Einführung in die Wahrscheinlichkeitstheorie und Statistik

Dr. C.J. Luchsinger

1 Wahrscheinlichkeit

Literatur Kapitel 1

* gelegentlich lesen: Cartoon Guide: Kapitel 1 und 2 (diese Fragen behandle ich in meiner Vlsg nicht)
* Cartoon Guide: Kapitel 3
* Krengel: § 1 und 2
* Storrer: 32, 33, 34, 35, 36, 48

1.1 Motivation zur Wahrscheinlichkeitstheorie und Statistik

1.1.1 Was fällt ein bei den Begriffen "Wahrscheinlichkeit" und/oder "Statistik"?

Input:

Ordnung:

1.1.2 Wo wird Statistik eingesetzt, statistische Fragestellungen?

* theoretische, mathematische, spielerische Fragen: Münzwurf, Würfelspiele, weitere

* Biologie / Medizin: Wirkung von Medikamenten zur Blutdrucksenkung (Stichprobengrösse, Nebenwirkungen, Spätfolgen, viele Jobs bei Novartis, Roche etc)

* Soziologie / Demographie / Statistische Ämter: Volkszählung (muss man alle befragen, wie Stichprobe auswählen)

* Finanzmathematik: Wie soll man Geld anlegen? Preis von Finanzprodukten (Black-Scholes)? Risikomanagement?

* Versicherungsmathematik: Wie viel Prämie soll man für eine Leistung verlangen? Verwendung der Prämie als Versicherung?

* Technik / Physik / Qualitätskontrolle in der Produktion

Mehr zu den Anwendungen und Berufen für MathematikerInnen:

  * www.math-jobs.ch/beruf.html
  * www.math-jobs.com

1.2 Geschichtliches

Für Jahrhunderte haben Menschen statistische Argumente (aus der Testtheorie, mehr in Kapitel 6) benutzt. Dazu zwei Beispiele:

a) (Gregor von Tours, 6. Jh.; nach Robert Ineichen) 20 Krieger, "die Gott nicht fürchteten", setzten über einen Fluss, um ein Kloster zu plündern und die Mönche zu töten. Auf dem Rückweg kamen alle bis auf einen in den Fluten um: Gregor: "Nur einer von ihnen blieb unverletzt, jener, der die anderen wegen ihres Vorhabens getadelt hatte. Wenn jemand denkt, dies sei durch Zufall geschehen, so möge er bedenken, dass *ein unschuldiger* Mann gerettet wurde, *inmitten so vieler Verbrecher*". [Hypothesen aufstellen (es gibt keinen Gott), Testen (genau der hat überlebt, welcher an Gott glaubt/unschuldig ist), sog. P-Wert sehr klein (Wahrscheinlichkeit hierfür berechnen unter der Annahme "gibt keinen Gott"), also Null-Hypothese ablehnen (also hier: "es gibt doch einen Gott - der für Gerechtigkeit sorgt")]

b) Ankläger vor Gericht: "Gut, nehmen wir einmal an, der Angeklagte ist unschuldig und war die ganze Nacht zu Hause [Null-Hypothese, In dubio pro reo; Strategie des Anklägers ist offenbar, die Argumente der Verteidigung ad absurdum zu führen, lächerlich zu machen]. Ist es dann nicht sonderbar unwahrscheinlich, dass das Messer in seiner Wohnung gefunden wurde [kleiner P-Wert, Wahrscheinlichkeit für diese Ereignis, gegeben die Null-Hypothese]? Deshalb ist die Annahme, dass der Angeklagte unschuldig ist falsch [Null-Hypothese verwerfen und Alternativ-Hypothese annehmen].

17. Jahrhundert: Abbruch eines Glückspiels mitten im Spiel (Bullen sind unerwartet aufgetaucht), und man ist in Führung. Wieviel vom denkbaren Gewinn steht einem *fairerweise* zu?

19. Jahrhundert: Landwirtschaftliche Versuche, Variabilität im Pflanzenwachstum, Düngemethoden

20. Jahrhundert: saubere Fundierung der Wahrscheinlichkeitstheorie und Siegeszug durch alle denkbaren (und undenkbaren (Geisteswissenschaften!)) Anwendungsgebiete.

21. Jahrhundert wird das Jahrhundert der Biologie / Medizin *(wenn die Mathematik / Statistik die dabei auftretenden Probleme lösen kann (z.B. in der Genetik)).*

Namen: Jakob Bernoulli (Uni Basel), Moivre, Laplace, Gauss, Poisson, Tschebyschew, Markov, Kolmogoroff, Fisher und viele weitere mehr.

1.3 Kombinatorik [Permutation, Kombination, Variation; ohne und mit Wiederholungen]

Die Kombinatorik untersucht, auf wieviele verschiedene Arten man Gegenstände, Buchstaben, Zahlen, Objekte anordnen kann. Also z.B. "Wie viele verschiedene, natürliche Zahlen kann man unter Verwendung von bis zu 4 Ziffern der Menge $\{1, 3, 7\}$ bilden (man darf die gleiche Ziffer mehrmals benutzen)? Man kann dann dabei mehr oder weniger geschickt vorgehen. Die Antwort ist übrigens:

Die Kombinatorik sollte bereits aus der Mittelschule bekannt sein. Das erste Übungsblatt ist diesem Gebiet gewidmet. Begriffe wie "Permutation, Kombination, Variation; ohne und mit Wiederholungen" gehören in dieses Gebiet. Ein Professor der Wahrscheinlichkeitstheorie/Statistik weiss normalerweise (im Gegensatz zu einem Mittelschullehrer oder Maturand) nicht auswendig, was hinter welchem Begriff steht. Dies ist auch nicht notwendig. Die Kombinatorik ist für uns nur ein Hilfsmittel und wir finden jeweils immer wieder ad hoc von neuem heraus, wie eine Formel sein muss. Dies macht man am Besten mit einem Beispiel mit kleiner Anzahl von etwa $n = 5$, wo n die Anzahl involvierter Objekte ist. Die einzigen beiden Ausdrücke, welche man kennen muss, sind:

Fakultät

$$n! := n(n-1)(n-2)\cdots 2 \cdot 1,$$

wir *definieren* weiter $0! := 1$.

Binomialkoeffizient

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

Classroom Assessment: Sei c eine beliebige, konstante, reelle Zahl grösser 1. n gehe in den folgenden 3 Ausdrücken gegen ∞. Ordnen Sie die drei Terme nach Ihrer Wachstumsgeschwindigkeit, d.h. welcher dominiert bei grossem n welche anderen und in welcher Reihenfolge.

$$n! \qquad n^c \qquad c^n.$$

Auch wenn wir die Begriffe im Titel zu 1.3 nicht auswendig lernen werden (dies ist an der Prüfung auch nicht verlangt!), wollen wir doch je ein Beispiel anschauen:

1. **Permutationen mit verschiedenen Elementen:** Jede Zusammenstellung einer endlichen Anzahl von Elementen in irgendeiner Anordnung, in der sämtliche Elemente verwendet werden, heisst Permutation der gegebenen Elemente. TSP im OR:

2. **Permutationen mit zum Teil gleichen Elementen:** Im ersten Fall haben wir ja $n!$ Permutationen erhalten. Treten jetzt aber Gruppen von gleichen Elementen auf, haben wir weniger Permutationen: Sei $e_1 = a, e_2 = b, e_3 = a$, so haben wir auf den ersten Blick durchaus 3 Elemente und erwarten $3! = 6$ Permutationen. Leider ist aber $e_1e_2e_3$ gleich wie $e_3e_2e_1$. Wir haben $2!$ Anordnungen (e_1e_3 und e_3e_1), welche eigentlich gleich sind. Deshalb gibt es hier nur $3!/2! = 3$ Permutationen.

Wir fassen die Permutationen zusammen:

* Anzahl Permutationen von n verschiedenen Elementen: $n!$

* n Elemente, k Gruppen von je $l_1, l_2, \ldots, l_k$ gleichen Elementen ($\sum_{i=1}^{k} l_i = n$). Dann haben wir nur noch

$$\frac{n!}{l_1!l_2! \cdots l_k!}$$

Permutationen.

Vorbereitende Bemerkung für 3., 4., 5. und 6. Jede Zusammenstellung von k Elementen aus n Elementen heisst eine Kombination $k-$ter Klasse oder $k-$ter Ordnung. Von den vier Elementen a, b, c, d sind Kombinationen 2. Klasse: $ab, ac, ad, bc, bb, aa, \ldots$. Werden nur verschiedene Elemente zur Zusammenstellung ausgewählt, so liegen *Kombinationen ohne Wiederholung*, sonst *Kombinationen mit Wiederholung* vor. Werden zwei Kombinationen, die gleiche Elemente, aber in verschiedener Anordnung enthalten, als verschieden betrachtet, so werden sie *Variation* genannt

3. **Variation ohne Wiederholung:** Das erste Element kann auf n verschiedene Arten gewählt werden. Das nächste auf $(n-1)$ verschiedene Arten und so weiter. Wir erhalten: Anzahl Variationen ohne Wiederholung von n Elementen zur $k-$ten Klasse ist

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

4. **Variation mit Wiederholung:** Sind Wiederholungen zugelassen, so kann auch das zweite, dritte usw. Element auf n Arten gewählt werden. Wir erhalten: Anzahl Variationen mit Wiederholungen von n Elementen zur $k-$ten Klasse ist

$$n^k.$$

Man betrachte nochmals das einleitende Beispiel zu Sektion 1.3.

5. **Kombination ohne Wiederholung:** Hier wird die Anordnung ausser Betracht gelassen. Würde man die k gewählten Elemente noch permutieren, hätte man ja den Fall 3. Es gibt $k!$ solche Permutationen. Die Anzahl Kombinationen ohne Wiederholung muss deshalb $k!$ mal kleiner sein als die Anzahl bei Fall 3:

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

wir erhalten hier den Binomialkoeffizienten!

6. **Kombination mit Wiederholung:** Hier ordnet man am Besten die Elemente (z.B. $k = 2$ und $n = 3$ resp. n allgemein) in einem Dreieck an:

Wir erhalten für die Anzahl: $3 + 2 + 1 = 6$ resp. $n + (n-1) + \cdots + 1 = (n+1)n/2$. Allgemein erhält man (Beweis durch vollständige Induktion, $k = 2$ als Verankerung und $k > 2$ zu zeigen): Die Anzahl Kombinationen mit Wiederholung ist

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

Sektion 1.3 erfolgte in Anlehnung an Kapitel 27 aus "Handbuch der Mathematik"; Buch und Zeit Verlag, Köln (1986).

1.4 Das Konzept der Wahrscheinlichkeit P

Wir werden jetzt die naive Vorstellung von Wahrscheinlichkeit und Zufall formalisieren, in die Sprache der Mathematik übersetzen. Als erstes müssen wir, ohne die Wahrscheinlichkeiten anzuschauen, Versuchsausgänge formalisieren.

1.4.1 Ereignisraum (auch Menge der Versuchsausgänge)

* Sei Ω eine nichtleere Menge.

* Sie steht für die Menge der Versuchsausgänge; wir nennen sie auch Ereignisraum. Es findet jeweils in einem Experiment genau ein sogenanntes Elementarereignis statt, z.B. $\omega_1 \in \Omega$ oder $\omega_2 \in \Omega$ etc.

* Bei einem Würfel wird man sinnvollerweise mit $\Omega := \{1, 2, 3, 4, 5, 6\}$ die Menge der möglichen Anzahl Augen wählen.

* Bei Münzwurf wird man sinnvollerweise $\Omega := \{K, Z\}$ wählen. $k-$facher Münzwurf: $\Omega := \{K, Z\}^k$.

* Die Menge Ω kann von endlicher oder unendlicher Kardinalität sein; sogar überabzählbar unendlich (die Begriffe "überabzählbar unendlich" und weitere folgen noch im ersten Semester in der Infinitesimalrechnung I, sie sind hier nicht weiter von Belang):

  * abzählbar unendliches Ω: zufällige Zahl aus $\Omega := \mathbb{N}$
  * überabzählbar unendliches Ω: Zeit bis Atom zerfällt ist im Intervall $\Omega := (0, \infty)$ und zufällig, zumindest in gewissen physikalischen Modellen. Unendlicher Münzwurf: $\Omega := \{K, Z\}^\infty$.

* Beim Würfel können wir fortfahren mit: A ist mit $A := \{2, 4, 6\}$ die Menge der geraden Zahlen und $B := \{4, 5, 6\}$ die Menge der Zahlen streng grösser 3.

* A bzw. B nennen wir ein Ereignis. Ereignisse sind allgemein Teilmengen von Ω. Dies ist verwirrend. Es kann ja (Ereignis A) nicht gleichzeitig $2, 4$ und 6 geworfen werden. Es gilt

nach wie vor, dass genau ein sogenanntes Elementarereignis stattfindet. Wenn dieses ω_1 z.B. gleich 2 ist, so sagen wir: A ist eingetreten - es ist ja auch eine gerade Zahl geworfen worden; eine reicht!

Wir fassen die mengentheoretischen Ausdrücke und ihre Bedeutung für die Wahrscheinlichkeitstheorie in folgender Tabelle zusammen:

Symbol	**Mengentheorie / Bedeutung für die WT**
Ω	Menge / Ereignisraum, Menge der Versuchsausgänge
ω	Element von Ω / Elementarereignis, Versuchsausgang
A	Teilmenge von Ω / Ereignis; falls $\omega \in A$, sagt man, dass das Ereignis A eingetreten ist
A^c	Komplement von A / kein Elementarereignis aus A findet statt
$A \cap B$	Schnittmenge von A und B / ein Elementarereignis aus A und B findet statt
$A \cup B$	Vereinigung von A und B / ein Elementarereignis aus A oder B findet statt
$A \backslash B$	A ohne B / ein Elementarereignis aus A tritt ein, aber nicht aus B
$A \subset B$	A ist Teilmenge von B / Wenn ein Elementarereignis aus A stattfindet, dann immer auch ein Elementarereignis aus B
$\emptyset$	leere Menge / unmögliches Ereignis
Ω	ganze Menge / sicheres Ereignis (etwas muss passieren)

1.4.2 σ-Algebra

Wir wollen den Ereignissen (z.B. A aus Ω) später eine Wahrscheinlichkeit ($P[A]$) zuordnen. Wenn wir mehrere Ereignisse vorgegeben haben, wollen wir auch die Wahrscheinlichkeiten von deren Vereinigungen, Durchschnitten oder Komplementen angeben können. An die Menge der Teilmengen von Ω, welche wir untersuchen, stellen wir also ein paar wenige Bedingungen:

Definition 1.1 [**σ-Algebra**] *Ein Teilmengensystem $\mathcal{A}$ von Ω heisst σ-Algebra, wenn folgende 3 Bedingungen erfüllt sind:*

a) $\Omega \in \mathcal{A}$

b) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$

c) $A_1, A_2, \ldots \in \mathcal{A} \Rightarrow \cup_{n \geq 1} A_n \in \mathcal{A}$.

1. Wieso muss $\emptyset$ immer in einer σ-Algebra enthalten sein?

2. Welches ist die kleinste σ-Algebra überhaupt?

3. Wieso muss mit A und B immer auch $A \cap B$ in einer σ-Algebra enthalten sein?

4. Welches ist die kleinste σ-Algebra, welche Ereignis A enthält?

Falls $|\Omega| = n < \infty$, so hat die Potenzmenge (Menge aller Teilmengen) von Ω bekanntlich Kardinalität 2^n, ist also wiederum endlich. Wir werden im Fall $|\Omega| = n < \infty$ in dieser Vorlesung einfach als $\mathcal{A}$ die Potenzmenge von Ω wählen.

Nur nebenbei sei erwähnt, dass man z.B. im Fall $\Omega = \mathbb{R}$ etliche Klippen überwinden muss(te), um sowohl diese Konzepte wie auch die dazugehörigen Wahrscheinlichkeitsmasse $P[.]$ korrekt und ohne Widersprüche zu definieren. Dies wird in der Vlsg WT behandelt. Wir gehen hier nicht weiter darauf ein, sondern präsentieren im Eilzugtempo: Wenn wir $\Omega = \mathbb{R}$ wählen, dann ist die sogenannte *Borel-σ-Algebra* $\mathcal{B}(\mathbb{R})$ die kleinste σ-Algebra auf $\mathbb{R}$, welche z.B. alle geschlossenen Intervalle enthält. Dieses Teil-Mengensystem enthält auch alle halboffenen und alle offenen Intervalle, auch "fast alle" sonstigen Teilmengen von $\mathbb{R}$. Man muss aktiv und gezielt Mengen konstruieren, welche in dieser σ-Algebra *nicht* enthalten sind!

1.4.3 Wahrscheinlichkeit $P[.]$

Definition 1.2 [Wahrscheinlichkeit P] *Eine Wahrscheinlichkeit P ist eine reellwertige Funktion auf den Mengen aus $\mathcal{A}$. Dabei müssen folgende 3 Bedingungen erfüllt sein:*

a) $A \in \mathcal{A} \Rightarrow 0 \leq P[A] \leq 1$

b) $P[\Omega] = 1$

c) Sei $\{A_i\}_{i=1}^{\infty}$ eine abzählbare Folge von disjunkten Mengen aus $\mathcal{A}$, dann muss gelten:

$$P[\cup_{i=1}^{\infty} A_i] = \sum_{i=1}^{\infty} P[A_i].$$

Man darf in Definition 1.2 c) z.B. auch $A_i = \emptyset, i \geq 3$ wählen!

Es gibt eine extrem praktische Interpretation von **Wahrscheinlichkeiten und Proportionen**: Wir haben 7.5 Millionen EinwohnerInnen in der Schweiz. Wenn man jetzt zufällig (jedeR EinwohnerIn hat gleiche Wahrscheinlichkeit, ausgewählt zu werden, also (1/7'500'000)) eineN EinwohnerIn herausgreift, so ist beispielsweise die Wahrscheinlichkeit, einen Basler herauszugreifen, genau gleich der Proportion der Basler an der Schweizerischen Bevölkerung. Wir werden diese Analogie in vielen Fällen einsetzen.

Aus Definition 1.2 lassen sich nützliche Eigenschaften ableiten, welche wir im folgenden Lemma zusammenfassen.

Lemma 1.3 [nützliche Eigenschaften von P] *Mit $A, B \in \mathcal{A}$ gelten folgende Aussagen:*

a) Prinzip der Gegenwahrscheinlichkeit: $P[A] = 1 - P[A^c]$; v.a. $P[\emptyset] = 0$

b) Sei $\{A_i\}_{i=1}^{\infty}$ eine abzählbare Folge von Mengen aus $\mathcal{A}$, dann muss gelten:

$$P[\cup_{i=1}^{\infty} A_i] \leq \sum_{i=1}^{\infty} P[A_i].$$

c) $A \subseteq B \Rightarrow P[B] = P[A] + P[B \backslash A]$

d) $A \subseteq B \Rightarrow P[A] \leq P[B]$

e) $P[A \cup B] = P[A] + P[B] - P[A \cap B]$.

Beweis von Lemma 1.3 Wir beweisen hier nur Aussage a); Aussagen b), c), d) und e) sind als einfache Übungen zu lösen.

Es gelten $A \cup A^c = \Omega$ und $A \cap A^c = \emptyset$. Wir wählen $A_1 = A$, $A_2 = A^c$ und $A_i = \emptyset, i \geq 3$ in Definition 1.2 c). Deshalb gilt:

$$1 = P[\Omega] = P[A \cup A^c] = P[A] + P[A^c] + \sum_{i=3}^{\infty} P[\emptyset].$$

Daraus folgt die Behauptung.

$\square$

1.4.4 Wahrscheinlichkeitsraum

Definition 1.4 [Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$] *Wir nennen ein Tripel der Form $(\Omega, \mathcal{A}, P)$ Wahrscheinlichkeitsraum.*

Bevor wir uns den *bedingten* Wahrscheinlichkeiten zuwenden, wollen wir noch untersuchen, von welcher Gestalt die Wahrscheinlichkeiten in zwei wichtigen Fällen sind:

1. $|\Omega|$ endlich: $\Omega := \{1, 2, \ldots, n\}$.

$\underline{p_i \Rightarrow P, P[A]}$ Wir wählen zuerst beliebige Werte $p_1, \ldots, p_n \in [0, 1]$ und wollen fortfahren mit $P[\{i\}] := p_i; 1 \leq i \leq n$. Worauf ist jetzt noch zu achten?

$\underline{P \Rightarrow p_i, P[A]}$

Ein wichtiger, einfacher Fall ist die **Uniform-Verteilung auf** $\Omega := \{1, \ldots, n\}$: $p_i := 1/n, 1 \leq i \leq n$. Wir erhalten dann für beliebiges $A \subset \{1, 2, \ldots, n\}$:

$$P[A] = \sum_{i \in A} p_i = \sum_{i \in A} \frac{1}{n} = \frac{|A|}{n}.$$

2. $|\Omega|$ abzählbar unendlich, z.B. $\Omega := \mathbb{N}$: Mit den gleichen Schritten wie oben erhalten wir einerseits, dass wir beliebige $p_i \in [0, 1], i \geq 1$, wählen können, sodass $\sum_{i=1}^{\infty} p_i = 1$ gelten muss. Andererseits gilt auch hier, dass alle Wahrscheinlichkeiten von der Form $P[A] = \sum_{i \in A} p_i$ sind, wo $A \subseteq \Omega$.

   * Kann es hier eine sinnvolle Uniform-Verteilung geben?

   * Geben Sie 2 Beispiele von möglichen p_i's an im Fall $\Omega := \mathbb{N}$.

1.4.5 Bedingte Wahrscheinlichkeit $P[A|B]$

Betrachten wir ein sehr einfaches Beispiel: Wurf eines fairen Würfels. Wie ist die Wahrscheinlichkeit, eine gerade Zahl zu werfen (Ereignis A), gegeben wir haben eine Zahl kleiner als 4 (Ereignis B). Die Antwort ist:

Wie sind wir darauf gekommen? Klar war, dass wir uns nur um die Zahlen $\{1, 2, 3\}$ kümmern müssen (gegeben kleiner 4). Das ist unser neuer Ereignisraum. Jedes Elementarereignis hat jetzt Wahrscheinlichkeit 1/3 (früher 1/6). Die Antwort ist demnach 1/3, weil 2 die einzige gerade Zahl kleiner 4 (und grösser 0) ist. Was haben wir aber wirklich gemacht?

$$\frac{1}{3} = 2 * \frac{1}{6} = \frac{1/6}{1/2} = \frac{P[A \cap B]}{P[B]}.$$

Wir müssen durch "$P[B]$" teilen, weil wir wieder eine Wahrscheinlichkeit haben wollen (muss auf 1 summieren). Deshalb definieren wir:

Definition 1.5 [Bedingte Wahrscheinlichkeit $P[A|B]$]

$$P[A|B] := \frac{P[A \cap B]}{P[B]},$$

falls $P[B] > 0$. Man nennt $P[A|B]$ die bedingte Wahrscheinlichkeit von A gegeben B.

Es gilt:

$$P[A|B]P[B] = P[A \cap B] = P[B|A]P[A].$$

Der Leser / die Leserin zeige: $P[.|B]$ ist selber auch eine Wahrscheinlichkeit:

Sei $B \subset A$. Zeigen Sie, dass dann gilt: $P[A|B] = 1$. Warum muss dies so sein (anschaulich)?

Wir leiten über zu 1.4.6; untersuchen Sie $P[A|B]$, falls $P[A \cap B] = P[A]P[B]$.

1.4.6 Unabhängigkeit von Ereignissen

Anschaulich und in der Modellierung bedeutet unabhängig, dass zwei Ereignisse von unabhängigen Mechanismen erzeugt wurden.

Definition 1.6 [Unabhängigkeit von Ereignissen] *Ereignisse A und B sind (paarweise) unabhängig voneinander, wenn*

$$P[A \cap B] = P[A]P[B].$$

Falls $P[B] > 0$ resp. $P[A] > 0$, so ist Unabhängigkeit von A und B gleichbedeutend mit

$$P[A|B] = P[A] \qquad \text{("who cares" – Variante)}$$

resp. $P[B|A] = P[B]$.

Bei mehr als 2 involvierten Ereignissen ist Vorsicht am Platz. Beginnen wir zuerst mit der Definition: Die Ereignisse $A_1, \ldots, A_n$ sind in der Gesamtheit unabhängig voneinander, wenn für beliebige Auswahl $k = 1, \ldots, n$ und $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ gilt:

$$P[A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}] = P[A_{i_1}]P[A_{i_2}] \ldots P[A_{i_k}].$$

Aus der paarweisen Unabhängigkeit folgt jedoch nicht zwingend die (gesamtheitliche) Unabhängigkeit! Entwickeln Sie aus der folgenden Ausgangslage ein Gegenbeispiel dazu: Sei $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$. Alle elementaren Ereignisse seien gleichwahrscheinlich (1/4). Untersuchen Sie die Ereignisse $A = \{\omega_1, \omega_2\}$, $B = \{\omega_1, \omega_3\}$ und $C = \{\omega_1, \omega_4\}$ unter diesem Gesichtspunkt.

1.4.7 Formel von Bayes

Bsp von Cartoon Guide, p 46-50:

* A: Patient hat bestimmte Krankheit (HIV, Grippe, etc.)
* B: Test hierfür ist positiv
* $P[A] = 0.001$ [1 aus 1000 hat Krankheit]
* $P[B|A] = 0.99$ [W'keit für positiven Test wenn Patient wirklich krank 99 %]
* $P[B|A^c] = 0.02$ [W'keit für positiven Test obschon Patient gesund 2 %]

$$P[A|B] = ? \quad \text{[W'keit Patient hat Krankheit gegeben der Test ist positiv]}$$

Können wir dies berechnen? Falls ja, wie gross ist W'keit?

Angenommen wir kennen $P[B|A]$, wollen aber eigentlich $P[A|B]$ wissen. Es gilt nach Definition

$$P[A|B] = \frac{P[A \cap B]}{P[B]}.$$

Des weiteren gelten folgende Mengengesetze: $B = (B \cap A) \cup (B \cap A^c)$, wobei die Mengen $(B \cap A)$ und $(B \cap A^c)$ disjunkt sind. Also gilt

$$P[B] = P[B \cap A] + P[B \cap A^c]$$

und weiter

$$P[B \cap A] = P[A \cap B] = P[B|A]P[A];$$

analog erhalten wir

$$P[B \cap A^c] = P[B|A^c]P[A^c].$$

Zusammen ergibt sich die *Formel von Bayes*:

$$P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{P[B|A]P[A]}{P[B|A]P[A] + P[B|A^c]P[A^c]}.$$

1.4.8 Formel von der totalen Wahrscheinlichkeit FTW

In der letzten Formel haben wir im 2. Gleichheitszeichen im Nenner bereits einen Spezialfall der FTW angewendet. Allgemein formulieren wir die FTW in folgendem

Lemma 1.7 [Formel von der totalen Wahrscheinlichkeit FTW] $B_1, B_2, \ldots$ *sei eine Partition von* Ω *(die* B_i*'s sind disjunkt und* $\cup_{i=1}^{\infty} B_i = \Omega$*). Weiter sei für alle* $B_i, i \geq 1$, $P[B_i] > 0$ *erfüllt. Dann gilt für jedes* $A \in \mathcal{A}$*:*

$$P[A] = \sum_{i=1}^{\infty} P[A|B_i]P[B_i]. \tag{FTW}$$

Wie man am Beweis unten sofort sieht, gilt ein analoges Resultat auch für eine endliche Partition.

Gutes Beispiel dazu: Proportion von SchweizerInnen mit Vermögen über 1 Million CHF und wir haben leider nur die kantonalen Daten. Lsg: A ist Ereignis, eine Person auszuwählen mit Vermögen über 1 Million. $P[A|B_i]$ Wahrscheinlichkeit hierfür, gegeben suche nur unter Einwohnern im Kanton i, $P[B_i]$ Anteil Bewohner in Kanton i an gesamter Bevölkerung der Schweiz.

Beweis von Lemma 1.7:

$\square$

Wahrscheinlichkeitsbäume (Rauchen und KKW)

Wir fügen noch zwei miteinander eng verwandte Formeln hinzu:

Lemma 1.8 [$\lim P[A_n]$ **bei ab- oder aufsteigenden Folgen**] *Sei* $A_1, A_2, \ldots$ *eine aufsteigende Folge von Ereignissen, d.h.* $A_1 \subseteq A_2 \subseteq A_3 \ldots$. *Wir definieren in dem Fall*

$$A := \lim_{i \to \infty} A_i := \cup_{i \geq 1} A_i.$$

Dann gilt $P[A] = \lim_{i \to \infty} P[A_i]$.

Analog gilt: Sei $B_1, B_2, \ldots$ *eine absteigende Folge von Ereignissen, d.h.* $B_1 \supseteq B_2 \supseteq B_3 \ldots$. *Wir definieren in dem Fall*

$$B := \lim_{i \to \infty} B_i := \cap_{i \geq 1} B_i.$$

Dann gilt $P[B] = \lim_{i \to \infty} P[B_i]$.

Beweis von Lemma 1.8: $A = A_1 \cup (A_2 \backslash A_1) \cup (A_3 \backslash A_2) \cup \ldots$, wobei dies Vereinigungen disjunkter Mengen sind. Damit gilt:

$$\begin{aligned} P[A] &= P[A_1] + \sum_{i \geq 1} P[A_{i+1} \backslash A_i] \\ &= P[A_1] + \lim_{n \to \infty} \sum_{i=1}^{n} [P[A_{i+1}] - P[A_i]] \\ &= \lim_{n \to \infty} P[A_n]. \end{aligned}$$

Das analoge Resultat einer fallenden Folge kann man mit Komplementbildung und dem ersten Teil beweisen. Dies ist in den Übungen zu lösen.

$\square$