

Grundlagen der Mathematik II

Lösungsvorschlag zum 8. Übungsblatt

Aufgabe 1. Als Ergebnisraum bietet sich $\Omega := \{1, 2, 3, 4, 5, 6\}$ an. Zum Ermitteln der Wahrscheinlichkeitsverteilung P auf Ω werten wir die Informationen aus, die über den Würfel gegeben sind:

- Daß die Ergebnisse 2, 3, 4, 5 alle gleich wahrscheinlich sind, bedeutet, daß es eine Zahl $t \in [0, 1]$ gibt mit $P(\{\omega\}) = t$ für $\omega = 2, 3, 4, 5$.
- Daß doppelt so oft eine 6 erscheint wie eine 1 (also die Wahrscheinlichkeit für eine 6 doppelt so groß ist wie die für eine 1) bedeutet, daß es eine Zahl $s \in [0, 1]$ gibt mit $P(\{1\}) = s$ und $P(\{6\}) = 2s$.
- Es ist

$$\begin{aligned}P(\{1, 6\}) &= P(\{1\}) + P(\{6\}) = s + 2s = 3s, \\P(\{2, 5\}) &= P(\{2\}) + P(\{5\}) = t + t = 2t, \\P(\{3, 4\}) &= P(\{3\}) + P(\{4\}) = t + t = 2t.\end{aligned}$$

Daß diese Wahrscheinlichkeiten alle gleich groß sind, bedeutet $3s = 2t$.

- Außerdem ist $1 = P(\Omega) = P(\{1\}) + \dots + P(\{6\}) = s + 4t + 2s = 3s + 4t$.

Damit haben wir zwei Gleichungen für die Zahlen s und t , die wir damit berechnen können: Wegen $3s = 2t$ ist $1 = 3s + 4t = 2t + 4t = 6t$, also $t = \frac{1}{6}$ und damit $s = \frac{2}{3}t = \frac{1}{9}$, so daß sich die folgende Wahrscheinlichkeitsverteilung P auf Ω ergibt:

| | | | | | | |
|-----------------|---------------|---------------|---------------|---------------|---------------|---------------|
| ω | 1 | 2 | 3 | 4 | 5 | 6 |
| $P(\{\omega\})$ | $\frac{1}{9}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{2}{9}$ |

Für das Ereignis

A : „Es wird eine ungerade Zahl gewürfelt“

gilt $A = \{1, 3, 5\}$ und damit

$$\begin{aligned}P(A) &= P(\{1, 3, 5\}) \\&= P(\{1\}) + P(\{3\}) + P(\{5\}) \\&= \frac{1}{9} + \frac{1}{6} + \frac{1}{6} \\&= \frac{4}{9}.\end{aligned}$$

Aufgabe 2.

- Es bietet sich $\Omega := \{1, \dots, 900\}$ und (da alle Zahlen die gleiche Wahrscheinlichkeit haben sollen) die Laplace-Verteilung P auf Ω an, also $P(\{\omega\}) = \frac{1}{|\Omega|} = \frac{1}{900}$ für alle $\omega \in \Omega$.
- Es ist $A_d = \{d, 2d, 3d, \dots, 900\}$ (die letzte Zahl in A_d ist stets 900, weil 900 nach Voraussetzung ein Vielfaches von d ist!). Damit gilt $|A_d| = \frac{900}{d}$, also $P(A_d) = \frac{|A_d|}{|\Omega|} = \frac{900/d}{900} = \frac{1}{d}$.

c) Die gesuchte Wahrscheinlichkeit ist nach Sylvester

$$\begin{aligned} P(A_2 \cup A_3 \cup A_5) &= P(A_2) + P(A_3) + P(A_5) \\ &\quad - P(A_2 \cap A_3) - P(A_2 \cap A_5) - P(A_3 \cap A_5) \\ &\quad + P(A_2 \cap A_3 \cap A_5) \\ &= \dots \end{aligned}$$

Nun ist allgemein eine Zahl genau dann durch d und durch e teilbar, wenn sie durch $\text{kgV}(d, e)$ teilbar ist; dies bedeutet $A_d \cap A_e = A_{\text{kgV}(d,e)}$. Damit können wir weiter umformen:

$$\begin{aligned} \dots &= P(A_2) + P(A_3) + P(A_5) \\ &\quad - P(A_6) - P(A_{10}) - P(A_{15}) \\ &\quad + P(A_{30}) \\ &= \frac{1}{2} + \frac{1}{3} + \frac{1}{5} - \frac{1}{6} - \frac{1}{10} - \frac{1}{15} + \frac{1}{30} \\ &= \frac{11}{15} \approx 0,733 \end{aligned}$$

d) Es sei B die Menge aller Zahlen in Ω , die weder durch 2 noch durch 3 noch durch 5 teilbar sind. Dann ist die Wahrscheinlichkeit, daß eine vom Zufallsgenerator gelieferte Zahl in B liegt, genau

$$P(B) = P(\overline{A_2} \cap \overline{A_3} \cap \overline{A_5}) = P(\overline{A_2 \cup A_3 \cup A_5}) = 1 - P(A_2 \cup A_3 \cup A_5) = 1 - \frac{11}{15} = \frac{4}{15}.$$

Andererseits ist aber

$$P(B) = \frac{|B|}{|\Omega|} = \frac{|B|}{900}.$$

Also ist $\frac{|B|}{900} = \frac{4}{15}$, also $|B| = \frac{4}{15} \cdot 900 = 240$. Damit gibt es 240 solche Zahlen im Bereich $1, \dots, 900$. (Das sind genau die Zahlen, die im Sieb des Eratosthenes für die Zahlen ≤ 900 übriggeblieben sind, nachdem die ersten drei Primzahlen 2, 3, 5 bearbeitet wurden.)

Aufgabe 3. Jede Packung enthält fünf Sticker, deren Auswahl man modellieren kann durch das Ziehen von fünf Kugeln aus einer Urne mit *allen* 640 verschiedenen Sticker-Motiven. Da garantiert niemals das gleiche Motiv mehrfach gewählt wird, wird *ohne Zurücklegen* gezogen; da aber die Reihenfolge der Karten innerhalb der Packung ohne Belang ist, ziehen wir *ohne Beachtung der Reihenfolge*.

Es sei M die Menge aller verschiedenen Motive; uns interessiert hauptsächlich, daß $|M| = 640$ ist. (Man kann sich die Motive auch einfach durchnummeriert denken und, entsprechend der offiziellen Zählung, $M = \{0, 1, \dots, 639\}$ verwenden.)

Ein möglicher Ergebnisraum für das Zufallsexperiment „Kauf (oder Zusammenstellung) einer Stickerpackung“ ist dann

$$\Omega := \{X \subset M \mid |X| = 5\},$$

die Menge aller fünf-elementigen Teilmengen von M . Es ist

$$|\Omega| = \binom{640}{5} = \frac{640 \cdot 639 \cdot 638 \cdot 637 \cdot 636}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 880.880.128.128,$$

und es ist aufgrund der Garantien des Herstellers anzunehmen, daß alle Elemente von Ω die gleiche Wahrscheinlichkeit haben, daß die Wahrscheinlichkeitsverteilung P auf Ω also die Laplace-Verteilung sein muß.

Von den insgesamt $17 \cdot 32 = 544$ Spielermotiven besitzt Jogi nun bereits $10 \cdot 32 = 320$, so daß ihm noch $550 - 320 = 224$ Spielermotive fehlen. Für das Ereignis

A : „Die Packung enthält nur bislang nicht in Jogis Sammlung vorhandene Spieler“

gilt also

$$|A| = \binom{224}{5} = \frac{224 \cdot 223 \cdot 222 \cdot 221 \cdot 220}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 4.493.032.544.$$

Damit ergibt sich

$$P(A) = \frac{\binom{224}{5}}{\binom{640}{5}} = \frac{224 \cdot 223 \cdot 222 \cdot 221 \cdot 220}{640 \cdot 639 \cdot 638 \cdot 637 \cdot 636} \approx 0,0051,$$

also eine Wahrscheinlichkeit von etwa einem halben Prozent.

(Jogi besitzt mit seinen 320 Spielern ja bereits die Hälfte aller überhaupt gedruckten Motive. Wenn man das Ereignis vergrößert und danach fragt, mit welcher Wahrscheinlichkeit Jogis in der Packung ausschließlich *irgendwelche* Motive findet, die er noch nicht besitzt, so ergibt sich eine Wahrscheinlichkeit von ziemlich genau 3 Prozent. Die Vergrößerung kommt natürlich dadurch zustande, daß nun nicht nur fünf der 224 noch fehlenden Spielerkarten, sondern fünf der 320 überhaupt noch fehlenden Karten erlaubt sind.)

Aufgabe 4.

- a) Diese Formel kann man, ähnlich wie in Aufgabe 3 vom 8. Tutoriumsblatt, aus der Formel von Sylvester für drei und zwei Ereignisse folgern, indem man *Klammern setzt*. Eine mögliche Rechnung lautet

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3 \cup A_4) &= P((A_1 \cup A_2 \cup A_3) \cup A_4) \\ &= P(A_1 \cup A_2 \cup A_3) + P(A_4) - P((A_1 \cup A_2 \cup A_3) \cap A_4) \\ &= P(A_1) + P(A_2) + P(A_3) \\ &\quad - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) \\ &\quad + P(A_1 \cap A_2 \cap A_3) \\ &\quad + P(A_4) \\ &\quad - P((A_1 \cap A_4) \cup (A_2 \cap A_4) \cup (A_3 \cap A_4)) \\ &= P(A_1) + P(A_2) + P(A_3) + P(A_4) \\ &\quad - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) \\ &\quad + P(A_1 \cap A_2 \cap A_3) \\ &\quad - P(A_1 \cap A_4) - P(A_2 \cap A_4) - P(A_3 \cap A_4) \\ &\quad + P(A_1 \cap A_2 \cap A_4) + P(A_1 \cap A_3 \cap A_4) + P(A_2 \cap A_3 \cap A_4) \\ &\quad - P(A_1 \cap A_2 \cap A_3 \cap A_4), \end{aligned}$$

und dies ist (bis auf die Sortierung der Summanden) die behauptete Formel. – Eine andere mögliche Rechnung klammert zu Beginn anders und verwendet dann ausschließlich die Formel für *zwei*

Ereignisse:

$$\begin{aligned}
P(A_1 \cup A_2 \cup A_3 \cup A_4) &= P((A_1 \cup A_2) \cup (A_3 \cup A_4)) \\
&= P(A_1 \cup A_2) + P(A_3 \cup A_4) - P((A_1 \cup A_2) \cap (A_3 \cup A_4)) \\
&= P(A_1) + P(A_2) - P(A_1 \cap A_2) \\
&\quad P(A_3) + P(A_4) - P(A_3 \cap A_4) \\
&\quad - P((A_1 \cap (A_3 \cup A_4)) \cup (A_2 \cap (A_3 \cup A_4))) \\
&= P(A_1) + P(A_2) + P(A_3) + P(A_4) \\
&\quad - P(A_1 \cap A_2) - P(A_3 \cap A_4) \\
&\quad - P(A_1 \cap (A_3 \cup A_4)) - P(A_2 \cap (A_3 \cup A_4)) \\
&\quad + P(A_1 \cap A_2 \cap (A_3 \cup A_4)) \\
&= P(A_1) + P(A_2) + P(A_3) + P(A_4) \\
&\quad - P(A_1 \cap A_2) - P(A_3 \cap A_4) \\
&\quad - P((A_1 \cap A_3) \cup (A_1 \cap A_4)) - P((A_2 \cap A_3) \cup (A_2 \cap A_4)) \\
&\quad + P((A_1 \cap A_2 \cap A_3) \cup (A_1 \cap A_2 \cap A_4)) \\
&= P(A_1) + P(A_2) + P(A_3) + P(A_4) \\
&\quad - P(A_1 \cap A_2) - P(A_3 \cap A_4) \\
&\quad - P(A_1 \cap A_3) - P(A_1 \cap A_4) + P(A_1 \cap A_3 \cap A_4) \\
&\quad - P(A_2 \cap A_3) - P(A_2 \cap A_4) + P(A_2 \cap A_3 \cap A_4) \\
&\quad + P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_4) \\
&\quad - P(A_1 \cap A_2 \cap A_3 \cap A_4),
\end{aligned}$$

und wieder ist dies (bis auf die Reihenfolge der Summanden) die behauptete Formel.

- b) Die aufgrund der bisher behandelten Spezialfälle $n = 2, 3, 4$ zu vermutende Formel lautet: Ist $(\Omega, \mathcal{P}(\Omega), P)$ ein (endlicher) Wahrscheinlichkeitsraum, und sind $A_1, \dots, A_n \subset \Omega$ Ereignisse, so gilt

$$\begin{aligned}
&P(A_1 \cup \dots \cup A_n) = \\
&= P(A_1) + \dots + P(A_n) \\
&\quad - (P(A_1 \cap A_2) + P(A_1 \cap A_3) + \dots + P(A_{n-1} \cap A_n)) \\
&\quad + (P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_4) + \dots + P(A_{n-2} \cap A_{n-1} \cap A_n)) \\
&\quad - \dots \\
&\quad + \dots \\
&\quad \vdots \\
&\quad + (-1)^{n-1} P(A_1 \cap A_2 \cap \dots \cap A_n) \\
&= \sum_{k=1}^n \left[(-1)^{k-1} \left(\sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k}) \right) \right]
\end{aligned}$$

Diese Formel ist sicher gültig für $n = 2, 3, 4$ (diese Fälle haben wir bewiesen); für $n = 1$ ist sie richtig, weil sie auf die Aussage $P(A_1) = P(A_1)$ zusammenschumpft, und ebenso sogar für $n = 0$: in diesem Fall besagt sie $P(\emptyset) = 0$ (denn unter einer „leeren Vereinigung“ ist die leere Menge zu verstehen, und eine leere Summe ergibt 0).

Zum Beweis dieser Formel bietet sich vollständige Induktion nach n an. Induktionsanfänge gibt es bereits in genügender Zahl; auf jeden Fall muß die Gültigkeit des Falles $n = 2$ schon gezeigt sein,

denn dieser wird im Induktionsschritt angewandt. Für diesen Induktionsschritt $n - 1 \rightarrow n, n \geq 3$, rechnen wir wie in der ersten unter a) vorgeschlagenen Rechnung (denn diese blieb überschaubarer) unter Verwendung der Formel $n - 1$ und für 2 Mengen:

$$\begin{aligned}
& P(A_1 \cup \dots \cup A_n) = \\
& = P((A_1 \cup \dots \cup A_{n-1}) \cup A_n) \\
& = P(A_1 \cup \dots \cup A_{n-1}) + P(A_n) \\
& \quad - P((A_1 \cup \dots \cup A_{n-1}) \cap A_n) \\
& = P(A_1 \cup \dots \cup A_{n-1}) + P(A_n) \\
& \quad + P(A_n) \\
& \quad - P((A_1 \cap A_n) \cup \dots \cup (A_{n-1} \cap A_n)) \\
& = \sum_{k=1}^{n-1} \left[(-1)^{k-1} \left(\sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n-1} P(A_{i_1} \cap \dots \cap A_{i_k}) \right) \right] \\
& \quad + P(A_n) \\
& \quad - \sum_{k=1}^{n-1} \left[(-1)^{k-1} \left(\sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n-1} P(A_{i_1} \cap \dots \cap A_{i_k} \cap A_n) \right) \right] \\
& \stackrel{(*)}{=} \sum_{k=1}^{n-1} \left[(-1)^{k-1} \left(\sum_{1 \leq i_1 < i_2 < \dots < i_k < n} P(A_{i_1} \cap \dots \cap A_{i_k}) \right) \right] \\
& \quad + P(A_n) \\
& \quad + \sum_{k=2}^n \left[(-1)^{k-1} \left(\sum_{1 \leq i_1 < i_2 < \dots < i_{k-1} < i_k = n} P(A_{i_1} \cap \dots \cap A_{i_k}) \right) \right] \\
& \stackrel{(**)}{=} \sum_{k=1}^n \left[(-1)^{k-1} \left(\sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k}) \right) \right].
\end{aligned}$$

Im Schritt (*) wurde dabei der zweiten großen Summe umindiziert, indem der vormals einzeln aufgeführte Schnitt-Faktor A_n nun in den Schnitt $A_{i_1} \cap \dots \cap A_{i_k}$ mit aufgenommen wurde. Dadurch hat sich k um 1 vergrößert, und es ist $i_k = n$.

Im Schritt (**) wurden die beiden großen Summen zu einer einzigen zusammengefaßt. Die erste große Summe umfaßt alle Summanden mit $i_k < n$, die zweite alle Summanden mit $i_k = n$; da die zweite Summe jedoch erst bei $k = 2$ beginnt, fehlt der entsprechende Summand für $k = 1$. Dieser lautet aber

$$(-1)^{k-1} \left(\sum_{1 \leq i_1 < i_2 < \dots < i_{k-1} < i_k = n} P(A_{i_1} \cap \dots \cap A_{i_k}) \right) \stackrel{k=1}{=} P(A_n),$$

und ist damit genau der einzeln aufgeführte Summand.

Für Interessierte: Zwei alternative Beweise für die allgemeine Formel von Sylvester

Der im Lösungsvorschlag für Aufgabe 4 b) angegebene Induktionsbeweis für die Sylvestersche Formel hat einige Nachteile:

- Er ist einigermaßen unübersichtlich.
- Man sieht, *daß* die Aussage stimmt, aber vielleicht weniger, *warum* sie stimmt. (Dies gilt oft für Induktionsbeweise.)

- Man muß die Aussage schon kennen, die man beweisen möchte, und bekommt sie nicht als „Ergebnis“ einer Argumentation. (Auch dies ist ein notorisches Problem bei Induktionsbeweisen.)

Aus diesem Grund hier zwei Vorschläge für alternative Beweise, die einige (wenn auch vielleicht nicht alle) dieser Nachteile umgehen:

Zweiter Beweis: Aufspalten in Elementarereignisse

$$P(A_1 \cup \dots \cup A_n) = \sum_{k=1}^n \left[(-1)^{k-1} \left(\sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k}) \right) \right]$$

schreiben wir die Wahrscheinlichkeit $P(A)$ jedes Ereignisses $A \subset \Omega$ als Summe von Elementarereignissen der Form

$$P(A) = \sum_{\omega \in A} P(\{\omega\})$$

und sortieren die große Summe in der Formel von Sylvester so, daß für jedes $\omega \in \Omega$ die Faktoren $P(\{\omega\})$ ausgeklammert werden. Das läßt sich am leichtesten notieren mit Hilfe der sogenannten *charakteristischen Funktionen*: Für jedes Ereignis $A \subset \Omega$ ist $\chi_A : \Omega \rightarrow \mathbb{R}$ die Funktion mit

$$\chi_A(\omega) = \begin{cases} 1 & \text{für } \omega \in A, \\ 0 & \text{für } \omega \notin A. \end{cases}$$

Es gilt dann $P(A) = \sum_{\omega \in A} P(\{\omega\}) = \sum_{\omega \in \Omega} \chi_A(\omega) \cdot P(\{\omega\})$, damit kann man die rechte Seite der Sylvesterschen Formel umformen zu

$$\begin{aligned} & \sum_{k=1}^n \left[(-1)^{k-1} \left(\sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \sum_{\omega \in \Omega} \chi_{A_{i_1} \cap \dots \cap A_{i_k}}(\omega) \cdot P(\{\omega\}) \right) \right] \\ &= \sum_{\omega \in \Omega} P(\{\omega\}) \cdot \left[\sum_{k=1}^n (-1)^{k-1} \left(\sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \chi_{A_{i_1} \cap \dots \cap A_{i_k}}(\omega) \right) \right] = \dots \end{aligned}$$

Der Term in großen runden Klammern im Inneren dieses Ausdrucks zählt die Anzahl der Schnitte aus k der Mengen A_1, \dots, A_n , die ω enthalten. Diese Anzahl beträgt genau $\binom{c(\omega)}{k}$, wobei wir mit

$$c(\omega) := |\{i \in \{1, \dots, n\} \mid \omega \in A_i\}|$$

die Anzahl derjenigen der Mengen A_1, \dots, A_n bezeichnen, die das Ergebnis ω enthalten.¹ Setzen wir dieses Resultat in die obige Summe Formel ein, erhält sie die Gestalt

$$\dots = \sum_{\omega \in \Omega} P(\{\omega\}) \cdot \left[\sum_{k=1}^n (-1)^{k-1} \binom{c(\omega)}{k} \right] = \sum_{\omega \in \Omega} P(\{\omega\}) \cdot \left[\sum_{k=1}^{c(\omega)} (-1)^{k-1} \binom{c(\omega)}{k} \right] = \dots$$

Nun ist aber nach der binomischen Formel für jedes $c \in \mathbb{N}_0$

$$\begin{aligned} \sum_{k=1}^c (-1)^{k-1} \binom{c}{k} &= 1 + \sum_{k=0}^c (-1)^{k-1} \binom{c}{k} = 1 - \sum_{k=0}^c (-1)^k \binom{c}{k} \\ &= 1 - (1-1)^c = 1 - 0^c \\ &= \begin{cases} 1 & \text{für } c > 0, \\ 0 & \text{für } c = 0, \end{cases} \end{aligned}$$

¹Wie üblich ist dabei $\binom{m}{k} = 0$ für $m < k$.

und da genau dann $c(\omega) > 0$ ist, wenn $\omega \in A_1 \cup \dots \cup A_n$ ist, bedeutet dies

$$\sum_{k=1}^c (-1)^{k-1} \binom{c(\omega)}{k} = \chi_{A_1 \cup \dots \cup A_n}(\omega) \quad \text{für alle } \omega \in \Omega,$$

und wir können wir schließlich weiter umformen zu

$$\dots = \sum_{\omega \in \Omega} P(\{\omega\}) \cdot \chi_{A_1 \cup \dots \cup A_n}(\omega) = P(A_1 \cup \dots \cup A_n),$$

womit die Formel bewiesen ist.

Auch dieser Beweis hat einige Nachteile:

- Der Rückgriff auf die Binomische Formel bedeutet die versteckte Anwendung eines Induktionsbeweises.
- Man fängt mit der „komplizierten“ Seite der Formel an und formt sie um, bis der eigentlich interessierende Ausdruck $P(A_1 \cup \dots \cup A_n)$ herauskommt. Insbesondere muß man auch hier die Formel schon kennen, um sie beweisen zu können!
- Das Argument ist zwar an sich nicht kompliziert, jedoch mühsam zu notieren.

Dritter Beweis: Erwartungswert von charakteristischen Funktionen

Ein sehr eleganter Beweis ist möglich, wenn man den (in der Vorlesung nicht behandelten) Begriff der *Erwartungswertes* verwendet. Ist $f : \Omega \rightarrow \mathbb{R}$ eine beliebige Funktion, so nennt man

$$\mathbb{E}^P(f) := \mathbb{E}(f) := \sum_{\omega \in \Omega} P(\{\omega\}) \cdot f(\omega)$$

den *Erwartungswert* von f (unter der Wahrscheinlichkeitsverteilung P). Wir benötigen nun einige Tatsachen, die man ohne Mühe direkt nachrechnen kann:

- *Linearität des Erwartungswertes:* Sind $f, g : \Omega \rightarrow \mathbb{R}$ zwei Funktionen, so ist $\mathbb{E}(f + g) = \mathbb{E}(f) + \mathbb{E}(g)$. Ist außerdem $\lambda \in \mathbb{R}$ eine beliebige Zahl, so ist $\mathbb{E}(\lambda f) = \lambda \cdot \mathbb{E}(f)$.
(Der Erwartungswert eines *Produktes* von Funktionen ist aber nicht so einfach zu berechnen!)
- *Erwartungswert von charakteristischen Funktionen:* Ist $A \subset \Omega$ ein Ereignis, so ist $\mathbb{E}(\chi_A) = P(A)$.
- *Charakteristische Funktionen von Komplement und Durchschnitt:* Für ein Ereignis $A \subset \Omega$ gilt $\chi_{\bar{A}} = 1 - \chi_A$, und für zwei Ereignisse $A_1, A_2 \subset \Omega$ gilt $\chi_{A_1 \cap A_2} = \chi_{A_1} \cdot \chi_{A_2}$.
(Die charakteristische Funktion einer *Vereinigung* von Ereignissen ist aber nicht so einfach zu berechnen – dies ist gewissermaßen der Kern der ganzen Aufgabe!)

Damit kann man nun die Wahrscheinlichkeit einer Vereinigung von Ereignissen berechnen, indem man verwendet, daß Komplementbildung Vereinigungen in Schnitte übersetzt:

$$\begin{aligned} P(A_1 \cup \dots \cup A_n) &= \mathbb{E}(\chi_{A_1 \cup \dots \cup A_n}) = \mathbb{E}(1 - \chi_{\overline{A_1 \cup \dots \cup A_n}}) = \mathbb{E}(1 - \chi_{\bar{A}_1 \cap \dots \cap \bar{A}_n}) \\ &= \mathbb{E}(1 - \chi_{\bar{A}_1} \cdot \dots \cdot \chi_{\bar{A}_n}) = \mathbb{E}(1 - (1 - \chi_{A_1}) \cdot \dots \cdot (1 - \chi_{A_n})) = \dots \end{aligned}$$

Nun gilt aber ganz allgemein

$$(1 - t_1) \cdot \dots \cdot (1 - t_n) = \sum_{k=0}^n (-1)^k \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} t_{i_1} \cdot \dots \cdot t_{i_k} \right),$$

wie man durch Ausmultiplizieren sieht.² Also können wir weiter umformen:

$$\begin{aligned}
 \dots &= \mathbb{E} \left(1 - \sum_{k=0}^n (-1)^k \left[\left(\sum_{1 \leq i_1 < \dots < i_k \leq n} \chi_{A_{i_1}} \cdot \dots \cdot \chi_{A_{i_k}} \right) \right] \right) \\
 &= \mathbb{E} \left(\sum_{k=1}^n \left[(-1)^{k-1} \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} \chi_{A_{i_1} \cap \dots \cap A_{i_k}} \right) \right] \right) \\
 &= \sum_{k=1}^n \left[(-1)^{k-1} \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} \mathbb{E}(\chi_{A_{i_1} \cap \dots \cap A_{i_k}}) \right) \right] \\
 &= \sum_{k=1}^n \left[(-1)^{k-1} \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} P(\chi_{A_{i_1} \cap \dots \cap A_{i_k}}) \right) \right].
 \end{aligned}$$

Nach der (unmaßgeblichen) Ansicht des Schreibers dieser Zeilen ist das in vielerlei Hinsicht der „gute“ Beweis für die Formel von Sylvester: Man sieht, wo die Formel herkommt; der entscheidende Schritt ist eine einfache algebraische Umformung (Ausmultiplizieren!); der Beweis beginnt auf der „richtigen“ Seite der Formel, und es gehen nur einfache, leicht zu merkende Rechenregeln ein. Diese Vorteile erkaufte man sich freilich mit der Verwendung des bislang nicht zur Verfügung stehenden Begriffs des Erwartungswertes.

²Allgemeiner besagt der (für $n = 2$ aus der Schule bekannte) *Satz von Vieta*: Multipliziert man das Polynom $P = (X - t_1) \cdot \dots \cdot (X - t_n)$ mit den Nullstellen t_1, \dots, t_n aus, also $P = c_n X^n + c_{n-1} X^{n-1} + \dots + c_0$, so gilt

$$c_{n-k} = (-1)^k \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} t_{i_1} \cdot \dots \cdot t_{i_k} \right).$$

Auch dieser „Satz“ wird im Wesentlichen durch sorgfältiges Ausmultiplizieren bewiesen; mittels Einsetzen von $X = 1$ liefert er unsere oben verwendete Formel.