Fachhochschule
Münster University of
Applied Sciences

Kombinatorik und Wahrscheinlichkeit

Prof. Dr. Martin Poppe

Aufgaben zur Einführung

1. Einfache Anwendung der Begriffe
Wie groß ist die Wahrscheinlichkeit, beim Skat Spiel mindestens einen Buben zugeteilt zu bekommen?

2. Ein Informatiker will mit Hilfe eines Zufallsgenerators komponieren. Er lässt einen Rechner beliebige Abfolgen von ¼ – Noten der C Dur-Tonleiter generieren. Wie groß ist die Wahrscheinlichkeit, dass er auf diese Wiese die 1. Zeile von „Freude schöner Götterfunken“ aus der 9. Symphonie van Beethovens findet (15 Noten)? Wie viele Versuche muss er unternehmen, um mit mindestens 50% Wahrscheinlichkeit die Tonfolge mindestens einmal zu erzeugen? Wie lange braucht er dazu, wenn pro Sekunde eine Melodie generiert wird?

3. Bisweilen werden die Begriffe „Wahrscheinlichkeit“ und „relative Häufigkeit“ synonym verwandt. Welche Begriffsdefinition liegt dem zugrunde?

4. Es wird mit 2 Würfeln gewürfelt. Welche Augen-Summe kommt öfter vor: die 9 oder die 10? (Gottfried Wilhelm Leibniz konnte einst das Ergebnis nicht glauben)

5. Noch immer gibt es Menschen, die Lotto spielen. Wie groß ist die Chance, 6 Richtige aus 49 zu finden?

6. Mit welcher Wahrscheinlichkeit würfelt man mit 3 Würfeln ein „Genie“, also zwei Einsen und eine Zwei?

Aufgaben zur Begriffsbildung

7. In der Elementarteilchen-Physik werden die Zerfälle sehr kurzlebiger Teilchen statistisch beschrieben. Beispielsweise zerfällt das so genannte „ρ^0–Meson“ in ca. 5 10^{-24} s entweder in ein „$\pi^+\pi^-$“ Paar (99,9% aller Zerfälle) oder ein „$\mu^+\mu^-$“ Paar (0,01%) , oder in ein Elektron-Positron Paar(0,09%). Ordnen Sie in diesem Beispiel die Begriffe Elementarereignis, Zufallsgröße, Ergebnismenge und Wahrscheinlichkeit zu. Kann man den Zerfall als „Laplace-Versuch“ bezeichnen?

8. Ein Flugzeug kann in der Luft fehlerfrei fliegen (99,99%), mit nicht fatalen Fehlern fliegen (0,0095%) oder abstürzen (0,0005%). Beschreiben Sie diese Möglichkeiten mit Hilfe der Begriffe des endlichen Wahrscheinlichkeitsraumes.

9. Die Wahrscheinlichkeit, dass mindestens eines der Ereignisse aus A oder B eintritt kann wie folgt berechnet werden: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
Bitte begründen sie diese Formel mit einer Skizze der Mengen A, B und Ω.

10. Geburtstags-Paradox
Im Folgenden wird angenommen, dass es keine Schaltjahre gibt und das jedes Datum für einen Geburtstag gleich wahrscheinlich ist.
Wie groß ist die Wahrscheinlichkeit, dass eine andere Person am gleichen Tag Geburtstag hat, wie ich selbst? Für dieses Problem benennen Sie bitte auch die Ergebnismenge, das Wahrscheinlichkeitsmaß, das Elementarereignis und den Wahrscheinlichkeitsraum.
Für welche Anzahl *n* von Personen wird die Wahrscheinlichkeit, dass mindestens eine andere Person am gleichen (beliebigen) Tag Geburtstag hat 50%? Wird die Wahrscheinlichkeit für irgend eine Anzahl = 1?

Aufgaben zur bedingten Wahrscheinlichkeit

11. Frauen in der Wissenschaft

An der Fachhochschule Münster sind 40% der Eingeschriebenen Studentinnen. Jeweils 25% der Studenten und 10% der Studentinnen studieren Wirtschaftswissenschaften.
Wie groß ist die Wahrscheinlichkeit, dass ein Studierender der Wirtschaftswissenschaften männlich ist? Wie groß ist die Wahrscheinlichkeit, dass ein Studierender, der nicht Wirtschaftswissenschaften studiert, eine Frau ist?

12. Screening Test
    Angenommen, es gibt eine Krankheit, die 0,2 Promille der Bevölkerung in sich tragen und es gibt einen Test auf diese Krankheit, der mit 99%-iger Wahrscheinlichkeit das korrekte Ergebnis liefert. Sie werden positiv getestet. Wie groß ist die Wahrscheinlichkeit, dass Sie wirklich krank sind?

13. Krebs-Behandlung oder nicht?
    Angenommen, 30% aller Männer erkranken gegen Ende Ihres Lebens an Krebs, jedoch sterben nur 3% der Erkrankten daran, denn er wächst sehr langsam. Wenn es eine Diagnose gibt, die mit 95%-iger Wahrscheinlichkeit korrekt ist. Wie gut muss dann die Behandlung sein, d.h. Wie hoch klein muss dann die Wahrscheinlichkeit *y* sein, dass die Behandlung selbst zum Tode führt, damit sich eine großflächige Untersuchung der Bevölkerung überhaupt auszahlt?

14. Spam Filter Statistik
    Man teilt die Mails in Ham und Spam ein. Dann schaut man die Wörter durch, die in den Ham Mails und in den Spam Mails vorkommen. Dann weiß man, dass z.B. in 30% aller Spam Mails das Wort *free* vorkommt und vielleicht in 1% aller Ham Mails.
    Wenn nun eine Mail ankommt, die das Wort Free enthält, mit welcher Wahrscheinlichkeit ist sie Spam?

Antworten

1. Am leichtesten ist es, die Wahrscheinlichkeit der Gegenhypothese „ich bekommen in 10 Versuchen keinen einzigen Buben“ zu berechnen. Dann gilt $P_{mindestens\,1\,Bube}=1-P_{kein\,Bube}$. Es gibt 32 Karten, 4 davon Buben. Also ist die Wahrscheinlichkeit, dass die erste Karte kein Bube ist $P_1=28/32$. Für die zweite Karte ist $P_2=27/31$, denn alle Buben sind noch da, aber bei einer Karte weniger. Für 10 Karten ist dann $P_{kein\,Bube}=P_1\cdot P_2\cdot\ldots P_{10}$, also $P_{kein\,Bube}=\dfrac{28\cdot 27\cdot 26\cdot\ldots 19}{32\cdot 31\cdot 30\cdot\ldots 23}\approx 0{,}203$ Man hat also nur in jedem fünften Spiel gar keinen Buben.
2. Für die erste Note gibt es 7 Möglichkeiten (c,d,e,f,h,a,h), ebenso für die zweite etc. Die Wahrscheinlichkeit, die Melodie durch Zufall zu entdecken ist also $P_{9.Synphonie}=(1/7)^{15}\approx 0{,}21\cdot 10^{-13}$. Und wie viele Versuche braucht man? Am besten, man betrachtet die Gegenhypothese: „Bei allen Versuchen ist die Symphonie nicht ein einziges Mal dabei“, denn sonst wird es wegen der Möglichkeit, die Melodie mehrmals zu finden kompliziert. Es ergibt sich für
   einen Versuch $P_{falsch}=1-(1/7)^{15}$,
   zwei Versuche $P_{falsch}=\left(1-(1/7)^{15}\right)^2$
   N Versuche $P_{falsch}=\left(1-(1/7)^{15}\right)^N$
   Wenn jetzt die Wahrscheinlichkeit 50% betragen soll, erhält man $\dfrac{1}{2}=\left(1-(1/7)^{15}\right)^N\Rightarrow N=\dfrac{ln(1/2)}{ln(1-(1/7)^{15})}$, zu berechnen mittels $ln(1-x)\approx -x$ als $N\approx 7^{15}\cdot ln(2)\approx 3{,}3\cdot 10^{12}$ (Das heißt: bei einem Versuch pro Sekunde wird es ca. 100 000 Jahre brauchen, bis die Melodie dabei ist)
3. Hier wird die so genannte „klassische Definition der Wahrscheinlichkeit“ (Huygens, Bernulli,Laplace): „Die Wahrscheinlichkeit eines günstigen Ereignisses ist die Anzahl der günstigen Ereignisse dividiert durch alle Ereignisse“.
4. Klar: Er dachte: Die 10 hat 2 Möglichkeiten 6+4 und 5+5. Die 9 hat auch zwei Möglichkeiten: 6+3 und 5+4. Er irrte, denn die Würfel sind unterscheidbar: Für die 10 gibt es die 3 Möglichkeiten 6+4 und 4+6 und 5+5. Für die 9 gibt es 4 Möglichkeiten: 6+3 und 3+6 und 5+4 und 4+5. daher kommt die 9 öfter vor als die 10.
5. Abzählen: erste Kugel: 49 Möglichkeiten, 6 davon richtig $P_1=6/49$ zweite Kugel: 48 Möglichkeiten, noch 5 richtig $P_2=5/48$ … sechste Kugel 44 Möglichkeiten, eine richtig $P_6=1/44$.
   Insgesamt ist die Chance, 6 Richtige zu bekommen also $P_{6\,Richtige}=P_1\cdot P_2\cdot\ldots\cdot P_6=\dfrac{6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}{49\cdot 48\cdot 47\cdot 46\cdot 45\cdot 44}=0{,}000\,007\,15\,\%$ (Zusatz: Die Anzahl der Versuche um eine Wahrscheinlichkeit von 50% zu erreichen ist $N=ln(2)/P_{6\,Richtige}\approx 9{,}7\cdot 10^6$. Bei einem Spiel pro Woche ergibt sich ein Zeitraum von 186 430 Jahren.)
6. Nennen wir (a,b,c) die Ergebnisse des Wurfs, so gibt es folgende drei Treffer: $(1,1,2),(1,2,1)$ und $(2,1,1)$, insgesamt jedoch 6^3 Möglichkeiten. So folgt $P_{Geni}=3/6^3=1/72\approx 1{,}4\,\%$.
7. Elementarereignis = Zerfall eines Mesons, Zufallsgröße: Identität der Zerfallsprodukte, Ergebnismenge = {„$\pi^+\pi^-$“ Paar, „$\mu^+\mu^-$“ Paar, „e^+e^-“ Paar}, Wahrscheinlichkeiten: 99,9%, 0,01% und 0,09%. Kein „Laplace-versuch“ , denn die Wahrscheinlichkeiten sind unterschiedlich.
8. Es ist hier Ω = {gut, defekt, ganz kaputt}, P = {gut, defekt, ganz kaputt, gut oder defekt, gut oder ganz kaputt, defekt oder ganz kaputt, gut oder defekt oder ganz kaputt}. Die Abbildung lautet P(gut) = 99,99%, P(defekt)=0,095%, P(ganz kaputt) = 0,0005%. Die (Elementar-)Ereignisse heißen „das Flugzeug fliegt fehlerfrei“, „das Flugzeug fliegt gerade noch“ und „Absturz“. {Ω,P} ist der Wahrscheinlichkeitsraum. Das sichere Ereignis heißt „das Flugzeug ist fehlerfrei, oder defekt oder ganz kaputt“, das unmögliche Ereignis existiert in diesem Falle nicht, da die Wahrscheinlichkeit 0 nicht auftritt .
9. Man muss die Wahrscheinlichkeiten aller in A oder in B enthaltenen Ereignisse summieren. Beginnt man mit allen von A, und summiert dann alle von B, so hat

man diejenigen, die in A und in B enthalten sind zweimal summiert. Sie müssen also einmal abgezogen werden.

10. Für jeden Tag ist die Wahrscheinlichkeit 1/365. Also auch für meinen Geburtstag. Die Ergebnismenge Ω ist die Menge aller möglichen Geburtstage. Das Elementarereignis ω_i ist „eine andere Person hat Geburtstag". Das Wahrscheinlichkeitsmaß, also die Wahrscheinlichkeitsverteilung lautet $P(\omega_i)=1/365$ für alle Elementarereignisse. Der Wahrscheinlichkeitsraum besteht aus Ω und $P(\omega_i)$.
Um die Wahrscheinlichkeit zu berechnen, dass eine andere Person genau meinen Geburtstag hat, legen wir diesen auf Tag 1 im Jahr und fragen zunächst nach der Wahrscheinlichkeit, mit der eine einzige Person diesen Tag trifft – also $P_{Treffer}=1/365$.
Dann überlegen wir allgemein, wie viele neue Treffer hinzu kommen, wenn die Personenzahl um 1 erhöht wird. Wenn wir mit $(a,b,c,..)$ die Geburtstage notieren, so dass zum Beispiel $(1,1,1)$ bedeutet, dass Es genau drei Personen gibt und dass diese beide am 1. Januar Geburtstag haben. Ein Treffer bedeutet, dass eine der Zahlen $a,b,c,..$ gleich eins ist.
Wenn nun die Zahl der Personen um 1 erhöht wird, werden aus allen alten Treffern 365 neue. Dazu kommen als neue Treffer alle N-Tupel von dem Typ $(a,b,d,..,1)$, Hiervon gibt es genau so viele neue, wie vorher Nicht-Treffer dabei waren. Nennen wir T_n die Anzahl der Treffer bei *n* Personen und N_n die Anzahl der Nicht-Treffer, so ergibt sich Personenzahl: $T_{n+1}=365\cdot T_n+N_n$. Es gilt: $T_1=1$ und $N_1=(365-1)=364$.
So erhält man sukzessive für zwei Personen $T_2=365\cdot 1+364$, $N_2=365^2-T_2=...=(365-1)^2$, für drei Personen $N_3=(365-1)^3$ und $T_3=1-N_3$ Insgesamt gibt es bei p Personen 365^n verschiedene N-Tupel, von denen $N_n=(365-1)^n$ der Aussage „Keine Person hat gleichzeitig mit mir Geburtstag „ entsprechen. Die Wahrscheinlichkeit, dass niemand am gleichen Tag Geburtstag feiert wie ich ist also für *n* Personen $P_{ohne\,mich}=\frac{364^n}{365^n}$. So folgt das Endergebnis:

$$P_{mit\,mir}=1-P_{ohne\,mich}=1-\left\{\frac{364}{365}\right\}^n \quad .$$

Man kann diese Gleichung umstellen, um herauszubekommen, wie viele Personen für eine bestimmte Wahrscheinlichkeit nötig sind: $p=\frac{ln(1-P_{mit\,mir})}{ln(364/365)}$.
Setzt man zum Beispiel eine Wahrscheinlichkeit von *50%* an, so werden dafür 253 Personen, also viel mehr als $365/2$ gebraucht. Ganz egal, wie groß *n* wird, die Wahrscheinlichkeit wird nie 1!

Um die Wahrscheinlichkeit zu berechnen, wie oft zwei Personen den gleichen (aber beliebigen) Geburtstag haben, macht man sich zunächst klar, dass es bei *n* Personen 365^n Geburtstagskombinationen gibt. Es ist nun einfacher, nach Unterschieden zu fragen, als nach Gleichheiten. Wir fragen also: „Für wie viele der 365^n Kombinationsmöglichkeiten sind alle Geburtstage verschieden?" Die erste Person kann 365 verschiedene Geburtstage haben, die zweite dann aber nur noch 364, die dritte nur noch 363 und so weiter. Die Gesamtwahrscheinlichkeit ist dann das Verhältnis $P_{alle\,verschieden}=\frac{365\cdot 364\cdot...\cdot(365-(n-1))}{365^n}$. Mindestens ein Doppelgeburtstag findet also mit einer Wahrscheinlichkeit $P_{Doppel}=1-P_{alle\,verschieden}$ statt. Man erhält beispielsweise $P_{Doppel}(n=23)=0{,}507$. Das ist das Paradoxe: schon bei viel weniger als 365/2 Personen ist die Chance eines Doppelgeburtstags 50%!
Für $n=366$ wird der Term $(365-(n-1))=0$. Bei 366 Personen kann man also sicher sein, das wenigstens zwei den gleichen Geburtstag haben – klar, wenn es mehr Leute als Tage gibt...

Vergleicht man die Ergebnisse für „meinen Geburtstag“ und „irgend ein Geburtstag“ , so kann man Depressionen bekommen.

11. Aus den Angaben folgt:
$P(\textit{Frau und Wirtschaft}) = 40\% \cdot 10\% = 4\%.$
$P(\textit{Mann und Wirtschaft}) = (1-40\%) \cdot 25\% = 15\%.$
Man kann diese Wahrscheinlichkeiten so lesen: Von 100 Studierenden sind 4 Frauen am Fachbereich Wirtschaft und 15 Männer am Fachbereich Wirtschaft.
Die Wahrscheinlichkeit, dass ein Wirtschaftsstudent weiblich ist, ist dann $P(\textit{Wirtschaftler ist eine Frau}) = 4\%/(4\%+15\%) = 4/19$. Oder das Gegenstück:
$P(\textit{Wirtschaftler ist ein Mann}) = (\textit{Mann} \mid \textit{Wirtschaftler}) = 15\%/(4\%+15\%) = 15/19$
Für die zweite Frage überlegen wir:
$P(\textit{Frau und nicht Wirtschaft}) = 40\% \cdot 90\% = 36\%.$
$P(\textit{Mann und nicht Wirtschaft}) = 60\% \cdot 75\% = 45\%.$
Die Wahrscheinlichkeit, dass ein Nicht-Wirtschaftler eine Frau ist somit $P(\textit{Nicht}-\textit{Wirtschaftler ist eine Frau}) = 36/(36+45) = 4/9$

12. Die Wahrscheinlichkeit krank zu sein ist $P(\textit{krank}) = 0{,}0002.$ Entsprechend ist $P(\textit{gesund}) = 0{,}9998.$ Nun gibt es insgesamt vier Möglichkeiten bzw. Gruppen von Personen: gesund und negativ getestet, gesund und positiv getestet, krank und positiv getestet, krank und negativ getestet. Deren Wahrscheinlichkeiten sind die jeweiligen Produkte:
$P(\textit{gesund}, \textit{negativ}) = 0{,}9998 \cdot 0{,}99 \approx 0{,}9898.$
$P(\textit{gesund}, \textit{positiv}) = 0{,}9998 \cdot 0{,}01 \approx 0{,}009998$
$P(\textit{krank}, \textit{negativ}) = 0{,}0002 \cdot 0{,}01 = 0{,}000002.$
$P(\textit{krank}, \textit{positiv}) = 0{,}0002 \cdot 0{,}99 = 0{,}000198$
damit ist die Chance, bei einem positiven Testergebnisses tatsächlich krank zu sein
$$\frac{P(\textit{krank}, \textit{positiv})}{P(\textit{gesund}, \textit{positiv}) + P(\textit{krank}, \textit{positiv})} = \frac{0{,}000198}{0{,}009998 + 0{,}000198}$$
Rechnet man dies aus, so kommt ein Wert von *0,0194,* also weniger als *2%* heraus. Man kann also trotz eines positiven Testergebnisses mit 98%-iger Sicherheit davon ausgehen, gesund zu sein!

13. Wir können am besten einen Entscheidungsbaum aufstellen:
alle

0,3 Krebs		0,7 kein Krebs	
0,95 positiv	0,05 negativ	0,05 positiv	0,95 negativ
(1-y) ok, y tot	0,97 ok, 0,03 tot	(1-y) ok y tot	ok

Insgesamt stirbt also ein Anteil $P_{tot} = (0{,}3 \cdot 0{,}95 + 0{,}7 \cdot 0{,}05) \cdot y + 0{,}3 \cdot 0{,}05 \cdot 0{,}03$
Dieser Anteil muss kleiner als $P_{\textit{kein Test}} = 0{,}3 \cdot 0{,}03 = 0{,}009$ sein. Nach *y* aufgelöst ergibt sich ein Grenzwert von 2,67%. Nur wenn mehr als 97% der Patienten die Behandlung verkraften, lohnt sich der Test überhaupt. Bleibt man unterhalb dieser Grenze, so wird letztlich viel Geld dafür ausgegebenen, dass die Menschen eher sterben...

14. Man teilt die Mails in Ham und Spam ein. Dann schaut man die Wörter durch, die in den Ham Mails und in den Spam Mails vorkommen. Dann weiß man, dass z.B. in 30% aller Spam Mails das Wort *free* vorkommt und vielleicht in 1% aller Ham Mails.
Hier die Komponenten:
alle 100%
h = Anteil Ham, (1-h) Anteil Spam

h * 1% = Ham,mit „free“,	h*99%=Ham, ohne „free“,
(1-h)*30%= Spam,mit „free“,	(1-h)*70%= Spam, ohne „free“

P(mit „free“)= h*1%+(1-h)*30%
Die Antwort hängt also davon ab, welcher Anteil der E-mails Spam ist.