ob nicht. Wir testen rechtsseitig, weil wir $p > 0{,}27$ vermuten.

Mit Hilfe eines Computers erhalten wir

$$P\left(X \geq 155\right) = \sum_{i=155}^{500} \binom{500}{i} \cdot 0{,}27^i \cdot 0{,}73^{500-i} \cong 0{,}02595 < 0{,}05$$

Daher kann die Nullhypothese mit einer Irrtumswahrscheinlichkeit von 5% verworfen werden. Wir können also von einem signifikanten Stimmenzuwachs ausgehen.

Im zweiten Fall gelangen wir zu folgendem Resultat

$$\alpha_0 \cong 0{,}02595 > 0{,}003$$

$\Rightarrow$ kein hochsignifikanter Stimmenzuwachs

In diesem Beispiel handelt es sich also um einen signifikanten, aber nicht um einen hochsignifikanten Stimmenzuwachs.

BEISPIEL 13.18 *Kanzlerfrage 2*

Wir bleiben bei unserer Umfrage von 500 Personen, in der 27% sich für einen bestimmten Politiker aussprachen. Allerdings ergab nun eine zweite Umfrage nur mehr 25% Zustimmung. Haben wir es jetzt mit einem signifikanten bzw. einen hochsignifikanten Verlust der Befürworter zu tun?

Wir stellen die beiden Hypothesen

$$H_0 : p \geq 0{,}27, H_1 : p < 0{,}27$$

auf.

25% Zustimmung bedeutet, dass 125 Personen für diesen Kandidaten gestimmt haben. Wie schon in Beispiel 13.17 legen wir erneut eine Binomialverteilung zugrunde. Dieses Mal interessieren wir uns jedoch für einen linksseitigen Anteilstest, da nun die Frage ist, ob dieser Politiker in der Wählergunst gefallen ist.

Gehen wir zunächst der Frage bezüglich einem siginifikanten Stimmenverlust nach:

$$P\left(X \leq 125\right) = \sum_{i=0}^{125} \binom{500}{i} \cdot 0{,}27^{i} \cdot 0{,}73^{500-i} \cong 0{,}16944 > 0{,}05$$

Die Nullhypothese kann mit einer Irrtumswahrscheinlichkeit von 0,05 nicht verworfen werden. Hier handelt es sich also nicht um einen signifikanten und somit auch nicht um einen hochsignifikanten Stimmenverlust.

BEISPIEL 13.19 *Kanzlerfrage 3*

Auch bei Hypothesentests empfiehlt sich die Verwendung der Normalverteilung statt der Binomialverteilung, um den Rechenaufwand erheblich zu reduzieren.

Wir sehen, dass die Approximation durch die Normalverteilung sowohl bei Beispiel 13.16 als auch bei Beispiel 13.17 gerechtfertigt ist, da

$$\mu = 500 \cdot 0{,}27 = 135,\ \sigma = \sqrt{500 \cdot 0{,}27 \cdot 0{,}73} \cong 9{,}927$$

$$\sigma^2 = 98{,}55 > 9$$

Im Falle von Beispiel 13.17 erhalten wir

$$P\left(X \geq 155\right) \approx 1 - \Phi\left(z\right)$$

$$z = \frac{155 - \mu}{\sigma} \cong 2{,}01$$

$$1 - \Phi\left(2{,}01\right) = 0{,}0222 < 0{,}05$$

$\Rightarrow$ signifikanter Stimmenzuwachs.

Für Beispiel 13.18 gilt

$$P\left(X \leq 125\right) \approx \Phi\left(z\right)$$

$$z = \frac{125 - \mu}{\sigma} \cong -1{,}01$$

$$\Phi\left(-1{,}01\right) = 0{,}1562 > 0{,}05$$

$\Rightarrow$ kein signifikanter Stimmenverlust.

BEMERKUNG 13.20 *Vorgangsweise bei zweiseitigen Anteilstests*

- ♠ Aufstellung einer Nullhypothese und einer Alternativhypothese $H_0 : p = p_0, H_1 : p \neq p_0$
- ♠ Festlegung eines Signifikanzniveaus γ_0 bzw. einer Irrtumswahrscheinlichkeit α_0
- ♠ Bestimmung der absoluten Häufigkeit *k* mit der gewünschten Eigenschaft anhand einer Stichprobe vom Umfang *n*
- ♠ Berechnung der beiden Irrtumswahrscheinlichkeiten $P\left(X \leq k\right) = \alpha_1$ und $P\left(X \geq k\right) = \alpha_2$ unter der Annahme, dass *X* binomialverteilt mit den Parametern *n* und p_0 ist
- ♠ Ablehnung der Nullhypothese, falls $\alpha_1 \leq \frac{\alpha_0}{2}$ oder $\alpha_2 \leq \frac{\alpha_0}{2}$

BEISPIEL 13.21 *Änderung des Wähleranteils*

Eine Partei A habe bei einer Wahl 28% der Stimmen erhalten. Ein Jahr danach soll mit einer Irrtumswahrscheinlichkeit von 3% in Erfahrung gebracht werden, ob sich der Wähleranteil verändert wird. Von 500 Personen würden 120 Partei A wählen. Was lässt sich aufgrund dieses Umfrageergebnisses über das Wählerverhalten aussagen? Hat sich der Wähleranteil verändert?

Wir stellen die folgenden beiden Hypothesen auf:

$$H_0 : p = 0{,}28, H_1 : p \neq 0{,}28$$

Aufgrund von

$$P(X \leq 120) = \sum_{i=0}^{120} \binom{500}{i} \cdot 0{,}28^i \cdot 0{,}72^{500-i} \cong 0{,}02474 > \frac{0{,}03}{2}$$

$$P(X \geq 120) = \sum_{i=120}^{500} \binom{500}{i} \cdot 0{,}28^i \cdot 0{,}72^{500-i} \cong 0{,}98065 > \frac{0{,}03}{2}$$

können wir die Nullhypothese nicht verwerfen. Folglich hat sich der Wähleranteil mit einer Irrtumswahrscheinlichkeit von 3% nicht verändert.

BEISPIEL 13.22 *Änderung des Wähleranteils 2*

Auch bei Beispiel 13.20 können wir die Binomialverteilung durch die Normalverteilung ersetzen, da

$$\mu = 500 \cdot 0{,}28 = 140, \sigma = \sqrt{500 \cdot 0{,}28 \cdot 0{,}72} \cong 10{,}04$$

$$\sigma^2 = 100{,}8 > 9$$

Wir erhalten also

$$P(X \leq 120) \approx \Phi(z)$$

$$z = \frac{120 - \mu}{\sigma} \cong -1{,}99$$

$$\Phi(-1{,}99) = 0{,}0233 > \frac{0{,}03}{2}$$

sowie

$$P(X \geq 120) \approx 1 - \Phi(-1{,}99) = 0{,}9767 > \frac{0{,}03}{2}$$

$\Rightarrow$ keine Änderung des Wähleranteils

14. RANDOMIZED-RESPONSE-METHODEN

Ein großes Problem bei Wahlumfragen sind falsche Antworten der Befragten, die den tatsächlichen Wähleranteil einer Partei verzerren. Eine Möglichkeit der Behebung dieses Problems bieten die so genannten Randomized-Response-Methoden, mit denen sich der ungefähre Wähleranteil einer Partei trotz falschen Angaben bei Umfragen bestimmen lässt.

Bei einer Randomized-Response-Methode muss die befragte Person aus mehreren Fragen durch ein Zufallsexperiment eine auswählen. Auf welche Frage sich die Antwort bezieht, weiß somit nur die befragte Person, nicht jedoch der Interviewer. Deshalb kann davon ausgegangen werden, dass die befragte Person auch auf heikle Fragestellungen eher eine richtige Antwort gibt als bei einer direkten Befragung.

Ich möchte nun auf zwei derartige Modelle näher eingehen. Zum einen auf die Warner-Methode, zum anderen die Devore-Methode.

Erstmals publizierte Stanley L. Warner im Jahre 1965 im „Journal of the American Statistical Association“ eine Randomized-Response-Methode zur Beseitigung von falschen Antworten bei heiklen Fragestellungen.[77] Ein großer Nachteil der Warner-Methode ist jedoch, dass nur Zufallsexperimente verwendet werden dürfen, die nicht mit einer Wahrscheinlichkeit von $\alpha = \frac{1}{2}$ eintreten, da sich sonst kein Schätzer angeben lässt.

1977 schlug Jay L. Devore im Journal „Communications in Statistics – Theory Methods“ eine alternative Methode vor, die den großen Vorteil besaß, dass das Zufallsexperiment zur Beantwortung einer Frage auch die Wahrscheinlichkeit 0,5 besitzen durfte.[78]

[77] WARNER, Stanley, Randomized response: A Survey technique for eliminating evasive answer bias, In: Journal of the American Statistical Association, Vol 60(1965), Heft März, S.63-69

[78] siehe: LANG, Stefan, Randomized Response. Befragungstechniken zur Vermeidung von Verzerrung bei sensitiven Fragen, Institut für Statistik, Universität München, München o.J.

14.1 PROBLEMSTELLUNG

BEISPIEL 14.1 *Anteilsschätzung bei „ja“/ „nein“-Fragen*

Gesucht sei der tatsächliche Wähleranteil p einer Partei X. Um diesen durch eine Umfrage abschätzen zu können, wird n Personen die Frage gestellt, ob sie besagte Partei wählen würden, wenn am kommenden Sonntag Wahlen wären. Gehen wir zunächst einmal vom (unrealistischen) Idealfall aus, dass alle befragten Personen wahrheitsgemäß antworten.

Im Falle von ehrlichen Antworten erhalten wir die folgenden bedingten Wahrscheinlichkeiten.

$$P\left(\text{Antwort "ja"} \middle| \text{Parteiwähler}\right) = 1$$

$$P\left(\text{Antwort "ja"} \middle| \text{kein Parteiwähler}\right) = 0$$

$$P\left(\text{Antwort "nein"} \middle| \text{Parteiwähler}\right) = 0$$

$$P\left(\text{Antwort "nein"} \middle| \text{kein Parteiwähler}\right) = 1$$

Daraus folgt

$$\begin{aligned} P\left(\text{Antwort "ja"}\right) &\underset{3.8}{\overset{\text{Satz}}{=}} P\left(\text{Antwort "ja"} \middle| \text{Parteiwähler}\right) \cdot P\left(\text{Parteiwähler}\right) + \\ &+ P\left(\text{Antwort "ja"} \middle| \text{kein Parteiwähler}\right) \cdot P\left(\text{kein Parteiwähler}\right) = \\ &= 1 \cdot p + 0 \cdot (1-p) = p \end{aligned}$$

Sei X nun die Anzahl der „ja“-Sager unter den Interviewten eine binomialverteilte Zufallsvariable.

Als Maximum-Likelihood-Schätzer erhält man dann

$$\hat{P} := \frac{1}{n} \cdot \sum_{i=1}^{n} X_i = \frac{X}{n}.$$

mit der Varianz

$$V\left(\hat{P}\right) = V\left(\frac{X}{n}\right) = \frac{1}{n^2} \cdot V(X) = \frac{1}{n^{\cancel{2}}} \cdot \cancel{n} \cdot p \cdot (1-p) = \frac{p \cdot (1-p)}{n}.$$

Gerade bei politischen Fragen neigen die Befragten jedoch oft dazu falsche Antworten zu geben. Wir treffen die Annahme, dass jemand, der die Partei nicht wählen würde, auch wahrheitsgemäß mit nein antworten wird, währenddessen davon ausgegangen werden kann, dass nicht jeder Wähler dies auch zugeben würde. Es gilt also

$$P\left(\text{Antwort "ja"} \middle| \text{Parteiwähler}\right) < 1$$

$$P\left(\text{Antwort "ja"} \middle| \text{kein Parteiwähler}\right) = 0$$

Die Verzerrung von der Wirklichkeit ist um größer, desto geringer die erstgenannte Wahrscheinlichkeit ist.

Genau genommen könnte jemand auch angeben, eine Partei zu wählen, obwohl er das gar nicht tatsächlich beabsichtigt, da gewisse Parteien in der Öffentlichkeit angesehener sind als andere. Um diesen Fall ausschließen zu können, wollen wir uns im Folgenden mit der Wählerschaft von rechtsgerichteten Parteien auseinandersetzen. Hier kann sehr wohl davon ausgegangen werden, dass die obigen beiden bedingten Wahrscheinlichkeiten erfüllt sind.

Bei den Randomized-Response-Methoden werden den interviewten Personen mehrere Fragen vorgelegt. Der Trick bei diesen Verfahren besteht nun darin, dass der Befragte nur eine der Fragen beantwortet und dass der Interviewer nicht weiß, auf welche Frage sich die Antwort des Interviewten bezieht.

Der Befragte wählt nämlich mit einem Zufallsexperiment (Würfel, Münzwurf, Kartenziehen, etc.) eine der Fragen aus. Der Interviewer kennt nur die einzelnen Wahrscheinlichkeiten für das Eintreten des Zufallsexperiments, nicht jedoch dessen Ausgang.

Im Folgenden gehen wir davon aus, dass jeweils zwei verschiedene Fragen zur Auswahl stehen. Als Zufallsexperiment wählen wir das Kartenziehen. Eine gewisse Anzahl von Karten sei mit Frage 1 bedruckt, der Rest mit Frage 2. Die zu interviewende Person werde nun gebeten, eine Karte zu ziehen und die darauf abgebildete Frage wahrheitsgemäß zu beantworten.

Die Wahrscheinlichkeit, dass dabei Frage 1 gezogen wird, wollen wir in weiterer Folge mit α bezeichnen. Frage 2 trete somit mit einer Wahrscheinlichkeit von $1-\alpha$ ein:

$$P(\text{Frage 1}) := \alpha, P(\text{Frage 2}) := 1-\alpha$$

Weiters bezeichne π die Wahrscheinlichkeit, dass eine Person die gezogene Frage mit „ja“ beantwortet.

$$P(\text{Antwort "ja"}) := \pi$$

14.2 DAS MODELL VON WARNER

Jeder befragten Person werden eine konkrete Frage sowie deren Negation zur Beantwortung vorgelegt, z.B:

- ♠ Frage 1: Wenn kommenden Sonntag Wahlen wären, würden Sie dann Partei *X* wählen?
- ♠ Frage 2: Wenn kommenden Sonntag Wahlen wären, würden Sie dann eine andere Partei als Partei *X* wählen?

SATZ 14.2 *Der Warner-Schätzer*

Es werden *n* Personen befragt. Sei *X* die Anzahl der „Ja"-Sager unter den Interviewten. Dann ist

$$\hat{P}_W = \frac{\frac{X}{n} - (1-\alpha)}{2\alpha - 1}$$

der Warner-Schätzer für den tatsächlichen Wähleranteil *p* in der Bevölkerung mit $\alpha \neq \frac{1}{2}$.

BEWEIS:

Zunächst berechnen wir mit Hilfe des Satzes von der totalen Wahrscheinlichkeit die Wahrscheinlichkeit, dass eine befragte Person mit „ja" antwortet.

$$\begin{aligned}
P(\text{Antwort "ja"}) &\underset{3.8}{\overset{\text{Satz}}{=}} P(\text{Antwort "ja"} \mid \text{Frage 1}) \cdot P(\text{Frage 1}) + \\
&\quad + P(\text{Antwort "ja"} \mid \text{Frage 2}) \cdot P(\text{Frage 2}) = \\
&= p\alpha + (1-p)\cdot(1-\alpha) = p\alpha + (1-\alpha) - p\cdot(1-\alpha) = \\
&= p\cdot(2\alpha - 1) + (1-\alpha) := \pi_W
\end{aligned}$$

Den Schätzer $\hat{P}_W$ erhalten wir, indem wir die relative Häufigkeit der „Ja"-Sager mit dem soeben berechneten π_W gleichsetzen.

$$P(\text{Antwort "ja"}) = \frac{X}{n}$$

$$\hat{P}_W \cdot (2\alpha - 1) + (1-\alpha) = \frac{X}{n}$$

$$\hat{P}_W = \frac{\frac{X}{n} - (1-\alpha)}{2\alpha - 1}$$

□

BEISPIEL 14.3 *Anteilsschätzung mit dem Warner-Schätzer*

Es werden 1000 Personen gefragt, ob sie eine bestimmte Partei wählen. Es sollen die beiden Fragen von oben verwendet werden, wobei Frage 2 doppelt so oft gezogen werden kann wie Frage 1. 620 Personen antworten nun mit „ja“. Somit gilt für den Warner-Schätzer:

$$\hat{P}_W = \frac{\frac{620}{1000}-\left(1-\frac{1}{3}\right)}{2\cdot\frac{1}{3}-1} = 0{,}14$$

Somit kann der ungefähre Wähleranteil dieser Partei mit ungefähr 14% geschätzt werden.

SATZ 14.4 *Erwartungswert und Varianz des Warner-Schätzers*

Der Warner-Schätzer besitzt die beiden folgenden Kennzahlen:

$$E\left(\hat{P}_W\right) = p, V\left(\hat{P}_W\right) = \frac{\pi_W\cdot(1-\pi_W)}{n\cdot(2\alpha-1)^2} = \frac{p\cdot(1-p)}{n} + \frac{\alpha\cdot(1-\alpha)}{n\cdot(2\alpha-1)^2}$$

BEWEIS:

Da X binomialverteilt ist mit den Kenngrößen $E(X) = n\cdot\pi_W$ und $V(X) = n\cdot\pi_W\cdot(1-\pi_W)$, erhalten wir

$$E\left(\hat{P}_W\right) = E\left(\frac{\frac{X}{n}-(1-\alpha)}{2\alpha-1}\right) \underset{5.4}{\overset{\text{Satz}}{=}} \frac{\frac{1}{n}\cdot E(X)-(1-\alpha)}{2\alpha-1} =$$

$$= \frac{\frac{1}{\cancel{n}}\cdot\cancel{n}\cdot\pi_W-(1-\alpha)}{2\alpha-1} = \frac{p\cdot\cancel{(2\alpha-1)}+\cancel{(1-\alpha)}-\cancel{(1-\alpha)}}{\cancel{2\alpha-1}} = p$$

$$V\left(\hat{P}_W\right) = V\left(\frac{\frac{X}{n}-(1-\alpha)}{2\alpha-1}\right) \underset{5.11}{\overset{\text{Satz}}{=}} \frac{1}{n^2\cdot(2\alpha-1)^2}\cdot V(X) =$$

$$= \frac{\cancel{n}\cdot\pi_W\cdot(1-\pi_W)}{n^{\cancel{2}}\cdot(2\alpha-1)^2} = \frac{\pi_W\cdot(1-\pi_W)}{n\cdot(2\alpha-1)^2}$$

Die Varianz lässt sich noch weiter umformen:

$$V\left(\hat{P}_W\right) = \frac{\pi_W\cdot(1-\pi_W)}{n\cdot(2\alpha-1)^2} = \frac{1}{n\cdot(2\alpha-1)^2}\cdot\left[\pi_W\cdot(1-\pi_W)\right] =$$

$$= \frac{1}{n\cdot(2\alpha-1)^2}\cdot\left[p\cdot(2\alpha-1)+(1-\alpha)\right]\cdot\left[1-p\cdot(2\alpha-1)-(1-\alpha)\right] =$$

$$= \frac{1}{n\cdot(2\alpha-1)^2}\cdot\left[p\cdot(2\alpha-1)+(1-\alpha)\right]\cdot\left[\alpha-p\cdot(2\alpha-1)\right] =$$

$$
\begin{aligned}
&= \frac{1}{n\cdot(2\alpha-1)^2}\cdot\Big[\alpha\cdot p\cdot(2\alpha-1)-p^2\cdot(2\alpha-1)^2+\\
&\qquad +\alpha\cdot(1-\alpha)-p\cdot(1-\alpha)\cdot(2\alpha-1)\Big]=\\
&= \frac{1}{n\cdot(2\alpha-1)^2}\cdot\Big[(2\alpha-1)\cdot p\cdot\big(\alpha-(1-\alpha)\big)-p^2\cdot(2\alpha-1)^2=\\
&\qquad +\alpha\cdot(1-\alpha)\Big]=\\
&= \frac{1}{n\cdot(2\alpha-1)^2}\cdot\Big[(2\alpha-1)^2\cdot p\cdot(1-p)+\alpha\cdot(1-\alpha)\Big]=\\
&= \frac{p\cdot(1-p)}{n}+\frac{\alpha\cdot(1-\alpha)}{n\cdot(2\alpha-1)^2}
\end{aligned}
$$

□

Man erkennt, dass der erste Summand der Varianz des Warner-Schätzers mit der Varianz des Schätzers bei lauter richtigen Antworten übereinstimmt. Der zweite Summand gibt an, wie sehr die ursprüngliche Varianz vergrößert wird, und zwar je mehr, desto näher der Wert α bei $\frac{1}{2}$ liegt. Dieser Sachverhalt lässt sich auch graphisch verdeutlichen.

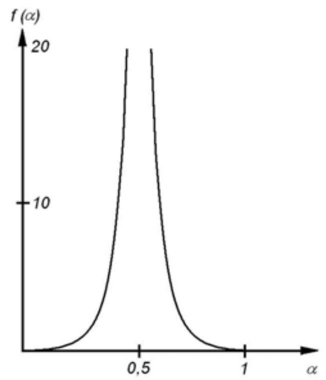


ABBILDUNG 27 Graphische Darstellung der Funktion $f(\alpha)=\frac{\alpha\cdot(1-\alpha)}{(2\alpha-1)^2}$

Das größte Manko der Warner-Methode ist offensichtlich, dass der Warner-Schätzer für $\alpha=\frac{1}{2}$ nicht definiert ist. Dieses Problem tritt Modell von Devore nicht auf.

14.3 DAS MODELL VON DEVORE

Neben der konkreten Frage wird dieses Mal jeder befragten wahlberechtigten Person auch eine Frage vorgelegt, die als Antwort stets ein „ja“ verlangt, z.B:

♠ Frage 1: Sind Sie wahlberechtigt?
♠ Frage 2: Wenn kommenden Sonntag Wahlen wären, würden Sie dann Partei *X* wählen?

SATZ 14.5 *Der Devore-Schätzer*

Es werden *n* Personen befragt. Sei *X* die Anzahl der „Ja“-Sager unter den Interviewten. Dann ist

$$\hat{P}_D = \frac{\frac{X}{n} - \alpha}{1 - \alpha}$$

der Devore-Schätzer für den tatsächlichen Wähleranteil *p* in der Bevölkerung.

BEWEIS:

Wie schon beim Warner-Schätzer berechnen wir zunächst die Wahrscheinlichkeit, dass eine befragte Person mit „ja“ antwortet:

$$\begin{aligned} P(\text{Antwort "ja"}) &\underset{3.18}{\overset{\text{Satz}}{=}} P(\text{Antwort "ja"} | \text{Frage 1}) \cdot P(\text{Frage 1}) + \\ &\quad + P(\text{Antwort "ja"} | \text{Frage 2}) \cdot P(\text{Frage 2}) = \\ &= 1 \cdot \alpha + p \cdot (1 - \alpha) := \pi_D \end{aligned}$$

Als Devore-Schätzer $\hat{P}_D$ erhalten wir somit

$$P(\text{Antwort "ja"}) = \frac{X}{n}$$

$$1 \cdot \alpha + \hat{P}_D \cdot (1 - \alpha) = \frac{X}{n}$$

$$\hat{P}_D = \frac{\frac{X}{n} - \alpha}{1 - \alpha}$$

□

BEISPIEL 14.6 *Anteilsschätzung mit dem Devore-Schätzer*

Es werden 1000 Personen gefragt, ob sie eine bestimmte Partei wählen. Seien die beiden oben genannten Fragen gleich wahrscheinlich. Wenn 590 Personen mit „ja“ antworten, dann erhalten wir den folgenden Devore-Schätzer:

$$\hat{P}_D = \frac{\frac{X}{n} - \alpha}{1 - \alpha} = \frac{\frac{590}{1000} - \frac{1}{2}}{\frac{1}{2}} = 0{,}18$$

Somit liegt der ungefähre Wähleranteil dieser Partei bei ungefähr 18%.

Zu dem gleichen Ergebnis wären wir auch mit folgender Überlegung gekommen: Wir gehen davon aus, dass 500 Personen Frage 1 gezogen haben und 500 Personen Frage 2. Somit stammen 90 der insgesamt 590 „Ja"-Antworten von denjenigen, die Frage 2 gezogen haben. Dies führt uns zu der Annahme, dass der Wähleranteil dieser Partei bei rund $\frac{90}{500} \mathrel{\hat{=}} 18\%$ liegt.

SATZ 14.7 *Erwartungswert und Varianz des Devore-Schätzers*

Der Devore-Schätzer besitzt die beiden folgenden Kennzahlen:

$$E\left(\hat{P}_D\right) = p, V\left(\hat{P}_D\right) = \frac{\pi_D \cdot \left(1-\pi_D\right)}{n \cdot \left(1-\alpha\right)^2}$$

BEWEIS:

Die binomialverteilte Zufallsvariable X besitze die Kenngrößen $E\left(X\right) = n \cdot \pi_D$ und $V\left(X\right) = n \cdot \pi_D \cdot \left(1-\pi_D\right)$. Somit folgt für den Erwartungswert und die Varianz des Devore-Schätzers

$$E\left(\hat{P}_D\right) = E\left(\frac{\frac{X}{n}-\alpha}{1-\alpha}\right) \overset{\text{Satz}}{\underset{5.4}{=}} \frac{\frac{1}{n}\cdot E\left(X\right)-\alpha}{1-\alpha} = \frac{\frac{1}{\cancel{n}}\cdot \cancel{n}\cdot \pi_D - \alpha}{1-\alpha} =$$

$$= \frac{\cancel{\alpha} + p \cdot \cancel{\left(1-\alpha\right)} \cancel{-\alpha}}{\cancel{1-\alpha}} = p$$

$$V\left(\hat{P}_D\right) = V\left(\frac{\frac{X}{n}-\alpha}{1-\alpha}\right) \overset{\text{Satz}}{\underset{5.11}{=}} \frac{1}{n^2 \cdot \left(1-\alpha\right)^2} \cdot V\left(X\right) = \frac{\cancel{n} \cdot \pi_D \cdot \left(1-\pi_D\right)}{n^{\cancel{2}} \cdot \left(1-\alpha\right)^2} =$$

$$= \frac{\pi_D \cdot \left(1-\pi_D\right)}{n \cdot \left(1-\alpha\right)^2}$$

□

Verwendete Literatur zu Randomized-Response-Methoden:

- ♠ BÜCHTER, Andreas / HENN, Hans-Wolfgang, Elementare Stochastik. Eine Einführung in die Mathematik der Daten und des Zufalls, Berlin / Heidelberg [2]2006, S.247
- ♠ EASTAWAY, Rob / WYNDHAM, Jeremy, How many people watch Coronation Street?, In: Why do buses come in threes? The Hidden Mathematics of Everyday Live, London 2006, S.29-40
- ♠ HESSE, Christian / MEISTER, Alexander, Übungsbuch zur angewandten Wahrscheinlichkeitstheorie. Aufgaben und Lösungen, Wiesbaden 2005, S.109-110; 133-135
- ♠ LANG, Stefan, Randomized Response. Befragungstechniken zur Vermeidung von Verzerrung bei sensitiven Fragen, Institut für Statistik, Universität München, München o.J.
- ♠ WARNER, Stanley, Randomized response: A Survey technique for eliminating evasive answer bias, In: Journal of the American Statistical Association, Vol 60(1965), Heft März, S.63-69

ZUSAMMENFASSUNG

Mit dieser Diplomarbeit setze ich mich mit den Grundlagen der Wahrscheinlichkeitsrechnung auseinander und gebe eine Einführung in die Stochastik. Ich habe versucht, einen möglichst anschaulichen Zugang zu wählen und dabei viele unterschiedliche Anwendungsbeispiele aus politischen Bereichen zu präsentieren, wo Stochastik eine entscheidende Rolle spielt.

Im Zuge dieser Arbeit bin ich auf die folgenden politischen Fragestellungen näher eingegangen:

- Absolute Mehrheit in einem Drei-Parteien-System (Kapitel 1)
- Stimmzettelproblem (Kapitel 2)
- Simpsonsche Paradoxon anhand der Zusammensetzung der Geschworeren-Jurys in Neuseeland (Kapitel 3)
- Manipulation von Wahlergebnissen (Kapitel 8)
- Mandatsverteilungen und eine stochastische Analyse der Fairness beim Hare-Verfahren (Kapitel 9)
- Verschiedene Machtindizes anhand der Machtverhältnisse in der EWG, Österreichs Einfluss im Weltsicherheitsrat und die Macht des amerikanischen Präsidenten (Kapitel 10)
- Gründe für falsche Wahlprognosen (Kapitel 11)
- Schätzen des Stichprobenumfangs (Kapitel 12)
- Mathematisches Modell von Meinungsumfragen (Kapitel 13)
- Bestimmung des tatsächlichen Wähleranteils einer Partei mithilfe der Randomized-Response-Methoden (Kapitel 14)

ABSTRACT

In this master thesis I'm dealing with the basics of the theory of probability and giving an introduction into stochastics. I tried to choose an approach as clear as possible and in doing so to present many different examples of use from political fields, where stochastics play a decisive role.

In the course of the work I elaborated on the following political questions:

♠ absolute majority in a three-parties-system (chapter 1)

♠ ballot problem (chapter 2)

♠ Simpson's paradox in connection with the composition of juries in New Zealand (chapter 3)

♠ manipulation of election results (chapter 8)

♠ distributions of mandates and a stochastic analysis of equity in Hare's method (chapter 9)

♠ different voting power indices on the basis of the relations of power in the EEC, Austria's influence in the Security Council and the power of the American president (chapter 10)

♠ reasons for inaccurate predictions of elections (chapter 11)

♠ estimation of the amount of a random sample (chapter 12)

♠ mathematical model of opinion polls (chapter 13)

♠ determination of the real number of voters of a party by using Randomized-Response-Methods (chapter 14)

LITERATURVERZEICHNIS

BÜCHER

- ♠ ARBEITSGEMEINSCHAFT MEDIA-ANALYSEN, Media-Analyse 2005 Jahresbericht, Wien 2005
- ♠ BARTH, Friedrich / HALLER, Rudolf, Stochastik Leistungskurs, München [12]1998
- ♠ BECK-BORNHOLDT, Hans-Peter / DUBBEN, Hans-Hermann, Der Hund, der Eier legt. Erkennen von Fehlinformation durch Querdenken, Reinbek bei Hamburg 2001
- ♠ BERNHARD, Martin / KOPP, Günther, Mathematik Repetitorium, Wien 1993
- ♠ BOL, Georg, Wahrscheinlichkeitstheorie. Einführung, München 2007
- ♠ BOSCH, Karl, Großes Lehrbuch der Statistik, München 1996
- ♠ BOSCH, Karl, Grundzüge der Statistik. Einführung mit Übungen, München 1996
- ♠ BOSCH, Karl, Statistik-Taschenbuch, München [2]1993
- ♠ BOURIER, Günther, Wahrscheinlichkeitsrechnung und schließende Statistik. Praxisorientierte Einführung mit Aufgaben und Lösungen, Wiesbaden [4]2005
- ♠ BRAMS, Steven, Game Theory and Politics, New York 1975
- ♠ BÜCHTER, Andreas / HENN, Hans-Wolfgang, Elementare Stochastik. Eine Einführung in die Mathematik des Zufalls, Berlin / Heidelberg [2]2006
- ♠ CHRISTOPH, Gerd / HACKEL, Horst, Starthilfe Stochastik, Stuttgart / Leipzig / Wiesbaden 2002
- ♠ CLAUß, Günter / FINZE, Falk-Rüdiger / PARTZSCH, Lothar, Statistik für Soziologen, Pädagogen, Pychologen und Mediziner. Grundlagen, Frankfurt am Main [4]2002
- ♠ DEHLING, Herold / HAUPT, Beate, Einführung in die Wahrscheinlichkeitstheorie und Statistik, Berlin / Heidelberg [2]2004
- ♠ DUBBEN, Hans-Hermann / BECK-BORNHOLDT, Hans-Peter, Mit an Wahrscheinlichkeit grenzender Sicherheit. Logisches Denken und Zufall, Reinbek bei Hamburg [3]2006
- ♠ FELSENTHAL, Dan S. / MACHOVER, Moshé, The Measurement of Voting Power. Theory and Practice, Problems and Paradoxes, Cornwall 1998
- ♠ GRIFFITH, Elmer, The Rise and Development of the Gerrymander, Dissertation, The University of Chicago, Chicago 1907
- ♠ HESSE Christian, Angewandte Wahrscheinlichkeitstheorie, Braunschweig / Wiesbaden 2003

- ♠ HESSE, Christian / MEISTER, Alexander, Übungsbuch zur angewandten Wahrscheinlichkeitstheorie. Aufgaben und Lösungen, Wiesbaden 2005
- ♠ HOLLER, Manfred / ILLING, Gerhard, Einführung in die Spieltheorie, Berlin ^{6}2006
- ♠ IRLE, Albrecht, Wahrscheinlichkeitstheorie und Statistik. Grundlagen – Resultate – Anwendungen, Stuttgart 2001
- ♠ KOHN, Wolfgang, Statistik. Datenanalyse und Wahrscheinlichkeitsrechnung, Berlin / Heidelberg 2005
- ♠ KOPFERMANN, Klaus, Mathematische Aspekte der Wahlverfahren. Mandatsverteilung bei Abstimmungen, Mannheim 1991
- ♠ KRENGEL, Ulrich, Einführung in die Wahrscheinlichkeitstheorie und Statistik, Wiesbaden ^{7}2003
- ♠ KRÖPFL, Bernhard, u.a., Angewandte Statistik. Eine Einführung für Wirtschaftswissenschaftler, München / Wien ^{2}1999
- ♠ LIPSCHUTZ, Seymour, Wahrscheinlichkeitsrechnung. Theorie und Anwendungen, Frankfurt am Main 1999
- ♠ MALLE, Günther, u.a., Mathematik verstehen 8, Wien 2007
- ♠ MOSLER, Karl / SCHMID, Friedrich, Wahrscheinlichkeitsrechnung und schließende Statistik, Berlin / Heidelberg 2006
- ♠ PAPULA, Lothar, Mathematik für Ingenieure und Naturwissenschaftler Band 3. Vektoranalysis, Wahrscheinlichkeitsrechnung, Mathematische Statistik, Fehler- und Ausgleichsrechnung, Wiesbaden ^{5}2008
- ♠ REICHEL, Hans-Christian u.a., Lehrbuch der Mathematik 7, Wien ^{2}1992
- ♠ REICHEL, Hans-Christian / MÜLLER, Robert / HANISCH, Günter, Lehrbuch der Mathematik 8, Wien ^{2}1993
- ♠ ROTH, Dieter, Empirische Wahlforschung. Ursprung, Theorien, Instrumente und Methoden, Opladen 1998
- ♠ SCHWÄRZLER, Jürgen, Methodische Überlegungen zu Wählerstromanalyse und Wahlhochrechnung, Diplomarbeit, Universität Wien, Wien 2000
- ♠ STATISTIK AUSTRIA (Hg.), Statistisches Jahrbuch Österreich 2006, Wien 2005
- ♠ SZEKELEY, Gábor, Paradoxa. Klassische und neue Überraschungen aus Wahrscheinlichkeitsrechung und mathematischer Statistik, Frankfurt am Main 1990
- ♠ TAYLOR, Alan, Mathematics and Politics. Strategy, Voting, Power and Proof, New York 1995
- ♠ THORWARTL, Wolfram / WAGNER, Günther / WAGNER, Helga, Mathematik positiv 8.Klasse AHS. Band 1: Musterbeispiele und Aufgaben, Wien 1998

- ♠ WOLPERS, Hans / GÖTZ, Stefan, Didaktik der Stochastik, In: Tietze, Uwe-Peter / Klika, Manfred / Wolpers, Hans (Hg.), Mathematik in der Sekundarstufe II. Band 3, Braunschweig / Wiesbaden 2002
- ♠ WÜSTENDÖRFER, Werner, Einführung in die Statistik. Für pädagogische und soziale Berufe, Nürnberg ^{2}2005

AUFSÄTZE UND ZEITSCHRIFTENARTIKEL

- ♠ Bundesministerium für Inneres, Amtsblatt vom 23.9.2004
- ♠ BLYTH, Colin, On Simpson's Paradox and the Sure-Thing Principle, In: Journal of the American Statistical Association Vol. 67(1972), Nummer 338, S.364-366
- ♠ CARNAL, Henri / RIEDWYL, Hans, Wer kommt ins Parlament?, In: Spektrum der Wissenschaft 09/2002, S.80-84
- ♠ EASTAWAY, Rob / WYNDHAM, Jeremy, How many people watch Coronation Street?, In: Why do buses come in threes? The Hidden Mathematics of Everyday Live, London 2006, S.29-40
- ♠ JAHNKE, Thomas, Das Simpsonsche Paradoxon verstehen – ein Beitrag des Mathematikunterrichts zur Allgemeinbildung, In: Journal für Mathematik-Didaktik Vol. 14(1993), Heft 3/4, S.221-242
- ♠ KRÄMER, Walter, Wie lügt man mit Statistik?, In: Stochastik in der Schule 1/1991, S.3-24 (vor allem S.11-12)
- ♠ KÜNZEL, Erhard, Über Simpsons Paradoxon, In: Stochastik in der Schule 1/1991, S.54-62
- ♠ LANG, Stefan, Randomized Response. Befragungstechniken zur Vermeidung von Verzerrung bei sensitiven Fragen, Institut für Statistik, Universität München, München o.J.
- ♠ LUCAS, William, Gewichtete Wahlsysteme: Wie man Macht messen kann, In: Garfunkel, Salomon / Steen, Lynn (Hg.), Mathematik in der Praxis. Anwendungen in Wirtschaft, Wissenschaft und Politik, Heidelberg, 1989, S.138-152
- ♠ MAUSBERG, Wolfgang (Hg.), Anteile, Zugriffe und Reihenfolgen (AZUR). Berechnungsverfahren für Sitzverteilungen und Reihenfolgen bei der Besetzung von Gremien und für die Zuteilung sonstiger Berechtigungen im parlamentarischen Bereich. ZI 5 – Arbeitspapier 1998/002, Bonn 1998
- ♠ MEYER, Jörg, Wahlen: Paradoxa bei der Sitzverteilung, In: Mathematica didactica Vol. 18(1995), Heft 1, S.21-34
- ♠ NEUWIRTH, Erich, Die Bedeutung der Statistik für politische Meinungsumfragen, In: Frei, Norbert / Heintel, Peter (Hg.), Politische Bildung als Unterrichtsprinzip. Konsequenzen für den Unterricht, Klagenfurt 1985, S.517-527
- ♠ PICHLER, Herbert, Wie „macht" man ein Wahlergebnis? Ein erster Blick in Wahlrecht, Wahlarithmetik und Wahlgeographie, In: Von Wahl zu Wahl. Informationen zur Politischen Bildung Nr. 21, Wien 2004, S.88

- ♠ PLASSER, Fritz / ULRAM, Peter, Die Wahlanalyse 2006. Wer hat wen warum gewählt?, Presseunterlage Fessel-GfK, Wien 2006
- ♠ PUKELSHEIM, Friedrich, Die Väter der Mandatszuteilungsverfahren, In: Spektrum der Wissenschaft 09/2002, S.83
- ♠ PUKELSHEIM, Friedrich, Divisor oder Quote? Zur Mathematik von Mandatszuteilungen bei Verhältniswahlen, Report 392, Institut für Mathematik, Universität Augsburg, Augsburg 1998
- ♠ RÖTTEL, Karl, Berechnung von Sitzverteilungen in Ausschüssen, In: Arbeitskreis Mathematik und Bildung (Hg.), Mathematik – unsichtbar, doch allgegenwärtig, Eichstätt 2002, S.75-98
- ♠ SCHICK, Karl, Wahlberechnungsverfahren, In: Praxis der Mathematik Vol. 28(1986), Heft 2, S.81-88, 105-118
- ♠ WARNER, Stanley, Randomized response: A Survey technique for eliminating evasive answer bias, In: Journal of the American Statistical Association, Vol 60(1965), Heft März, S.63-69
- ♠ WESTBROOKE, Ian, Simpson's Paradox. An example in a New Zealand Survey of Jury Composition. Research Report #1 (revised), Wellington 1997
- ♠ ZENTRUM POLIS. Politik Lernen in der Schule (Hg.), Nationalratswahlen und Wahlrecht in Österreich. Polis aktuell Nr. 7/2006

INTERNETLINKS

- ♠ http://de.wikipedia.org/wiki/BZÖ [05.02.2009]
- ♠ http://de.wikipedia.org/wiki/Gerrymander [30.07.2010]
- ♠ http://de.wikipedia.org/wiki/Liberales_Forum [05.02.2009]
- ♠ http://de.wikipedia.org/wiki/Liste_Martin [05.02.2009]
- ♠ http://de.wikipedia.org/wiki/Sicherheitsrat_der_Vereinten_Nationen [18.05.2007]
- ♠ http://de.wikipedia.org/wiki/Wahlforschung [05.02.2009]
- ♠ http://wahl06.bmi.gv.at/gkz_9.htm [03.02.2009]
- ♠ http://webarchiv.bundestag.de/archive/2005/0919/ftp/pdf_arch/azur.pdf [03.02.2009]
- ♠ http://www.bundestag.de/parlament/funktion/gesetze/bwahlg_pdf.pdf [02.03.2009]
- ♠ http://www.bundestag.eu/ftp/pdf_arch/azur.pdf [03.02.2009]
- ♠ http://www.bmi.gv.at/wahlen/ [01.07.2007]
- ♠ http://www.bmi.gv.at/wahlen/Nationalrats-Wahlordnung_1992.pdf [03.02.2009]
- ♠ http://www.bmi.gv.at/wahlen/nrw02info.asp [31.12.2008]
- ♠ http://www.bmi.gv.at/wahlen/NRW_06_gesamtergE.asp [03.02.2009]

- ♠ http://www.bmi.gv.at/wahlen/wahldownloads/bpwahl04_ErgE.pdf [29.12.2008]
- ♠ http://www.bmi.gv.at/wahlen/wahldownloads/NRW_06/NRW_06_Mandatsspiegel V.pdf [03.02.2009]
- ♠ http://www.bundeswahlleiter.de/bundestagswahl2002/deutsch/ergebnis2002/bund_land/btw2002/kru11_btw2002.htm [11.06.2007]
- ♠ http://www.loc.gov/exhibits/treasures/trr113.html [03.02.2009]
- ♠ http://www.math.uni-augsburg.de/stochastik/pukelsheim/2002g.html abrufbar [03.02.2009]
- ♠ http://www.nationalatlas.gov/printable/congress.html#list [03.03.2009]
- ♠ http://www.noe.gv.at/service/politik/landtag/Geschichte.htm [08.06.2007]
- ♠ http://www.politik-lernen.at/goto/polis/details/pa_nationalratswahlen/ [03.02.2009]
- ♠ http://www.statistik-berlin.de/wahlen/vorwahlen/btw-98/ergebnisse/karten/karte3.html [03.02.2009]
- ♠ http://www.stats.govt.nz/NR/rdonlyres/5DFEE5C8-A969-4684-B95F-FA2AFEDA649A/0/Simpdox.pdf [01.01.2009]
- ♠ http://www.wahlrecht.de/doku/download/index.htm [03.02.2009]
- ♠ http://www.westmiller.com/fairvote2k/in_gerry.htm [03.02.2009]

ABBILDUNGSVERZEICHNIS

TABELLENVERZEICHNIS

LEBENSDATEN DER VORKOMMENDEN PERSONEN

- ♠ ANDRAE, Carl (1812-1893): dänischer Mathematiker und Ministerpräsident
- ♠ BANZHAF, John F. (* 1940): amerikanischer Jurist
- ♠ BAYES, Thomas (1702-1761): englischer Pfarrer
- ♠ BERNOULLI, Jacob (1654-1705): Schweizer Mathematiker
- ♠ DEEGAN, John: amerikanischer Politologe des 20./21.Jahrhunderts
- ♠ DE MOIVRE, Abraham (1667-1754): französischer Mathematiker
- ♠ DE MORGAN, Augustus (1806-1871): englischer Mathematiker
- ♠ DEVORE, JAY L.: amerikanischer Statistiker des 20./21.Jahrhunderts
- ♠ D'HONDT, Victor (1841-1901): belgischer Rechtswissenschafter
- ♠ FEINSTEIN, Alvan R. (1925-2001): amerikanischer Mathematiker und Epidemiologe
- ♠ FELSENTHAL, Dan (* 1938): israelischer Politikwissenschafter
- ♠ GAUß, Carl Friedrich (1777-1855): deutscher Mathematiker, Physiker und Astronom
- ♠ GERRY, Elbridge (1744-1814): einer der Unterzeichner der amerikanischen Unabhängigkeitserklärung und fünfter Vizepräsident der Vereinigten Staaten von Amerika
- ♠ HAGENBACH-BISCHOFF, Eduard (1833-1910): Schweizer Physiker
- ♠ HAMILTON, Alexander (1755-1804): amerikanischer Politiker, einer der Väter der amerikanischen Verfassung
- ♠ HARE, Thomas (1806-1891): britischer Jurist
- ♠ JEFFERSON, Thomas (1743-1826): Verfasser der amerikanischen Unabhängigkeitserklärung, Begründer der demokratischen Partei und dritter Präsident der Vereinigten Staaten von Amerika
- ♠ JOHNSTON, Ronald John (* 1941): britischer Geograph
- ♠ KOLMOGOROW, Andrej Nikolajewitsch (1909-1987): russischer Mathematiker
- ♠ LAPLACE, Pierre Simon (1749-1827): französischer Mathematiker und Astronom
- ♠ MACHOVER, Moshé (* 1936): israelischer Professor für Mathematik und Philosophie
- ♠ NIEMEYER, Horst F. (*1931): deutscher Mathematiker
- ♠ PACKEL, Edward Wesler (* 1941): amerikanischer Mathematiker
- ♠ PASCAL, Blaise (1623-1662): französischer Naturwissenschafter
- ♠ PENROSE, Lionel (1898-1972): britischer Psychiater und Mathematiker

- ♠ QUETELET, Lambert Adolphe Jacques (1796-1874): belgischer Astronom und Statistiker
- ♠ ROGERS, Will (1879-1935): amerikanischer Komiker
- ♠ SHAPLEY, Lloyd S. (*1923): amerikanischer Mathematiker
- ♠ SHUBIK, Martin (*1926): amerikanischer Wirtschaftswissenschafter
- ♠ SIMPSON, Edward Hugh (*1922): britischer Statistiker
- ♠ STEINER, Jakob (1796-1863): Schweizer Mathematiker
- ♠ STIRLING, James (1692-1770): schottischer Mathematiker
- ♠ SYLVESTER, James Joseph (1814-1897): britischer Mathematiker
- ♠ TSCHEBYSCHEW, Pafnuti Lwowitsch (1821-1894): russischer Mathematiker
- ♠ VINTON, Samuel F. (1792-1862): amerikanischer Politiker
- ♠ VON MISES, Richard Edler (1883-1953): österreichischer Mathematiker
- ♠ WALLIS, John (1616-1703): englischer Mathematiker
- ♠ WARNER, Stanley L. (1928-1992): amerikanischer Wirtschaftswissenschafter

CURRICULUM VITAE

ALLGEMEINES

- ♠ Name: Markus Dorn
- ♠ Geburtsdatum: 9.Februar 1980
- ♠ Geburtsort: Wien
- ♠ Staatsbürgerschaft: Österreich

AUSBILDUNG

- ♠ 4 Klassen Albertus Magnus Volksschule, Michaelerstraße 12, 1180 Wien
- ♠ 8 Klassen Albertus Magnus Gymnasium, Semperstraße 45, 1180 Wien
  Englisch ab 1. Klasse, Latein ab 3. Klasse, Französisch ab 5. Klasse
  Matura im Juni 1998 bestanden
- ♠ ab Oktober 1998 Lehramtstudium Mathematik und Geschichte an der Universität Wien

BERUFLICHE TÄTIGKEITEN

- ♠ 1997-1998: Eisgeschäft Leonardelli
- ♠ August 1999 – September 2003: Arbeit im Nachhilfeinstitut Studienkreis (Inhaber: Ing. Dr. Zimmermann)
- ♠ Schülerbetreuer bei der Zahlenjagd im Juni 2003 und 2004 (siehe www.zahlenjagd.at)
- ♠ ab Oktober 2003: Arbeit im Nachhilfeinstitut Lernquadrat (Inhaber: Ing. Dr. Zimmermann) (siehe www.lernquadrat.at)
- ♠ WS 2003 – SS 2008: Stud.Ass. an der Universität für Bodenkultur Wien. Abhaltung von mathematischen Übungsgruppen für Kulturtechnik und Wasserwirtschaft und für Lebensmitteltechnologen (siehe www.boku.ac.at/math)
- ♠ ab Dezember 2007: Vertragslehrer für Mathematik und GZ am BRG Rahlgasse 4,1060 Wien (siehe www.ahs-rahlgasse.at)
- ♠ September 2009-August 2010: Vertragslehrer für Mathematik am Phönix-Realgymnasium (privatgeführtes Gymnasium), Knöllgasse 20-24, 1100 Wien (siehe www.phoenixrealgymnasium.at)