6. Grundlagen des Wahrscheinlichkeitsrechnens

6 Grundlagen des Wahrscheinlichkeitsrechnens
6.1 Einführung
6.2 Wozu benötigen wir Wahrscheinlichkeitsangaben?
6.3 Der Ursprung der Wahrscheinlichkeitsrechnung
6.4 Begriffserklärungen und Notationen
6.5 Wahrscheinlichkeiten
6.5.1 Die empirische Wahrscheinlichkeit
6.5.2 Die klassische Wahrscheinlichkeit
6.5.3 Die axiombasierte Wahrscheinlichkeit
6.6 Irrtumswahrscheinlichkeit
6.7 Übungen

6.1 Einführung

Am Ende des 5. Kapitels steht, dass im Bereich $\bar{x} \pm 1 * s_{\bar{x}}$ mit 68,28 %iger Wahrscheinlichkeit der Mittelwert liegt, wobei wir nicht näher auf die Bedeutung des Begriffs Wahrscheinlichkeit eingegangen sind. Wir wollen uns nun mit den elementaren Grundlagen der Wahrscheinlichkeitsrechnung soweit beschäftigen, wie es für das Verständnis der folgenden Texte erforderlich ist. In Redewendungen sagen wir immer wieder, ein zukünftiges Ereignis träte wahrscheinlich ein, sehr wahrscheinlich ein, mit an Sicherheit grenzender Wahrscheinlichkeit ein oder es wäre eher unwahrscheinlich, dass es einträte.

A: Um meine Kopfschmerzen zu unterdrücken nehme ich Medikament X, weil ich aufgrund meiner Erfahrung meine, dass dann die Symptome wahrscheinlich zurückgehen.
B: Weil ich die Wettersituation schon oft erlebt habe, sage ich, dass es in einer halben Stunde sehr wahrscheinlich ein Gewitter geben wird.
C: Wenn ich einen sechseitigen Würfel werfe, sage ich intuitiv voraus, dass mit einer Wahrscheinlichkeit von 16,7% die Ziffer 5 oben liegen wird.
D: Wenn ich beim Roulette meinen Einsatz auf noir plaziere, dann weiß ich, dass die Wahrscheinlichkeit zu gewinnen gleich 48,65% ist.
E: Ich lese in einem Protokoll: „Der Unterschied in der Wirkung der beiden Medikamente ist mit 99% iger Wahrscheinlichkeit signifikant, daher wird das besser wirkende Medikament für die Therapie empfohlen.“
F: Oder ich lese: „Mit 5%iger Irrtumswahrscheinlichkeit liegt der Erwartungswert μ zwischen 13,6 mm und 14,2 mm.“

Bei A und B kann ich mich auf Empirie stützen. Meine Erfahrungen reichen aber nicht um die Wahrscheinlichkeitsaussagen zu quantifizieren. Zahlen sind aber notwendig, wenn ich mit Wahrscheinlichkeiten rechnen muss. Anders ist die Situation in den Fällen C und D. Hier kann ich ohne jede eigene Erfahrung auf dem Gebiet intuitiv eine Zahlenwert für die Wahrscheinlichkeit nennen. Und bei E und F konnten wir, wie auch bei den Aussagen zum Standardfehler des Mittelwertes in Kapitel 5 – nach vorangegangenen Berechnungen - präzisen Wahrscheinlichkeitsangaben machen. Offensichtlich gibt es Situationen, in denen wir quantitative Wahrscheinlichkeiten angeben können und solche, in denen das nicht geht.

6.2 Wozu benötigen wir Wahrscheinlichkeitsangaben in der Statistik?

In der **deskriptiven Statistik** benötigen wir quantitative Wahrscheinlichkeitsangaben um z.B. den Konfidenzbereich zu berechnen, das ist der Bereich um den Mittelwert herum, in dem der Erwartungswert μ mit z.B. 99%iger Wahrscheinlichkeit liegt. Oder um festzustellen, mit welcher Wahrscheinlichkeit eine beliebige Walnuss (Kapitel 3) weniger als 55,0 g wiegt. Ein Grundverständnis des Begriffs Wahrscheinlichkeit ist notwendig zur Berechnung der Form der idealen Normalverteilungskurve und z.B. zum Verständnis anderer Verteilungen wie Binominalverteilung, Poissonverteilung, t-Verteilung, Chi²-Verteilung.

In der **Inferenzstatistik** spielt die Wahrscheinlichkeit z.B. eine Rolle bei Hypothesenprüfungen wie in der folgenden Situation: 12 Mäuse erhielten das Narkotikum A und blieben im Mittel 75 min lang narkotisiert. 12 weitere Mäuse erhielten Narkotikum B und blieben 65 min in Narkose. Auf den ersten Blick wirkt A länger als B. Dieser Unterschied könnte zufallsbedingt sein (siehe Zufallsfehler). Zufällig könnten in der A-Gruppe die Tiere im Mittel etwas länger in Narkose geblieben sein. Den Einfluss des Zufalls müssen wir mit einem Test prüfen. Obwohl die Narkosezeiten etwas anderes sagen, stellen wir die negierende Hypothese auf, dass Präparat A nicht zu einer überzufälligen (statistisch gesicherten = signifikanten) längeren Narkosezeit führt als Präparat B. Dann berechnen wir mit den Werten der Urliste, mit welcher quantitativen Wahrscheinlichkeitsangabe diese Hypothese gegebenenfalls abgelehnt werden muss. Das Ergebnis der Berechnung hängt von den Daten der Urliste ab, es könnte sein, dass wir sagen müssen: „Die Hypothese ist falsch, sie wird abgelehnt. Präparat A führt mit 95%iger Wahrscheinlichkeit zu einer längeren Narkose als Präparat B“. (Warum die Hypothese negierend aufgestellt wird, das können wir hier in Kürze nicht erklären und verweisen auf Kapitel 14 in dem wie ausführlich darauf eingehen werden.) Bei der Prüfung von Medikamenten genügt es nicht zu sagen, Präparat A wirkt sehr wahrscheinlich besser als Präparat B. Wir benötigen zur Beurteilung der Wirkung Zahlen. Ob eine 95%ige Wahrscheinlichkeit ausreicht oder ob wir die Entscheidung mit 99,9%iger Wahrscheinlichkeit absichern wollen, darüber reden wir in einem späteren Kapitel.

6.3 Der Ursprung der Wahrscheinlichkeitsrechnung

Bevor wir uns näher mit der Frage beschäftigen, was der Begriff Wahrscheinlichkeit bedeutet und wie wir damit rechnen können, wollen wir einen kurzen Blick in die Geschichte der Wahrscheinlichkeitstheorie werfen. Ohne näher darauf einzugehen, führen wir hier einige Namen an, die mit diesem Gebiet der Mathematik eng verbunden sind. Oft werden die Anfänge in das 16.Jh. gelegt, als der italienische Arzt **Gerolamo Cardano** 1524 seine Erkenntnisse über Gewinnchancen bei Glücksspiel in „Das Buch vom Würfelspiel“ niederschrieb. Er behielt sein Wissen vor der Veröffentlichung relativ lange für sich und war damit seinen Glücksspielkameraden gegenüber beim Spiel im Vorteil. Es wird berichtet, dass er mit den auf dieser Grundlage erzielten Gewinnen Teile seines Studiums finanziert haben soll. Als Geburtsstunde der Wahrscheinlichkeitsrechnung wird aber meist das Jahr 1654 genannt, als die französischen Mathematiker **Blaise Pascal** und **Pierre de Fermat** sich auf Fragen eines routinierten Glücksspielers des **Chevalier de Méré Antoine Gombaud** zu Gewinnchancen bei Würfelspiel ebenfalls mit diesem Thema beschäftigten. Sie entwickelten die Grundlagen der klassischen Wahrscheinlichkeitsrechnung. In der Folgezeit leisteten viele Mathematiker wie Jakob Bernoulli (1713 „Die Kunst des Vermutens“), Bayes, Euler, Poisson, Gauß (1795 „Methode der kleinsten Quadrate“) und Laplace (1812, „Mathematische Wahrscheinlichkeitstheorie“), im Zusammenhang mit Glücksspielen, astronomischen und anderen Messungen wichtige Beiträge zum Thema. Die heute als etabliert geltenden axiomatischen Grundlagen der Wahrscheinlichkeitsrechnung entwickelte 1933 der russische Mathematiker Andrei N. Kolmogoroff.

6.4 Begriffserklärungen und Notationen

In diesem Kapitel werden wir nicht viel rechnen aber wir müssen ein paar Dinge formalistisch darstellen, wozu es in der Literatur leider unterschiedliche Notationen gibt. Daher stellen wir hier einige Begriffserklärungen und die in diesem Kapitel benutzten Notationen vor. Der Zusammenhang der einzelnen Notationen kann an dieser Stelle u.U. noch nicht erkannt werden, er wird aber im Laufe des Textes klar.

Zufallsexperimente
Dies sind z.B. Glücksspiele mit Würfeln, Karten oder Münzen, die wir auch zur Erklärung der Voraussetzungen heranziehen, die für Zufallsexperimente gelten.

Notwendige Eigenschaften von Zufallsexperimenten

1. Zufallsexperimente finden unter festgelegten Bedingungen statt, die immer die gleichen bleiben müssen. *Der Würfel muss fair sein. Das bedeutet, jede Seite muss die gleiche Chance haben, nach einem Wurf oben zu liegen. Genau genommen gibt es das gar nicht, da jeder technische Herstellungsvorgang des Würfels prinzipiell fehlerbehaftet ist.*

2. Bei einem Zufallsexperiment müssen mindestens zwei unterschiedliche Ergebnisse realisierbar sein. *Bei einem Münzwurf sind es 2, beim sechsseitigen Würfel 6.*

3. Alle realisierbaren Ereignisse des Experiments stellen den sogenannten Ergebnisraum Ω dar und müssen bekannt sein. *Beim sechsseitigen Würfel sind die Ergebnisse 1, 2, 3, 4, 5, 6*

4. Der Ergebnisraum Ω ist endlich. *Beim Würfel haben wir 6 mögliche Ergebnisse.*

5. Jedes Ergebnis muss die gleiche Chance haben, realisiert werden zu können (Gleichmöglichkeit). Siehe Punkt 1. *Intuitiv können wir sagen, dass bei einem fairen Würfel keine der Zahlen einen Vorteil hat, oben zu liegen.*

6. Es ist prinzipiell nicht vorhersagbar, welches Ereigniss bei einem bestimmten Wurf realisiert wird. *Diese Forderung ist unmittelbar einsichtig.*

7. Das Experiment muss beliebig oft wiederholbar sein. *Wir können, wenn wir genügend Zeit haben, „beliebig" oft würfeln. Wir müssen es nicht, aber es muss prinzipiell möglich sein.*

Nur wenn diese Eigenschaften gegeben sind, sprechen wir von einem Zufallsexperiment. Zufallsexperimente unterscheiden sich zu.B. Von physikalischen Experimenten dadurch, dass der Ausgang letzterer auf Grund von Gesetzen kausal determiniert ist. Wenn ich einen Stein aus einem Meter Höhe fallen lasse, dann kann ich – nach vorhergehender Berechnung über das Gravitationsgesetzt – mit 100%iger Wahrscheinlichkeit vorhersagen, wann er auf dem Boden aufschlägt, nämlich nach 0,452 s. (Lassen wir dem Seitenwind und Ungenauigkeit bei der Zeitmessung mal unberücksichtigt.) Wenn ich einen Würfel werfe, dann kann ich grundsätzlich nicht vorhersagen, welche Zahl oben liegen wird.

Venn-Diagramme

Wenn wir den Ausgang eines Wurfes (4 liegt oben) in einem Venn-Diagramm (Abb. 1) darstellen, dann sind folgende Begriffe gegeneinander abzugrenzen.

1. Das Liegenbleiben des Würfels nennen wir ein **Ereignis** (E oder F)
2. Das Ereignis (E) enthält das **Ergebnis** (4).
3. Da das Ereignis nur ein Ergebnis enthält, nennen wir es ein **elementares Ereignis**.
4. Es gibt sechs verschiedene Möglichkeiten für das „Obenliegen" einer Ziffer. Die Gesamtheit aller realisierbaren Ergebnisse eines Experiments nennen wir den **Ergebnisraum Ω**. Er enthält alle möglichen Ergebnisse des Experiments (1, 2, 3, 4, 5, 6).
5. Wenn wir zwei Würfel werfen, dann hat das folgende Ereignis (F) zwei Ergebnisse (5 und 6). Dieses Ereignis F nennen wir ein **zusammengesetztes Ereignis.**

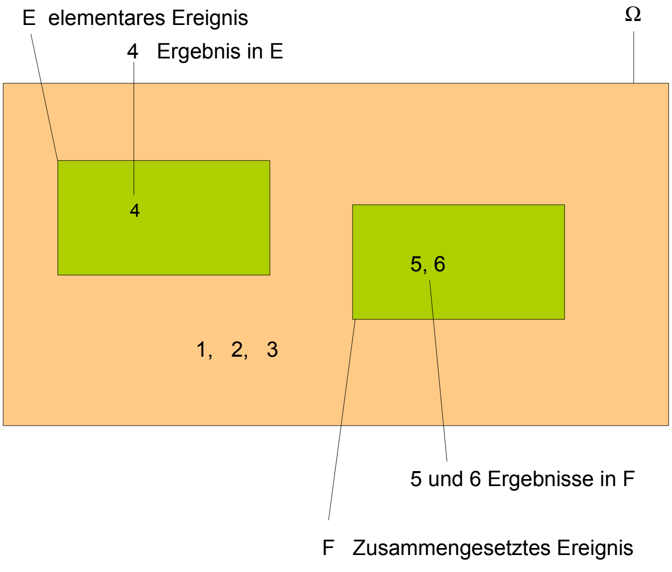

Abb.1

P, p, W	Wahrscheinlichkeit (lat. probabilitas)
Iq, W	Irrtumswahrscheinlichkeit
E	Eintritt eines Ereignisses, es können auch andere lateinische Großbuchstaben verwendet werden. Ereignis E ist eine Teilmenge des Ergebnisraumes Ω; $E \subseteq \Omega$
P(E)	Wahrscheinlichkeit für den Eintritt des Ereignisses E
Ω	Omega, Ergebnisraum, der alle möglichen Ereignisse enthält, er muss für jedes Experiment festgelegt werden.

$\lvert\Omega\rvert$	Anzahl aller möglichen Ergebnisse in Ω = n
n	Umfang der untersuchten Stichprobe, Anzahl der möglichen Ergebnisse die in einem Experiment realisiert werden können.
ν	das kleine griechische n gesprochen nü, Umfang der Stichprobe
H(E)	Absolute Häufigkeit von E
h(E)	relative Häufigkeit von E
$\subseteq$	Teilmenge von, wenn A = gerade Zahlen <10 und B = 6, dann gilt B $\subseteq$ A
$\neg$	Negationszeichen**,** $\neg$E wird gelesen: nicht E, $\neg$E ist das Komplement von E.
$\cap$	„und“ im Sinne von „sowohl als auch“, Schnittmenge E$\cap$F (gesprochen E geschnitten F) = Ergebnisse, die sowohl zu E wie auch zu F gehören. Wenn E = Zahlen <7 und F = Zahlen <3, dann gilt E$\cap$F = 1 und 2
$\cup$	„ oder“ im Sinne von „oder/und“ (nicht als entweder oder), nicht ausschließendes „oder“ Vereinigungsmenge E$\cup$F (gesprochen E vereinigt F) = Ergebnisse, die in E oder F oder beiden vorkommen. Wenn E = Zahlen <7 und B = Zahlen 10 bis 15, dann gilt E $\cup$ F = 1, 2, 3, 4, 5, 6, 7, 10, 11, 12, 13, 14, 15
$\emptyset$	unmögliches Ereignis, leere Menge, wenn A = Fallen einer 7 bei einem sechsseitige Würfel ist, dann ist das ein unmögliches Ereignis.
[1;6]	geschlossenes Intervall von 1 bis 6 einschließlich der beiden Grenzwerte
]1;6[	offenes Intervall von 1 bis 6 ausschließlich der beiden Grenzwerte

Tabelle 1

6.5 Wahrscheinlichkeit

Der Brockhaus schreibt: „Wahrscheinlichkeit ist ein Maß für den Grad der Möglichkeit noch nicht eingetretener Ereignisse,“ und bei Jacob Bernoulli lesen wir „Die Wahrscheinlichkeit ist nämlich ein Grad der Gewißheit ...“ Wie wir sehen werden, beschäftigt sich die Wahrscheinlichkeitsrechnung mit den Gesetzmäßigkeiten des Auftretens von Ereignissen, bei denen der Zufall eine Rolle spielt.

Im Laufe der Zeit haben sich unter verschiedenen Perspektiven diverse Theorien zur Wahrscheinlichkeitsrechnung entwickelt, von denen wir hier drei gegeneinander abgrenzen wollen.

Die empirische Wahrscheinlichkeit
Die klassische Wahrscheinlichkeit
Die axiomatische Wahrscheinlichkeit

Auf andere „Formen „ der Wahrscheinlichkeit gehen wir hier nicht ein.

6.5.1 Die empirische Wahrscheinlichkeit

(statistische Wahrscheinlichkeit, a-posteriori-Wahrscheinlichkeit (a-posteriori bedeutet: im Nachhinein, z.B. nach einem Experiment berechenbar.)

Führen wir noch einmal die beiden Sätze A und B aus dem Text weiter oben vor Augen, die Sache mit den Kopfschmerzen und die mit dem Wetter. Je öfter ich solche Situationen erlebt habe, um so größer wird die Sicherheit (Wahrscheinlichkeit), mit der ich Aussagen über zukünftiges Eintreffen machen kann. Um mit solchen qualitativen Wahrscheinlichkeitsangaben („Symptome gehen wahrscheinlich zurück“ und „sehr wahrscheinlich ein Gewitter“) rechnen zu können, müssen wir sie quantifizieren, ihnen also Zahlen zuweisen. Wie wir das machen, zeigt das Gedankenexperiment in Beispiel 1.

Beispiel 1
Wir wollen wissen, wie groß die Wahrscheinlichkeit dafür ist, dass eine beliebige Person aus der Population X die Blutgruppe A hat. Oder anders gefragt, wie viel % der Population die Gruppe A hat. Dazu untersuchen wir eine Zufallsstichprobe von 10 Personen der Population und finden bei zwei Personen die Gruppe A. Da

wir den wahren Wert nicht kennen, zweifeln wir das Ergebnis (2 von 10 = 20%) nicht an. Aber wir fragen uns, ob der Stichprobenumfang groß genug war, um eine "sichere" Antwort zu erhalten. In unserem Gedankenexperiment verfahren wir nun im Hinblick auf die Erklärung des Begriffs Wahrscheinlichkeit wie folgt. Wir untersuchen neun weitere Zufallsstichproben mit je n = 10. Zusammen haben wir damit 100 Personen untersucht. Die Zählergebnisse (H) für jede einzelne Stichprobe finden wir in folgender Tabelle.

Stichprobe	1	2	3	4	5	6	7	8	9	10
n	10	10	10	10	10	10	10	10	10	10
H	2	10	4	3	7	8	1	0	4	4

n = Stichprobenumfang, H = absolute Häufigkeit für Blutgruppe A

Tabelle 2

Wir werden nun die aufeinanderfolgenden Stichproben der Reihe nach zu Gruppen von 20, 30, 40, … 100 zusammenfassen (kumulieren) und erhalten die folgende Darstellung, in der die relative Häufigkeit (h) für das Auftreten von Blutgruppe A nach h = H/n berechnet ist.

	Absolute Häufigkeit H	Stichprobenumfang	relative Häufigkeit h
	Blutgruppe A	n kumuliert	h = H/n
Stichprobe 1	2	10	0.200
kumuliert bis Stpr. 2	12	20	0.600
kumuliert bis Stpr. 3	16	30	0.533
kumuliert bis Stpr. 4	19	40	0.475
kumuliert bis Stpr. 5	26	50	0.520
kumuliert bis Stpr. 6	34	60	0.567
kumuliert bis Stpr. 7	35	70	0.500
kumuliert bis Stpr. 8	35	80	0.438
kumuliert bis Stpr. 9	39	90	0.433
kumuliert bis Stpr. 10	43	100	0.430

Tabelle 3

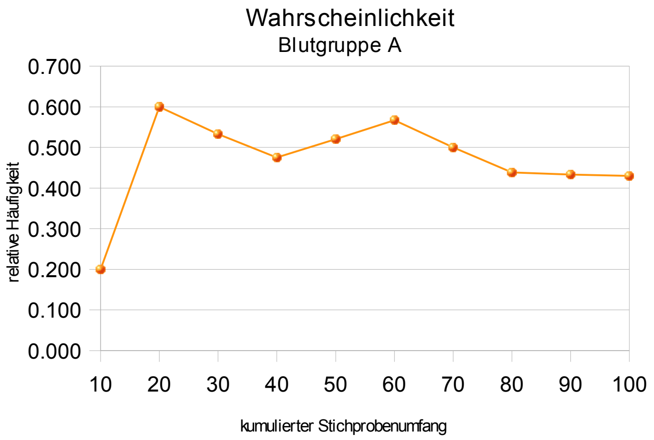

Abb. 2

Abb.2 zeigt die Entwicklung der relativen Häufigkeit bei zunehmendem n. Wir folgern daraus, dass die kumulierten relativen Häufigkeiten sich mit größer werdendem n einem unbekannten Endwert, dem Erwartungswert nähern. Dies wird durch die folgende Gleichung ausgedrückt, die wir lesen: Wenn n gegen unendlich strebt, sich der Quotient h/n einem Grenzwert P(E). Hier ist p(E) = 0,43 eine Näherung.

$$P(E) = \lim_{n \to \infty} H(E)/n$$

Die Wahrscheinlichkeit dafür, dass eine beliebige Person der Population X die Blutgruppe A hat, wird um so sicherer, je größer die Zahl der untersuchten Probanden ist. Das führt und direkt zur empirischen Wahrscheinlichkeit.

> Die Wahrscheinlichkeit für das Auftreten des Ereignisses Blutgruppe A ist gleich der relativen Häufigkeit für das Auftreten des Ereignisses bei hinreichend großem n.
>
> Die empirische Wahrscheinlichkeit entspricht also der relativen Häufigkeit.

Diese Wahrscheinlichkeitsdefinition eignet sich gut für Schätzungen von Erwartungswerten empirischer Daten. Eine wesentliche Erkenntnis ist, das ein Zusammenhang besteht zwischen Stichprobenumfang und der Sicherheit einer Aussage (siehe Kapitel 2). Diesen Zusammenhang hat der Schweizer Mathematiker Jacob Bernoulli mit dem "**Gesetz der großen Zahlen**" formuliert, welches 1713 posthum veröffentlicht wurde. Es entspricht in etwa der Aussage:

> Je mehr empirische Einzelwerte vorliegen, desto deutlicher tritt das Typische, welches der untersuchten Erscheinung in der Grundgesamtheit zugrunde liegt, hervor.

Nun können wir nach unserer empirischen Untersuchung (fiktiv)zwei Aussagen machen:

1. Eine beliebige Person der Population hat mit 43%iger Wahrscheinlichkeit die Blutgruppe A.
2. 43% der Population hat die Blutgruppe A.

Da wir ja immer von der Stichprobe auf die Population schließen, ist der gefundene Wahrscheinlichkeitswert natürlich nur eine Schätzung für den Erwartungswert. Eine Schätzung, die mit steigendem n besser wird. Zu bedenken ist immer, dass der Schluss ein Induktionsschluss ist und als solcher immer mit einer Unsicherheit behaftet. (siehe Kap.14)
Nach dieser Herleitung der Wahrscheinlichkeit als relative Häufigkeit nun zur formellen Darstellung der Berechnung der empirischen Wahrscheinlichkeit. Die Wahrscheinlichkeit ist also der Quotient aus der absoluten Häufigkeit von E dividiert durch den Stichprobenumfang n. In Beispiel 1 ist das die Häufigkeit von Blutgruppe A bezogen auf die Gesamtzahl der untersuchten Personen

also
$$P(E) = \frac{H(E)}{n} = \frac{43}{100} = 0{,}43$$

Im normalen Sprachgebrauch wird die so ermittelte, im Intervall [0;1] liegende Wahrscheinlichkeit mit 100 multipliziert damit wir eine Prozentangabe erhalten.

Die Definition der empirischen Wahrscheinlichkeit lautet

> Die Wahrscheinlichkeit für den Eintritt eines Ereignisses ist gleich der relativen äufigkeit des Ereignisses bei hinreichend großem n. P(E) = H/n

6.5.2 Die klassische Wahrscheinlichkeit

frequentitische Wahrscheinlichkeit, a-priori-Wahrscheinlichkeit. a priori bedeutet, sie kann vor dem Experiment berechnet werden.

Die klassische Wahrscheinlichkeit entwickelte sich vor einem anderen theoretischen Hintergrund und bedarf im Gegensatz zur empirischen Wahrscheinlichkeit keiner Empirie. Klassische Wahrscheinlichkeitsaussagen können intuitiv gemacht werden (a-priori-Wahrscheinlichkeit).

Typische Beispiele sind Glücksspiele

1. Wenn ich einen sechseitigen Würfel werfe, dann kann ich intuitiv voraussagen, dass mit einer Wahrscheinlichkeit von 1/6 = 16,7% die Ziffer 4 oben liegen wird.

2. Wenn ich eine beliebige Karte aus einem Skatkartenspiel entnehme, dann ist es sofort einleuchtend, dass es sich mit der Wahrscheinlichkeit 4/32 = 12,5% um ein Ass handelt.

3. Wenn ich beim Roulette meinen Einsatz auf noir platziere, dann weiß ich, dass meine Chance zu gewinnen gleich 18/37 = 48,65% ist.

Die klassische Wahrscheinlichkeit wurde, wie wir schon wissen, als ursprüngliches Wahrscheinlichkeitskonzept im Zusammenhang mit Glücksspiel entwickelt. Sie ist eng mit dem Namen Laplace verbunden der 1812 seine Erkenntnisse in „Mathematische Wahrscheinlichkeitstheorie" veröffentlichte. Danach sprechen wir von „Laplace-Wahrscheinlichkeiten", „Laplace-Experimenten" (Zufallsexperimenten) und „Laplace-Würfeln" (faire Würfel.

Wir wollen an Beispiel 2 in die formelle Berechnung von Laplace-Wahrscheinlichkeiten einführen.

Beispiel 2
Beim einmaligen Wurf eines zehnseitigen „Laplace-Würfels" ist für den Spielausgang eine „2" günstig. Da der Würfel die Ziffern 0, 1, 2, 3, 4, 5, 6, 7, 8 und 9 trägt, wird eines dieser zehn Ereignisse eintreten. Die Wahrscheinlichkeit, dass es eine 2 ist, beträgt, a priori 1/10 der Gesamtmöglichkeiten, also 1:10 = 0,1. Wir stellen nun den a priori gefundenen Wert durch eine Rechnung dar und legen fest:

Ω = {0,1,2,3,4,5,6,7,8,9} = <u>mögliche Ereignisse</u>
$|\Omega|$ = 10
E = Ereignis Fallen einer „2", für den Spielausgang <u>günstiges Ergebnis</u>.
H = 1, denn es gibt nur auf einer Würfelseite eine „2".

Nun gilt

$$P(E) = \frac{H(E)}{n} = \frac{1}{10} = 0{,}1$$

$$P(E) = \frac{|E|}{|\Omega|} = \frac{1}{10} = 0{,}1$$

$$P(E) = \frac{\text{günstige Fälle}}{\text{mögliche Fälle}} = \frac{1}{10} = 0{,}1$$

Welche Gleichung angewendet wird ist unerheblich. Mit der Wahrscheinlichkeit 0,1 = 10% wird beim nächsten Wurf eine 2 realisiert. Ob der konkrete Wurf dann wirklich eine 2 bringt, das wissen wir nicht, das hängt vom Zufall ab. Auf <u>lange</u> Sicht (Gesetz der großen Zahlen) wird es aber so sein, dass von 100000 Würfen ca. 10000 mal die 2 gewürfelt wird und natürlich auch jede der anderen Ziffern: 0, 1, 3, 4, 5, 6, 7, 8, 9 des Würfels ebenso oft. (Vielleicht würden wir für dieses Ergebnis auch mehr als 100000 Würfe benötigen.) Wir erinnern uns in diesem Zusammenhang sicher an die Gleichverteilung und daran, dass auch bei den a-posteriori-Wahrscheinlichkeiten das Gesetzt der großen Zahlen eine Rolle spielte. Die Berechnungh der klassischen Wahrscheinlichkeit führt zum gleichen Ergebnis wie die der empirischen Wahrscheinlichkeit.

Wir müssen jedoch beachten: Voraussagen im Sinne der Laplace-Wahrscheinlichkeit können aber nur gemacht werden, wenn die folgende beiden Voraussetzungen gegeben sind:

1. Jedes mögliche Ereignis muss a-priori die gleiche Chance haben.
2. Der Ergebnisraum, d.h. die Anzahl der möglichen Ereignisse muss endlich sein.

Und damit kommt für die Praxis ein Problem. Diese beiden Einschränkungen führen dazu, dass Laplace-Wahrscheinlichkeiten nur bedingt einsetzbar sind, da die Praxis diese beiden Voraussetzungen, abgesehen von Glücksspielen, nirgendwo bietet. Wenn wir 10 Mäuse mit einem Medikament behandeln welches die Entwicklung der Blutzellen beeinflußt, dann haben nicht alle 10 Tiere die gleichen Chancen identisch darauf zu reagieren. Der Stoffwechsel der einzelnen Tiere unterliegt der interindividuellen biologischen Variation. Wir wählen die Tiere zwar at random, aber das führt nicht in diesem Sinne zu gleichen Chancen. Und wir können den Versuch nicht beliebig oft wiederholen (Ethik, Zeit, Geld). Zuletzt ist es auch so, dass der Ergebnisraum nicht begrenzt ist. Bedingt durch die biologische Variabilität können unendlich viele Ergebnisse (reelle Zahlen) bei der Messung des Durchmessers von Lymphozyten auftreten.

Definition der klassischen Wahrscheinlichkeit:

Die **Wahrscheinlichkeit** für das Eintreten des Ereignisses E ist gleich dem Quotienten aus der Zahl der günstigen Ereignisse und der Zahl der möglichen (gleich**wahrscheinlichen**) Ereignisse.

P(E) = günstige Fälle/mögliche Fälle

Bei diese Definition der Wahrscheinlichkeit ist kritisch zubemerken, dass der Wahrscheinlichkeitsbegriff durch eine Formulierung definiert wird, die den Begriff Wahrscheinlichkeit enthält. Eine solche Zirkeldefinition ist an sich unbefriedigend.

Zusammenfassung der empirischen und klassischen Wahrscheinlichkeit.

	Die Wahrscheinlichkeit für den Eintritt eines Ereignisses ist
Empirische Wahrscheinlichkeit	die relative Häufigkeit, mit der das Ereignis bei hinreichend großem n eintritt. **P(E) = H/n**
Klassische Wahrscheinlichkeit	der Quotient aus der Anzahl der günstigen Fälle und der Anzahl der möglichen Fälle. **P(E) = günstige Fälle/mögliche Fälle**

6.5.3 Axiombasierte Wahrscheinlichkeit

Da die Definitionen der empirischen und klassischen Wahrscheinlichkeit keine wie in der Mathematik übliche axiomatische Grundlage hatten, wollte in den 30er Jahren des 20 Jh. u.a. der russische Mathematiker A.Kolmogoroff die bekannten Wahrscheinlichkeitstheorien in die „allgemeine Begriffsbildung der Mathematik“ einbinden. Dazu formulierte er auf der Basis der Mengenlehre ein Axiomsystem dessen erste 5 Axiome wir aus seiner Schrift “Grundbegriffe der Wahrscheinlichkeitsrechnung“ zitieren.

“1. Axiom: F ist ein Mengenkörper.
2. Axiom: F enthält die Menge E.
3 .Axiom: Jeder Menge A in F ist eine nicht negative reelle Zahl P(A) zugeordnet. Diese Zahl P(A) nennt man die Wahrscheinlichkeit des Ereignisses A.
4. Axiom. $P(E) = 1$
5. Axiom: Wenn A und B disjunkt sind, so gilt: $P(A+B) = P(A) + P(B)$”

P(A+B) schreiben wir heute $P(A \cup B)$

Das 6. Axiom, auf welches wir hier nicht eingehen, bezieht sich auf nichtendliche Ereignisse. Dadurch ist nach Kolmogoroff sichergestellt, dass die folgenden Berechnungen auch auf nicht endliche Ereignisse anwendbar sind. Bei den folgenden Erklärungen werden wir wieder auf sechseitige Würfel, Münzen und ein 32er Kartenspiel zurückgreifen.

Aus den Axiomen lassen sich eine Vielzahl von Regeln zum Rechnen mit Wahrscheinlichkeiten ableiten. Auf einige dieser Regeln werden wir hier eingehen. Sie entsprechen in der rechnerischen Anwendung dem, was wir von der empirischen und klassischen Wahrscheinlichkeit her schon kennennur sind sie nun auf wine axiomatische Basis gestellt und damit mathematisch abgesichert.

Regeln zum Rechnen mit Wahrscheinlichkeiten

A: Die Wahrscheinlichkeit ist eine nicht negative reelle Zahl $P(E) \geq 0$

B: Das sichere Ereignis $P(\Omega) = 1$

Würfel: Wie groß ist die Wahrscheinlichkeit dafür mit einem Würfel eine beliebige Zahl zu würfeln?
$\Omega = \{1,2,3,4,5,6\}$
$E = \Omega$
$|\Omega| = |E| = 6$
$P(E) = |E|/|\Omega| =$ 1,2,3,4,5,6) $= 6/6 = 1$

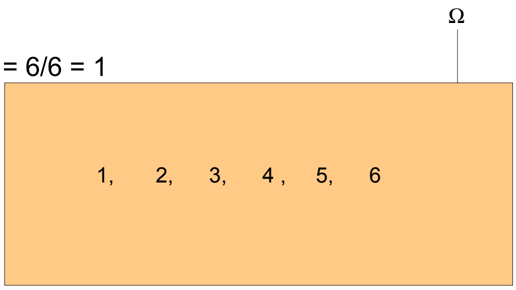

Abb.3

Mit der Wahrscheinlichkeit 1 werden wir eine der Zahlen aus dem Intervall [1;6] erwürfeln. Siehe Punkt E.

C: Das unmögliche Ereignis $P(E) = 0$

Würfel: Wie groß ist die Wahrscheinlichkeit, mit einem Würfel eine 7 zu werfen?
$\Omega = \{1,2,3,4,5,6\}$
$E = \{7\}$
$|\Omega| = 6$
$|E| = 0$
$P(E) = |E|/|\Omega| = 0/6 = 0$

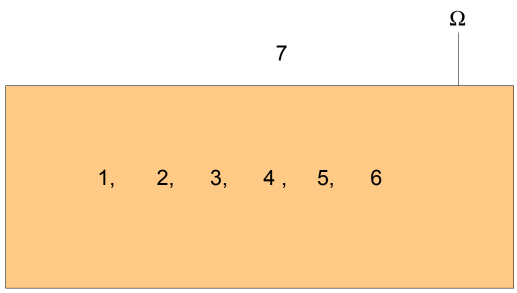

Abb.4

Die Wahrscheinlichkeit für das Werfen einer 7 ist Null, da die 7 kein Ergebnis des Ergebnisraumes Ω ist.

D: Die Wahscheinlichkeit ist ein Wert zwischen 0 und 1 $0 \leq P(E) \leq 1$

Das bedeutet, für jedes Ereignis (E) gibt es bezüglich seines Auftretens eine Wahrscheinlichkeit (P). P kann alle nicht negativen reellen Zahlen im Intervall [0;1] einnehmen.

Würfel: Wie groß ist die Wahrscheinlichkeit dafür eine 2 zu würfeln?

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

$|\Omega| = 6$

$E = \{2\}$

$|E| = 1$

$P(2) = |E|/|\Omega| = 1/6 = 0{,}167$

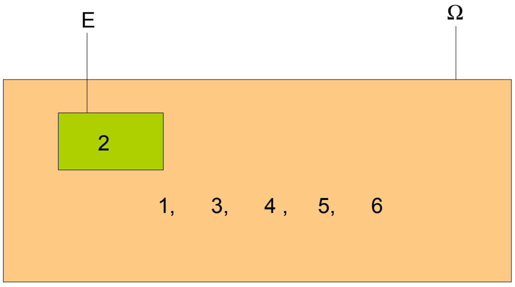

Abb.5

E Spezieller Additionsatz für Wahrscheinlichkeiten: $P(E \cup F) = P(E) + P(F)$
Oder-Verknüpfung für disjukte Ereignisse

Würfel: Wie groß ist die Wahrscheinlichkeit dafür, daß beim Wurf eines Würfels eine 2 <u>oder</u> eine 3 realisiert wird? Beide Ereignisse können nicht gleichzeitig eintreten, sie sind ja disjunkt.

$P(E \cup F) = P(E) + P(F)$

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

$|\Omega| = 6$

$E = \{2\}$

$F = \{3\}$

$P(E) = P(2) = 1/6$

$P(F) = P(3) = 1/6$

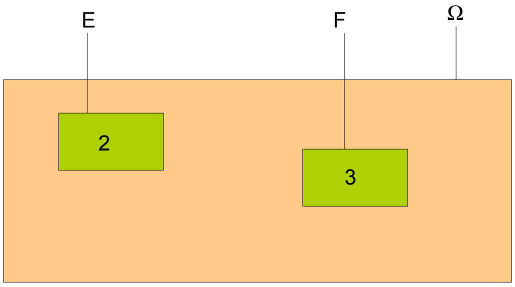

Abb.6

$P(2 \cup 3) = P(2) + P(3) = 1/6 + 1/6 = 2/6 = 0{,}333$

Wir können mit der Wahrscheinlichkeit 0,333 vorhersagen, dass in folgenden Wurf entweder eine 2 oder 3 realisiert wird.

Das bedeutet, die Wahrscheinlichkeit für das alternative Auftreten zweier **disjunkter** (sich gegenseitig ausschließender) Ereignisse E und F ist gleich der Summe der beiden Einzelwahrscheinlichkeiten.

Für beliebig viele Ereignisse gilt $P(E1 \cup E_2 \ldots En) = \sum_{i=1}^{n} P(E_i)$

Hieraus folgt

$P(1 \cup 2 \cup 3 \cup 4 \cup 5 \cup 6) = 1/6 + 1/6 + 1/6 + 1/6 + 1/6 + 1/6 = 6/6 = 1$

also ein sicheres Ereignis. Sie Punkt B

Zweites Beispiel hierzu

Eine Münze wird geworfen. Wie groß ist die Wahrscheinlichkeit dafür, dass A (Adler) oder Z (Zahl) oben liegt?

$P(A \cup Z) = P(A) + P(Z) = 1/2 + ½ = 1$

Eines der beiden Ereignisse wir mit P=1, also mit Sicherheit, eintreten.

F Allgemeiner Additionsatz für Wahrscheinlichkeiten $P(E \cup F) = P(E) + P(F) - P(E \cap F)$
Oder-Verknüpfung für nicht disjunkte Ereignisse

Das bedeutet, die Wahrscheinlichkeit für das <u>alternative</u> Auftreten zweier Ereignisse E und

F, die beide gemeinsame Ergebnisse enthalten, **die sich gegenseitig also nicht ausschließen**, ist gleich der Summe der beiden Einzelwahrscheinlichkeiten vermindert um die Wahrscheinlichkeit für die Schnittmenge von E und F, also für die in beiden Ereignissen vorkommenden Ergebnisse.

Würfel: Wie groß ist die Wahrscheinlichkeit dafür, dass bei einem Wurf mit einem Würfel eine der Zahlen x<3 oder eine der Zahlen 4>x>1 realisiert wird?

$\Omega = \{1,2,3,4,5,6\}$
$|\Omega| = 6$
E ={x<3 = 1, 2}
F = {4>x>1 = 2, 3}

P(E) = P(1, 2) = 2/6

P(F) = P(2, 3) = 2/6

$P(E \cap F) = 1/6$

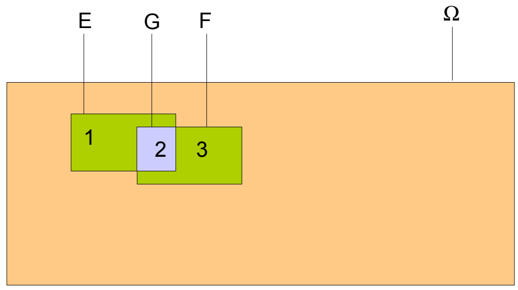

Abb.7

Würden wir P(E) und P(F) addieren, so würden wir das Ergebnis, welches in E und in F vorkommt, also die 2, doppelt zählen. Daher muss die Wahrscheinlichkeit für das beiden Ereignissen gemeinsame Ergebnis (2), (= die Schnittmenge von E und F) subtrahiert werden.

Also folgt

$P(E \cup F) = P(E) + P(F) - P(E \cap \mathbf{F})$
$P(2 \cup 3) = P(1, 2) + P(2, 3) = 2/6 + 2/6 - P(E \cap \mathbf{F})$
$P(E \cap F) = 1/6$ (die 2 ist eine der 6 Ziffern)
$P(2 \cup 3) = P(1, 2) + P(2, 3) = 2/6 + 2/6 - 1/6 = 3/6 = 1/2$
$P(2 \cup 3) = 20/36 = 0,5$
Hier $P(E \cap F)$ nicht durch Multiplikation berechnen. Das nur bei disjukten Ereignissen.

Es ist intuitiv einsichtig, dass die drei günstigen Ziffern, 50% der möglichen Ziffern ausmachen.

G: Multiplikation von Wahrscheinlichkeiten Und-Verknüpfung bei disjunkten Ereignissen	$P(E \cap F) = P(E) * P(F)$

Würfel: Wie groß ist die Wahrscheinlichkeit, dass wir bei zwei nacheinander auszuführenden Würfen mit einem Würfel eine 2 und eine 3 in beliebiger Reihenfolge erhalten? Die beiden Ereignisse sind voneinander unabhängig, da das Ereigniss des 1. Wurfes keinen Einfluß darauf hat, welche Ziffer beim 2. Wurf fällt. Der Würfel hat kein Gedächtnis!

E = {2}
F = {3}
$\Omega = \{1,2,3,4,5,6\}$
$|\Omega| = 6$
$P(E \cap F) = P(E) * P(F)$
P(E) = 1/6
P(F) = 1/6
$P(E \cap F) = 1/6 * 1/6 = 1/36 = 0,0277$

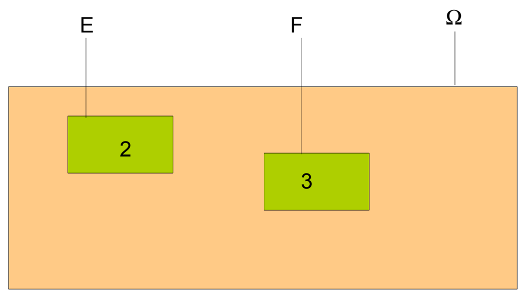

Abb.8

Wir können also mit der Wahrscheinlichkeit 0,0277 vorhersagen, dass bei den folgenden 2 Würfen eine 2 und eine 3 realisiert wird.

Das bedeutet, die Wahrscheinlichkeit für das gleichzeitige Auftreten voneinander unabhängiger, diskunkter Ereignisse ist gleich dem Produkt der Einzelwahrscheinlichkeiten.

H: Komplementäres Ereigniss, Komplement, Negation P(¬E) + P(E) = 1

Würfel:

erwürfelt: 2
nicht erwürfelt:1,3,4,5,6
Ω = {1,2,3,4,5,6}
|Ω| = 6
E = {2}
¬E = {1,3,4,5,6}

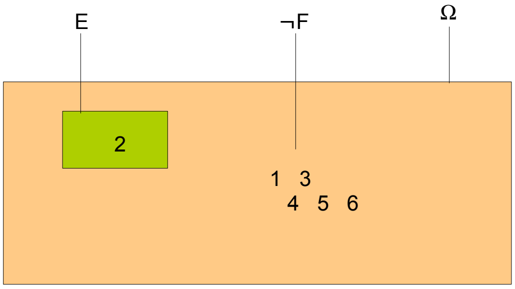

Abb.9

1, 3, 4, 5, 6 sind die zu 2 komplementären Ereignisse.
Wenn E bedeutet "3 liegt oben",
dann bedeutet ¬E "2 liegt nicht oben"
oder, was dem entspricht "¬2 liegt oben"
und das entspricht E = „1 oder 3 oder 4 oder 5 oder 6" liegen oben.

Die Summe der Wahrscheinlichkeiten komplementärer Ereignisse ist immer 1.

Nach P(¬E) + P(E) = 1
ist die Wahrscheinlichkeit für die 2: P(2) = 1/6 = 0,167
und die Wahrscheinlichkeit für !,3,4,5,6: P(¬2)= 5/6 = 0,833

Das bedeutet, zu dem eingetretenen Würfelereignis "E = 2 liegt oben" gibt es die 5 nicht eingetretenen Ereignisse: 1 oder 3 oder 4 oder 5 oder 6 liegt oben". Die Gesamtheit der Ereignisse in Ω, die nicht eingetreten sind, aber hätten eintreten können, nennen wir die komplementären Ereignisse zu dem eingetretenen Ereignis oder dessen Komplement oder seine Negation.

6.6 Irrtumswahrscheinlichkeit

In der Statistik nennen wir die Wahrscheinlichkeit P(E), mit der wir eine Aussage machen können, auch **Aussagewahrscheinlichkeit** [W], die Negation P(¬E) dagegen **Irrtumswahrscheinlichkeit** [IW]. Der Begriff Irrtumswahrscheinlichkeit spielt eine wesentliche Rolle bei Hypothesenprüfungen.

Wir ziehen aus eine 32-Blatt-Kartenspiel eine Karte.

Die Wahrscheinlichkeit für "Karo-Dame“ ist **W(Karo-Dame) = 1/32 = 0,03125**
Die Wahrscheinlichkeit dafür, dass die Karte <u>nicht Karo-Dame</u> ist, ist **W(¬Karo-Dame) = 31=32 = 0,96875**

Die Wahrscheinlichkeit dafür dass etwas nicht eintritt nennen wir Irrtumswahrscheinlichkeit.

Die **Irrtums**wahrscheinlichkeit (IW) dafür die Karo-Dame gezogen zu haben ist also 0,96875

Es gilt also:

P(E)	+	P(¬E)	=	1
W	+	IW	=	1
0,03125	+	0,96875	=	1

Umrechnung der Wahrscheinlichkeit in Irrtumswahrscheinlichkeit

In der Inferenzstatistik wollen wir z.B. berechnen, ob eine Substanz eine Heilwirkung hat. Die Wahrscheinlichkeit mit der wir sagen wollen „Die Substanz hat eine Heilwirkung“ legen wir auf z.B. 0,95 fest.

Damit wir den Wert W = 0,95 in die Rechnung mit einbeziehen können, benötigen wir die sogenannte t-Tabelle, in der „Rechenfaktoren“ für bestimmte Wahrscheinlichkeiten stehen. Dieser t-Tabelle – auf die wir in

einem späteren Kapitel ausführlich eingehen - müssen wir für die Berechnung den Faktor für W = 0,95 entnehmen. In der t-Tabelle stehen die Wahrscheinlichkeitsangaben aber in Form der Irrtumswahrscheinlichkeiten IW. Dort steht in der ersten Zeile hinter 2P statt 0,95 (Wahrscheinlichkeit) der Wert 0,05 (Irrtumswahrscheinlichkeit).

Ausschnitt aus der t-Tabelle (F. Keller, Statistik für naturwissenschaftliche Berufe, pmi-Verlag Frankfurt/M)
Auf die Frage, warum oben links „2P“ steht, gehen wir später ein.

2 P	0,40	0,20	0,10	**0,05**	0,025	0,02	0,01	0,005	0,001
ν									
1	1,376	3,078	6,3138	12,706	25,452	31,821	63,657	127,32	636,598
2	1,061	1,886	2,9200	4,3027	6,2153	6,965	9,9248	14,089	31,600
3	0,978	1,638	2,3534	**3,1825**	4,1765	4,541	5,8409	7,4533	12,924

Tabelle 4

Wir müssen also zur Benutzung der Tabelle gedanklich umformen: W $\longrightarrow$ IW

0,95 $\longrightarrow$ **0,05**

Zu diesem Wert finden wir für ν = 3 (Stichprobenumfang) den t-Wert **3,1825.** Welche Bedeutung dieser Wert hat, das werden wir auch später kennenlernen. An dieser Stelle geht es nur darum, den Begriff Irrtumswahrscheinlichkeit einzuführen.

Wir suchen in der t-Tabelle den Faktor für IW = 0,05 und haben damit den Faktor für die Wahrscheinlichkeit 0,95.

W + IW = 1
IW = 1 - W
IW = 1 - 0,95
IW = 0,05

Solche Umrechnungen werden bei der Auswertung von statistischen Tests auftreten.

Dieser Abstecher in einige Aspekte der Wahrscheinlichkeitsrechnung war kein Selbstzweck. Er sollte zu einem Grundverständnis des Wahrscheinlichkeitsbegriffs führen.

6.7 Übungen

Übung 1. Wie groß ist die Wahrscheinlichkeit, daß beim Wurf eines sechseitigen Würfels keine der sechs Zahlen 1, 2, 3, 4, 5, 6 oben liegt ?

Übung 2. Wie groß ist die Wahrscheinlichkeit, daß beim Wurf eines sechseitigen Würfels eine der sechs Zahlen 1, 2, 3, 4, 5, 6 oben liegt ?

Übung 3. Sie haben ein Kartenspiel mit 32 Karten, aus dem nacheinander einzelne Karten gezogen werden. Die gezogenen Karten werden nicht zurückgelegt. Berechnen Sie die Wahrscheinlichkeit für das Ziehen von:

1. Zug Herz 10
2. Zug Eine beliebige Karte aber keinem Herz-König
3. Zug Herz-König
4. Zug Karo 10 oder Pik 9 oder Herz Ass oder einer beliebigen Kreuzkarte

Übung 4. Eine Urne enthält 40 rote und 60 grüne Kugeln.

a) Wie groß ist die Wahrscheinlichkeit, bei einem Zug eine rote Kugel zu ziehen? Die Kugel wird zurückgelegt.
b) Wie groß ist die Wahrscheinlichkeit, beim 1. Zug eine grüne Kugel und beim 2. Zug wieder eine grüne Kugel zu ziehen? Die 1. Kugel wird vor dem 2. Zug zurückgelegt.
c) Im folgenden Experiment werden gezogene Kugeln nicht zurückgelegt. Wie groß ist die Wahrscheinlichkeit, im 1. Zug eine grüne und im 2. Zug eine grüne und im 3. Zug eine rote Kugel zu ziehen ?

Übung 5. In einem menschlichen Bevölkerungsteil sind die 4 Blutgruppen des AB0-Systems wie folgt vertreten: Gruppe A = 42%; B = 13%; AB = 7% und 0 = 38%

a) Wie groß ist die Wahrscheinlichkeit, daß eine zufällig herausgegriffene Person die Gruppe A oder 0 hat ?
b) Wie groß ist die Wahrscheinlichkeit, daß bei zwei zufällig herausgegriffene Personen die Gruppen AB und 0 auftreten ?

Lösung 1: P(E) = günstige Fälle/mögliche Fälle = 0/6 = 0. Lösung 2: P(E) = günstige Fälle/mögliche Fälle = 6/6 = 1. Lösung 3: 1. Zug: P(Karodame) = 1/32; es gibt nur eine Karodame, n = 32. 2. Zug Es sind noch 31 Karten im Spiel. P(Nicht Herzkönig) = 30/31; es gibt bei n = 31 Karten 30 Alternativen zu Herzkönig. 3. Zug Es sind noch 30 Karten im Spiel P(Herzkönig) = 1/30; es gibt nur einen Herzkönig, n = 30. 4. Zug: Es sind noch 29 Karten im Spiel. $P(K10 \cup P9 \cup HA) = (1+1+1)/29 = 0{,}103$. Gündtige Fälle = 3, da unabhängige Ereignisse addiert werden.n=29. Lösung 4: a) P(R) = H/n = 40/100;H = 0,4, da 40 rote Kugeln vorhanden sind, n = 100, da insgesamt 100 Kugeln vorhanden sind. b) $P(R_1) = 60/100 = 0{,}6$. $P(R_2) = 60/100 = 0{,}6$, da das Ziehen der zweiten Kugel unabhängig von der ersten ist. c) $P(G \cap G \cap R) = P(G) * P(G) * P(R) = 60/100 *59/99 * 40/98 = 0{,}1459$. Lösung 5: a) $P(A \cup 0) = 42/100 + 38/100 = 80/100 = 0{,}8 =$ 80%; bei der Oder-Verknüpfung werden die Einzelwahrscheinlichkeiten addiert. b) $P(AB \cap 0) = 7/100 * 38/100 = 0{,}0266 = 2{,}66\%$; bei Und-Verknüpfungen werden die Einzelwahrscheinlichkeiten multipliziert.