Technische Universität München
Fakultät für Informatik
Lehrstuhl für Effiziente Algorithmen
Prof. Dr. Ernst W. Mayr
Dr. Werner Meixner

Sommersemester 2011
Lösungsblatt 4
3. Juni 2011

# Diskrete Wahrscheinlichkeitstheorie

## Hausaufgabe 1 (5 Punkte)

Wir betrachten drei 6-seitige Würfel $A$, $B$ und $C$. Würfel $A$ hat 4 rote und 2 blaue Seiten. Würfel $B$ hat 2 rote und 4 blaue Seiten. Der Würfel $C$ ist ein üblicher Würfel, der die Augenzahlen 1 bis 6 zeigt. Wir nehmen an, dass die Ergebnisse von Würfen Laplace-verteilt sind bezüglich des Auftretens der Würfelseiten.

Experiment: Es wird zunächst $C$ geworfen. Falls die 6 geworfen wurde, so wird Würfel $A$ gewählt, ansonsten wird Würfel $B$ gewählt. Mit dem gewählten Würfel werden dann $n \geq 3$ Würfe durchgeführt. Das Ergebnis ist ein Wort $w \in \{\text{rot}, \text{blau}\}^*$ der Länge $n$.

1. Wir sagen, dass das Ereignis $R_i$ eintritt, wenn das $i$-te Zeichen der Ausgabe $w$ des Experiments 'rot' ist. Wir nehmen an, dass die Ereignisse $R_1$ und $R_2$ eingetreten sind. Wie hoch ist die Wahrscheinlichkeit, dass auch das Ereignis $R_3$ eintritt?

2. Wir nehmen an, dass das Ereignis $\bigcap_{i=1}^{n} R_i$ eingetreten ist. Wie groß ist die Wahrscheinlichkeit, dass im Experiment der Würfel $A$ gewählt wurde?

**Lösung**

Der Ergebnisraum ist $\Omega = \{(x_1, x_2, \ldots, x_n) \mid x_i \in \{\text{rot}, \text{blau}\}\}$. Die Wahrscheinlichkeitsfunktion Pr über $\Omega$ ist gegeben mittels Fallunterscheidung der totalen Wahrscheinlichkeit und des Produkts von Wahrscheinlichkeiten der Komponenten des Ergebnisses in den jeweiligen Fällen.

$$\Pr[x_i|\text{C=6}] = \begin{cases} \frac{2}{3} : x_i = \text{rot} \\ \frac{1}{3} : x_i = \text{blau} \end{cases}, \quad \Pr[x_i|C \neq 6] = \begin{cases} \frac{1}{3} : x_i = \text{rot} \\ \frac{2}{3} : x_i = \text{blau} \end{cases},$$

und damit

$$\Pr[(x_1, x_2, \ldots, x_n)] \;=\; \Pr[C = 6] \cdot \prod_{i=1}^{n} \Pr[x_i|C = 6] \;+\; \Pr[C \neq 6] \cdot \prod_{i=1}^{n} \Pr[x_i|C \neq 6]\,.$$

1. Speziell für $\Pr[R_1 \cap R_2]$ und $\Pr[R_1 \cap R_2 \cap R_3]$ ergibt sich

$$\begin{aligned} \Pr[R_1 \cap R_2] &= \frac{1}{6} \cdot \left(\frac{2}{3}\right)^2 + \frac{5}{6} \cdot \left(\frac{1}{3}\right)^2 = \frac{1}{6}\,, \\ \Pr[R_1 \cap R_2 \cap R_3] &= \frac{1}{6} \cdot \left(\frac{2}{3}\right)^3 + \frac{5}{6} \cdot \left(\frac{1}{3}\right)^3 = \frac{13}{162}\,. \end{aligned}$$

Nach Definition der bedingten Wahrscheinlichkeit erhalten wir

$$\Pr[R_3 \mid R_1 \cap R_2] = \frac{\Pr[R_1 \cap R_2 \cap R_3]}{\Pr[R_1 \cap R_2]} = \frac{13}{27}\,.$$

2. Wir suchen $\Pr[C = 6 \mid \bigcap_{i=1}^n R_i]$. Nun gilt

$$\Pr[C = 6 \mid \bigcap_{i=1}^n R_i] = \frac{\Pr[C = 6] \cdot \Pr[\bigcap_{i=1}^n R_i \mid C = 6]}{\Pr[\bigcap_{i=1}^n R_i]} .$$

Sämtliche Werte der rechten Seite der Gleichung sind uns bekannt und wir erhalten

$$\begin{aligned}\Pr[C = 6] \cdot \Pr[\bigcap_{i=1}^n R_i \mid C = 6] &= \frac{1}{6} \cdot \left(\frac{2}{3}\right)^n ,\\ \Pr[\bigcap_{i=1}^n R_i] &= \frac{1}{6} \cdot \left(\frac{2}{3}\right)^n + \frac{5}{6} \cdot \left(\frac{1}{3}\right)^n ,\\ \Pr[C = 6 \mid \bigcap_{i=1}^n R_i] &= \frac{2^n}{2^n + 5} .\end{aligned}$$

*Interpretation*: Je öfter rot gewürfelt wird, desto wahrscheinlicher ist es, dass am Anfang $C = 6$ gewürfelt wurde. Dies ist eine typische Argumentation nach dem Satz von Bayes.

## Hausaufgabe 2 (5 Punkte)

Sei $n \in \mathbb{N}$. Man zeige:
Paarweise verschiedene Ereignisse $A_1, A_2, \ldots, A_n$ sind genau dann unabhängig, wenn die Indikatorvariablen $I_{A_1}, I_{A_2}, \ldots, I_{A_n}$ unabhängig sind.

**Lösung**

$\Rightarrow$:
Seien $A_1, A_2, \ldots, A_n$ paarweise verschieden und unabhängig. Dann haben wir für alle $s_1 \in W_{I_{A_1}}, \ldots s_n \in W_{I_{A_n}}$ zu zeigen, dass gilt

$$\Pr[I_{A_1} = s_1, \ldots, I_{A_n} = s_n] = \Pr[I_{A_1} = s_1] \cdot \ldots \cdot \Pr[I_{A_n} = s_n] .$$

Für Indikatorvariablen gilt $W_{I_{A_i}} \subseteq \{0, 1\}$, d. h. $s_i \in \{0, 1\}$ für alle $i$. Mit der Bezeichnungsweise von Lemma 23 der Vorlesung folgt

$$A^{s_i} = \{\omega\, ;\ I_{A_i}(\omega) = s_i\} .$$

Mithin folgt durch Anwendung von Lemma 23

$$\begin{aligned}\Pr[I_{A_1} = s_1, \ldots, I_{A_n} = s_n] &= \Pr[A^{s_1} \cap \ldots \cap A^{s_n}]\\ &= \Pr[A^{s_1}] \cdot \ldots \cdot \Pr[A^{s_n}]\\ &= \Pr[I_{A_1} = s_1] \cdot \ldots \cdot \Pr[I_{A_n} = s_n] .\end{aligned}$$

$\Leftarrow$:
Seien die Indikatorvariablen $I_{A_1}, I_{A_2}, \ldots, I_{A_n}$ unabhängig. Dann folgt durch Anwendung von Satz 46 der Vorlesung für die Mengen $S_i = \{1\}$ für alle $i$, dass die Ereignisse $X_1 \in S_1$, $\ldots$, $X_n \in S_n$ unabhängig sind.
Da die Ereignisse $X_i \in S_i$ identisch sind mit den entsprechenden Ereignissen $A_i$, folgt die zu beweisende Aussage.

## Hausaufgabe 3 (5 Punkte)

Sei $\langle K, +, \cdot \rangle$ ein Körper. Man beweise durch vollständige Induktion, dass für alle $n \in \mathbb{N}$ und $x_1, x_2, \ldots, x_n \in K$ gilt:

$$\begin{aligned}
\prod_{i=1}^{n}(1-x_i) \quad = \quad & 1 - \sum_{1\leq i_1\leq n} x_{i_1} \\
& + \sum_{1\leq i_1<i_2\leq n} x_{i_1}\cdot x_{i_2} \\
& \vdots \\
& + (-1)^l \cdot \sum_{1\leq i_1<\ldots<i_l\leq n} x_{i_1}\cdot \ldots \cdot x_{i_l} \\
& \vdots \\
& + (-1)^n \cdot x_1 \cdot \ldots \cdot x_n\,.
\end{aligned}$$

**Lösung**

Wir beweisen die Eigenschaft $P(n)$ für alle $n \in \mathbb{N}$, wobei $P(n)$ genau dann wahr sei, wenn die Gleichung für alle $x_1, \ldots, x_n$ gilt.

$\underline{n = 1:}$

alle Summen sind leer mit Ausnahme der ersten Summe, d. h. $P(1)$ ist äquivalent mit der Gleichung $(1 - x_1) = 1 - x_1$, die natürlich gültig ist.

$\underline{n \to n+1:}$

Wir setzen $P(n)$ voraus und beweisen $P(n + 1)$.

$$\begin{aligned}
\prod_{i=1}^{n+1}(1-x_i) \quad = \quad & (1-x_{n+1})\left(1-\textstyle\sum_{1\leq i_1\leq n} x_{i_1}\right) \\
& + (1-x_{n+1})\textstyle\sum_{1\leq i_1<i_2\leq n} x_{i_1}\cdot x_{i_2} \\
& \vdots \\
& + (1-x_{n+1})(-1)^l\textstyle\sum_{1\leq i_1<\ldots<i_l\leq n} x_{i_1}\cdot\ldots\cdot x_{i_l} \\
& \vdots \\
& + (1-x_{n+1})(-1)^n\cdot x_1\cdot\ldots\cdot x_n\,. \\
= \quad & 1 - x_{n+1} - \textstyle\sum_{1\leq i_1\leq n} x_i \\
& + \textstyle\sum_{1\leq i_1\leq n} x_i x_{n+1} \; + \; \textstyle\sum_{1\leq i_1<i_2\leq n} x_{i_1}\cdot x_{i_2} \\
& \vdots \\
& + (-1)^l\cdot \sum_{1\leq i_1<\ldots<i_{l-1}\leq n} x_{i_1}\cdot\ldots\cdot x_{i_{l-1}}x_{n+1} \; + \; (-1)^l\cdot\sum_{1\leq i_1<\ldots<i_l\leq n} x_{i_1}\cdot\ldots\cdot x_{i_l} \\
& \vdots \\
& + (-1)^n\textstyle\sum_{1\leq i_1<\ldots<i_{n-1}\leq n} x_{i_1}\cdot\ldots\cdot x_{i_{n-1}}x_{n+1} \; + \; (-1)^n\cdot x_1\cdot\ldots\cdot x_n \\
& + (-1)^{n+1}\cdot x_1\cdot\ldots\cdot x_n x_{n+1}\,.
\end{aligned}$$

Wir fassen zusammen zu

$$\begin{aligned}\prod_{i=1}^{n+1}(1-x_i) &= 1-\sum_{1\le i_1\le n+1} x_{i_1} \\ &\quad+\sum_{1\le i_1<i_2\le n+1} x_{i_1}\cdot x_{i_2} \\ &\quad\vdots \\ &\quad+(-1)^l\cdot \sum_{1\le i_1<\ldots<i_l\le n+1} x_{i_1}\cdot\ldots\cdot x_{i_l} \\ &\quad\vdots \\ &\quad+(-1)^{n+1}\cdot x_1\cdot\ldots\cdot x_{n+1}\,.\end{aligned}$$

## Hausaufgabe 4 (5 Punkte)

Sie besitzen $n$ Passwörter, von denen genau eines den Zugang zu einem Rechner gestattet. Sie wollen Zugang zu dem Rechner erhalten und wählen zufällig so lange ein Passwort aus, bis Sie schließlich das Richtige gefunden haben.
Berechnen Sie den Erwartungswert der Zufallsvariablen $X_n$, die die Anzahl der benötigten Versuche zählt, wenn

1. jedes Passwort höchstens einmal verwendet wird.
2. jedes Passwort beliebig oft verwendet werden kann. Unterscheiden Sie dabei, ob der erste Versuch erfolgreich war oder nicht.

**Lösung**

1. Es gilt $W_{X_n}=\{1,2,\ldots,n\}$.

   Für die Wahrscheinlichkeit $\Pr[X_n>i]$ eines Misserfolgs in allen ersten $i$ Schritten gilt
   $$\Pr[X_n>i]=\frac{n-1}{n}\cdot\frac{n-2}{n-1}\cdot\ \ldots\ \cdot\frac{n-i}{n-(i-1)}=\frac{n-i}{n}\,.$$
   Entsprechend gilt für die Wahrscheinlichkeit $\Pr[X_n=i]$ eines Erfolgs im $i$-ten Schritt
   $$\Pr[X_n\!=\!i]=\Pr[X_n>i]\cdot\frac{1}{n-i}=\frac{1}{n}\,.$$
   Es folgt
   $$\begin{aligned}\mathbb{E}[X_n] &= \sum_{i\in W_{X_n}} i\cdot\Pr[x_n=i] \\ &= \frac{n(n+1)}{2}\cdot\frac{1}{n} \\ &= \frac{n+1}{2}\,.\end{aligned}$$

2. Es gilt $W_{X_n} = \mathbb{N}$. Sei $E_1$ bzw. $\overline{E_1}$ das Ereignis, dass beim ersten Versuch das Passwort getroffen bzw. nicht getroffen wird. Es gilt

$$\mathbb{E}[X_n] = \mathbb{E}[X_n|E_1] \cdot \Pr[E_1] + \mathbb{E}[X_n|\overline{E_1}] \cdot \Pr[\overline{E_1}]\,.$$

Mit $\mathbb{E}[X_n|E_1] = 1$ und $\mathbb{E}[X_n|\overline{E_1}] = 1 + \mathbb{E}[X_n]$ bzw. $\Pr[E_1] = \frac{1}{n}$ und $\Pr[\overline{E_1}] = \frac{n-1}{n}$ folgt

$$\mathbb{E}[X_n] \;=\; 1 \cdot \frac{1}{n} + (1 + \mathbb{E}[X_n]) \cdot \frac{n-1}{n}$$

Auflösung nach $\mathbb{E}[X_n]$ liefert $\mathbb{E}[X_n] = n$.

***Hinweis:*** *Auf den Übungsblättern in diesem Semester wird es grundsätzlich die drei Aufgabentypen Vorbereitungsaufgabe, Tutoraufgabe und Hausaufgabe geben. Die als Vorbereitung bezeichneten Aufgaben dienen der häuslichen Vorbereitung der Tutoraufgaben. Tutoraufgaben werden in den Übungsgruppen bearbeitet. Dabei wird die Lösung der Vorbereitungsaufgaben vorausgesetzt. Die Vorbereitungsaufgaben werden in der Zentralübung unterstützt.*

## Vorbereitung 1

In der statistischen Physik pflegt man die Verteilung gewisser Teilchen (Moleküle, Photonen, Elektronen, usw.) zu betrachten. Man nimmt an, dass sich jedes Teilchen durch einen Vektor von Kenngrößen charakterisieren läßt, der der Zuordnung des Teilchens in eine von $N$ „Zellen" eines „Phasenraumes" entspricht. Der Zustand eines physikalischen Systems wird dann dadurch beschrieben, dass man angibt, wie viele Teilchen sich jeweils in einer Zelle befinden.

1. Die Annahme, dass jede Verteilung von $n$ Gasmolekülen auf $N$ Zellen gleichwahrscheinlich ist, nennt man *Maxwell-Boltzmannsche Statistik*. Dabei wird die Unterscheidbarkeit der Gasmoleküle vorausgesetzt.

   Bestimmen Sie die Wahrscheinlichkeit, dass sich $k$ Gasmoleküle in einer bestimmten Zelle befinden.

2. Photonen und Elektronen werden nicht als unterscheidbare Teilchen betrachtet. Demgemäß gilt die Maxwell-Boltzmannsche Statistik nicht. Die auf der Annahme der Nichtunterscheidbarkeit von Teilchen beruhende Statistik nennt man *Bose-Einsteinsche Statistik*. Sie wird für Photonen angewendet.

   Bestimmen Sie die Wahrscheinlichkeit, dass sich in einer bestimmten Zelle $k$ Photonen befinden.

3. Für Elektronen hat man noch die zusätzliche Einschränkung (*Paulisches Prinzip*), dass sich in einer Zelle stets höchstens ein Teilchen befinden kann. Diesen Umstand berücksichtigt die *Fermi-Diracsche Statistik*.

   Bestimmen Sie die Wahrscheinlichkeit, dass sich in einer bestimmten Zelle ein Elektron befindet.

### Lösung

1. Jede örtliche Zuordnung der $n$ Gasmoleküle $G_1, \ldots, G_n$, so dass sich jedes Molekül in einer der $N$ Zellen $Z_1, \ldots, Z_N$ befindet, stellt ein Elementarereignis dar. Wir codieren die Gasmoleküle durch natürliche Zahlen $1, 2, \ldots, n$ und die Zellen durch natürliche Zahlen $1, 2, \ldots, N$. Die Elementarereignisse $e \in \Omega$ codieren wir durch $n$-Tupel $e = (e_1, e_2, \ldots, e_n)$ mit $e_i \in [N]$ und der Bedeutung, dass der Wert der $i$-ten Komponente von $e$ die Zelle liefert, in der sich das Molekül $G_i$ befindet. Es gilt mit Anwendung des Laplace Prinzips

   $$e \in \Omega = \{(e_1, e_2, \ldots, e_n)\,;\, e_i \in [N]\}\,,\ \Pr[e] = N^{-n}\,.$$

   Wir betrachten nun eine bestimmte Zelle $Z$.

Für jedes der $n$ Moleküle definieren wir eine Bernoulli Variable $X_i$, die angibt, ob sich bei Ereignis $e$ das Molekül $G_i$ in der Zelle $Z$ befindet oder nicht. Die Zufallsvariablen $X_i$ sind Bernoulli-verteilt mit Erfolgswahrscheinlichkeit $p = \frac{1}{N}$.

Die Zufallsvariablen $X_1, X_2, \ldots, X_n$ sind unabhängig.

Nun betrachten wir die Zufallsvariable

$$X_Z = \sum_{i=1}^{n} X_i \,.$$

$X_Z$ ist binomialverteilt, i. Z. $X \sim \mathrm{Bin}(n, \frac{1}{N})$, mit Dichtefunktion

$$f_{X_Z}(k) = b(k; n, \tfrac{1}{N}) := \binom{n}{k} \cdot \left(\frac{1}{N}\right)^{k} \cdot \left(\frac{N-1}{N}\right)^{n-k} \,.$$

Natürlich ist $f_{X_Z}(k)$ die gesuchte Wahrscheinlichkeit.

2. Das Elementarereignis, $n$ Photonen auf $N$ Zellen verteilt zu haben, ist genau dann unterscheidbar definiert, falls für jede Zelle die Anzahl der enthaltenen Photonen feststeht. Ein Elementarereignis entspricht also genau einer $n$-elementigen Multiteilmenge der Menge der Zellen. Die Anzahl $anz_{MTM}(n, N)$ von $n$-elementigen Multiteilmengen einer $N$-elementigen Menge ist bekanntlich

   $$anz_{MTM}(n, N) = \frac{N^{\overline{n}}}{n!} = \binom{N+n-1}{n} \,.$$

   Die Relation „Multiteilmenge“ bezeichnen wir mit $\subseteq_\mu$. Wir definieren

   $$m \in \Omega = \{m\, ;\, m \subseteq_\mu [N], |m| = n\}\,, \ \Pr[m] = \binom{N+n-1}{n}^{-1} \,.$$

   Wir betrachten nun wieder eine bestimmt Zelle $Z$.

   Wenn $Z$ genau $k$ Photonen enthält, dann sind $n-k$ Photonen auf die restlichen $N-1$ Zellen verteilt und dafür gibt es genau

   $$anz_{MTM}(n-k, N-1) = \frac{(N-1)^{\overline{n-k}}}{(n-k)!} = \binom{N+n-k-2}{n-k}$$

   Möglichkeiten.

   Sei $X_Z$ die Zufallsvariable, die für die Zelle $Z$ die Anzahl der enthaltenen Photonen angibt. Dann gilt für die Dichtfunktion von $X_Z$

   $$f_{X_Z}(k) = \binom{N+n-k-2}{n-k} \cdot \binom{N+n-1}{n}^{-1} \,.$$

   Natürlich ist wieder $f_{X_Z}(k)$ die gesuchte Wahrscheinlichkeit.

3. Das Elementarereignis, $n$ Elektronen auf $N$ Zellen verteilt zu haben, ist genau dann unterscheidbar definiert, falls für jede Zelle feststeht, ob sie ein Elektron enthält oder nicht. Ein Elementarereignis entspricht also genau einer $n$-elementigen Teilmenge

der Menge der Zellen. Die Anzahl $anz_{TM}(n, N)$ der $n$-elementigen Teilmengen einer $N$-elementigen Menge ist bekanntlich

$$anz_{TM}(n, N) = \binom{N}{n} .$$

Wir definieren mit Anwendung des Prinzips von Laplace

$$m \in \Omega = \{m\,;\, m \subseteq [N]\}\,,\ \Pr[m] = \binom{N}{n}^{-1} .$$

Wir betrachten nun wieder eine bestimmt Zelle $Z$.

Die Anzeige der Anzahl der in $Z$ enthaltenen Elektronen reduziert sich nun auf eine Bernoulli-verteilte Indikatorvariable $X_Z$. Wenn $Z$ ein Elektron enthält, dann befinden sich in den restlichen $N - 1$ Zellen die restlichen $n - 1$ Elektronen. Dafür gibt es genau

$$anz_{TM}(n - 1, N - 1) = \binom{N - 1}{n - 1}$$

Möglichkeiten. Die Erfolgswahrscheinlichkeit $p$ von $X_Z$ ist also

$$p = \binom{N - 1}{n - 1} \cdot \binom{N}{n}^{-1} = \frac{n}{N} .$$

$p$ ist die gesuchte Wahrscheinlichkeit.

## Vorbereitung 2

Inwiefern kann man behaupten, dass die negative Binomialverteilung sowohl die geometrische Verteilung als auch die Binomialverteilung als Spezialfälle enthält?

### Lösung

Wir werden auf den Folien zur Zentralübung am 3. Juni alle damit zusammenhängenden Verteilungen darstellen.

## Tutoraufgabe 1

1. Zeigen Sie für alle $j \in \mathbb{N}$:

$$|\{(s_1,\ldots,s_j) \in \{1,\ldots,n\}^j\,;\, s_1+\ldots+s_j \leq n\}| = \binom{n}{j}\,.$$

2. Sie führen das folgende mehrstufige Experiment durch:

   In jedem Schritt wählen Sie zufällig und gleichverteilt eine Zahl zwischen 1 und $n$ aus, d. h., in jedem Schritt und für alle $i$ mit $1 \leq i \leq n$ ist die Wahrscheinlichkeit, dass Sie die Zahl $i$ ziehen, gleich $\frac{1}{n}$. Das Experiment endet, nachdem die Summe der gezogenen Zahlen das erste Mal größer als $n$ ist. Die Zufallsvariable $X$ gibt an, nach wie vielen Schritten das Experiment endet.

   Zeigen Sie:

$$\Pr[X \geq j+1] = \frac{\binom{n}{j}}{n^j} \quad \text{für alle } j = 0,1,\ldots\ .$$

3. Folgern Sie

$$\mathbb{E}[X] = \left(1+\frac{1}{n}\right)^n\,.$$

**Lösung**

Wir benützen die folgende Bezeichnung.

$$S_{j,n} = \{(s_1,\ldots,s_j) \in \{1,\ldots,n\}^j\,;\, s_1+\ldots+s_j \leq n\}\,.$$

1. Es gibt zwei alternative Ansätze für den Beweis von $|S_{j,n}| = \binom{n}{j}$.

   (a) Man stellt $n$ als Summe von Einsen dar und markiert davon $j$ Einsen. Zu einer gegebenen Markierung werden für alle $i$ mit $1 \leq i \leq j$ die Zahlen $s_i$ definiert als die Summe aller Einsen, die rechts von der Eins mit Nummer $i-1$ stehen bis einschließlich der Eins mit Nummer $i$. Auf diese Weise kann man alle möglichen Summen $s_1+\ldots+s_j \leq n$ darstellen. Die Markierungen lassen sich eineindeutig den $j$-Tupeln $(s_1,\ldots,s_j)$ zuordnen. Damit gilt für die Anzahl $anz_M(j,n)$ der Markierungen von $j$ Einsen unter den $n$ Einsen die Gleichung

$$anz_M(j,n) = |S_{j,n}|\,.$$

   Nun ist aber die Anzahl $anz_M(j,n)$ der Markierungen von $j$ Einsen unter den $n$ Einsen genau gleich der Anzahl von $i$-elementigen Teilmengen von $[n]$. D. h., es gilt

$$anz_M(j,n) = \binom{n}{j}\,.$$

   (b) Auch ein induktiver Ansatz führt zum Ziel.
   Für $j = 1$ gilt offensichtlich $|S_{j,n}| = n = \binom{n}{1}$.

Für jedes $s_{j+1} \in [n-j]$ gibt es genau $\binom{n-s_{j+1}}{j}$ passende $j$-Tupel $(s_1, \ldots, s_j)$. Damit folgt

$$\begin{aligned} |S_{j+1,n}| &= \sum_{i=1}^{n-j} \binom{n-i}{j} = \sum_{k=j}^{n-1} \binom{k}{j} \\ &= \binom{n}{j+1} . \end{aligned}$$

Wir erinnern uns, dass die letzte Gleichung als Berechnung der Summe der Binomialkoeffizienten in einer „Spalte“ des Pascalschen Dreiecks vom Diagonalelement bis einschließlich dem Eintrag in der $(n-1)$ten Zeile bekannt ist.

2. Für $j = 0$ ist die Gleichung offenbar richtig, denn die Werte sind auf beiden Seiten gleich 1.

   Wir betrachten die Elementarereignisse des Ziehens von $n$ Zahlen, d. h. wir setzen $\Omega = \{(s_1, s_2, \ldots, s_n)\,;\, s_i \in [n]\}$ mit $\Pr[x] = \frac{1}{n^n}$ mit $x \in \Omega$.

   Für $j > 0$ tritt das Ereignis $E :=$ „$X \geq j+1$“, d. h. $(s_1, s_2, \ldots, s_j, s_{j+1}, \ldots, s_n) \in E$, genau dann ein, wenn $(s_1, s_2, \ldots, s_j) \in S_{j,n}$ gilt. Da alle $s_k$ mit $j+1 \leq k \leq n$ beliebige Werte annehmen können, erhalten wir

$$|E| = |S_{j,n}| \cdot n^{n-j} .$$

   Es folgt

$$\Pr[X \geq j+1] = \frac{|S_{j,n}| \cdot n^{n-j}}{n^n} = \frac{\binom{n}{j}}{n^j} .$$

3. Zur Berechnung des Erwartungswertes von $X$ wenden wir Satz 34 der Vorlesung an.

$$\begin{aligned} \mathbb{E}[X] &= \sum_{i=1}^{\infty} \Pr[X \geq i] \\ &= \sum_{i=1}^{\infty} \frac{\binom{n}{i-1}}{n^{i-1}} = \sum_{i=0}^{\infty} \frac{\binom{n}{i}}{n^i} \\ &= \sum_{i=0}^{n} \frac{\binom{n}{i}}{n^i} = \sum_{i=0}^{n} \binom{n}{i} \left(\frac{1}{n}\right)^i \cdot 1^{n-i} \\ &= \left(1 + \frac{1}{n}\right)^n . \end{aligned}$$

## Tutoraufgabe 2

Wir nehmen eine zufällige Auswahl $P'$ eines Parameters $n \in \mathbb{N}$ mit Wahrscheinlichkeit $\Pr[n] = \frac{2}{3}(\frac{1}{3})^{n-1}$ an. Dann definieren wir einen Prozess $P''$ dadurch, dass zunächst $P'$ aufgerufen wird und der ausgegebene Wert als Eingabeparameter $n$ für den Aufruf von $P_n$ verwendet wird. Dabei wählt $P_n$ $n$-mal einen Buchstaben aus der Menge $\{a, b, c\}$ aus, und zwar ein $a$ bzw. $b$ bzw. $c$ mit Wahrscheinlichkeit $\frac{1}{2}$ bzw. $\frac{1}{3}$ bzw. $\frac{1}{6}$.

Berechnen Sie die Wahrscheinlichkeit, dass der Prozess $P''$ ein Wort $w$ ausgibt, das genau ein $a$ enthält. Geben Sie insbesondere den zugrundeliegenden Wahrscheinlichkeitsraum an.

**Lösung**

Die Menge der Ergebnisse von $P''$ ist $\Omega = \{a, b, c\}^*$. Als Wahrscheinlichkeit für $w \in \Omega \backslash \{\epsilon\}$ setzen wir

$$\Pr[w] = \frac{2}{3}\left(\frac{1}{3}\right)^{|w|-1} \cdot \left(\frac{1}{2}\right)^{|w|_a} \cdot \left(\frac{1}{3}\right)^{|w|_b} \cdot \left(\frac{1}{6}\right)^{|w|_c} .$$

Dabei bezeichnet $|w_x|$ die Anzahl der Buchstaben $x$ im Wort $w$ und $|w|$ die Länge von $w$. Natürlich wird $\Pr[\varepsilon] = 0$ gesetzt.

Mit Hilfe des Multinomialsatzes rechnet man leicht

$$1 \;=\; \sum_{w \in \Omega, |w| = n} \left(\frac{1}{2}\right)^{|w|_a} \cdot \left(\frac{1}{3}\right)^{|w|_b} \cdot \left(\frac{1}{6}\right)^{|w|_c} .$$

Wegen $\sum_{i=1}^{\infty} \frac{2}{3}\left(\frac{1}{3}\right)^i = 1$ folgt, dass $W = \langle \Omega, \Pr \rangle$ ein Wahrscheinlichkeitsraum ist.

Sei $E_{n,a} \subseteq \Omega$ das Ereignis, dass $P''$ ein Wort der Länge $n$ mit genau einem enthaltenen $a$ ausgibt, also

$$E_{n,a} = \{w \in \Omega\,;\, |w| = n, |w|_a = 1\} .$$

Dann gilt wegen $|w|_a = 1$

$$\begin{aligned}
\Pr[E_{n,a}] &= \sum_{w \in E_{n,a}} \Pr[w] \\
&= \frac{2}{3}\left(\frac{1}{3}\right)^{n-1} \cdot \sum_{1+|w|_b+|w|_c = n} \frac{n!}{1!|w|_b!|w|_c!} \cdot \left(\frac{1}{2}\right)^{1} \cdot \left(\frac{1}{3}\right)^{|w|_b} \cdot \left(\frac{1}{6}\right)^{|w|_c} \\
&= \frac{2}{3}\left(\frac{1}{3}\right)^{n-1} \cdot \left(\frac{1}{2}\right) \cdot n \cdot \sum_{|w|_b+|w|_c = n-1} \frac{(n-1)!}{|w|_b!|w|_c!} \cdot \left(\frac{1}{3}\right)^{|w|_b} \cdot \left(\frac{1}{6}\right)^{|w|_c} \\
&= \frac{2}{3}\left(\frac{1}{3}\right)^{n-1} \cdot \left(\frac{1}{2}\right) \cdot n \cdot \sum_{i=0}^{n-1} \binom{n-1}{i} \cdot \left(\frac{1}{3}\right)^{i} \cdot \left(\frac{1}{6}\right)^{n-1-i} \\
&= \frac{2}{3}\left(\frac{1}{3}\right)^{n-1} \cdot \left(\frac{1}{2}\right) \cdot n \cdot \left(\frac{1}{3} + \frac{1}{6}\right)^{n-1} \\
&= \frac{1}{3} \cdot n \cdot \left(\frac{1}{6}\right)^{n-1} .
\end{aligned}$$

Sei nun $(X_a = 1) \subseteq \Omega$ das Ereignis, dass $P''$ ein Wort mit genau einem enthaltenen $a$ ausgibt, also

$$(X_a = 1) = \{w \in \Omega\,; w \in E_{n,a}, n \in \mathbb{N}\} .$$

Dann gilt

$$
\begin{aligned}
\Pr[X_a = 1] &= \sum_{n=1}^{\infty} \Pr[E_{n,a}] \\
&= \sum_{n=1}^{\infty} \frac{1}{3} \cdot n \cdot \left(\frac{1}{6}\right)^{n-1} \\
&= \frac{1}{3} \cdot \frac{1}{(1-\frac{1}{6})^2} \\
&= \frac{12}{25}\,.
\end{aligned}
$$

## Tutoraufgabe 3

Mit einem fairen Würfel mit Augenzahlen $\{1,2,3,4,5,6\}$ wird wie folgt gespielt. Beim Start wird der Würfel 1 Mal geworfen. Die geworfene Augenzahl sei $x$. Dem Wurf geben wir die Nummer 0. Nun wird so lange gewürfelt, bis wieder $x$ erscheint. Die dabei (nach dem Wurf Nummer 0) geworfenen Augenzahlen $y$ mit $y \neq x$ werden addiert. Das Ergebnis sei die Zufallsvariable $Z$.

1. Sei $n \in \mathbb{N}$ und sei $X_i$ die Augenzahl im $i$-ten Wurf mit $1 \leq i \leq n-1$. Sei $Z_n = \sum_{i=1}^{n-1} X_i$. Berechnen Sie den Erwartungswert von $Z_n$ unter der Bedingung, dass beim Start die Augenzahl $x$ geworfen wurde und das Spiel im $n$-ten Schritt endet.

2. Mit welcher Wahrscheinlichkeit endet das Spiel im $n$-ten Schritt?

3. Geben Sie für $Z$ den Erwartungswert an.

4. Beschreiben Sie mit Hilfe der Faltung von Dichtefunktionen $f_{X_i}, i = 1,2,\ldots$ ein Verfahren zur Berechnung der Dichtefunktion $f_Z$.

**Lösung**

Wir schreiben kurz $SpE = n$ für „Spielende im $n$-ten Schritt".
Der Startwurf mit Nummer 0 sei $X_0$.

1. Sei $x \in \{1,2,3,4,5,6\}$ und $n \in \mathbb{N}$.
   Wir berechnen zunächst den bedingten Erwartungswert $\mathbb{E}[X_i | X_0 = x \wedge SpE = n]$ mit $1 \leq i \leq n-1$.

   Jeder Wert $y \neq x$ tritt bei $X_i$ auf mit Wahrscheinlichkeit

   $$Pr[X_i{=}y | X_i{\neq}x] = \frac{\Pr[X_i = y \wedge X_i \neq x]}{\Pr[X_i \neq x]} = \frac{\Pr[X_i = y]}{\Pr[X_i \neq x]} = \frac{1/6}{5/6} = \frac{1}{5}\,.$$

   Wir erhalten

   $$\mathbb{E}[X_i | X_0 = x \wedge SpE = n] \;=\; \sum_{(1 \leq y \leq 6) \wedge (y \neq x)} \frac{1}{5} \cdot y = \frac{21-x}{5}\,,$$

   und damit

   $$\mathbb{E}[Z_n | X_0 = x \wedge SpE = n] \;=\; (n-1)\frac{21-x}{5}\,.$$

2. Es gilt für alle Startwerte $x$ und alle $n \geq 1$

$$\Pr[SpE = n] = \left(\frac{5}{6}\right)^{n-1} \cdot \frac{1}{6}\,.$$

3. Wir bestimmen zunächst den Erwartungswert von $Z_n$ unter der Bedingung $SpE = n$.

$$\begin{aligned}
\mathbb{E}[Z_n|SpE = n] &= \sum_{x\in[1..6]} \mathbb{E}[Z_n|SpE = n \wedge X_0 = x] \cdot \Pr[X_0 = x] \\
&= \sum_{x\in[1..6]} (n-1)\frac{21-x}{5} \cdot \frac{1}{6} = (n-1) \cdot 21 \cdot \frac{1}{6}\,.
\end{aligned}$$

Durch Anwendung des Satzes der totalen Wahrscheinlichkeit erhalten wir

$$\begin{aligned}
\mathbb{E}[Z] &= \textstyle\sum_{n=1}^{\infty} \mathbb{E}[Z_n|SpE = n] \cdot \Pr[SpE = n] \\
&= \textstyle\sum_{n=1}^{\infty} (n-1) \cdot 21 \cdot \frac{1}{6} \cdot \left(\frac{5}{6}\right)^{n-1} \cdot \frac{1}{6} \\
&= \textstyle\frac{35}{72} \cdot \sum_{n=1}^{\infty} n \left(\frac{5}{6}\right)^{n-1} \\
&= \textstyle\frac{35}{72} \cdot \left(\frac{1}{1-\frac{5}{6}}\right)^2 \;=\; \frac{35}{2}\,.
\end{aligned}$$

4. Die Berechnung der Dichte $f_Z$ erfolgt in zwei Schritten. Wir betrachten den folgenden 2. Schritt zuerst. Sei $f_{Z_n}$ die Dichte von $Z_n$. Dann gilt nach dem Satz von der totalen Wahrscheinlichkeit für alle $i \in \mathbb{N}$

$$f_Z(i) = \sum_{n=1}^{\infty} f_{Z_n}(i) \cdot \Pr[SpE = n]\,.$$

Der erste Schritt ist nun die Angabe der Dichte $f_{Z_n}$ auf der Grundlage der bedingten Dichten $f_{X_i|X_0=x}$. Es gilt

$$f_{Z_n}(i) = \sum_{x=1}^{6} \frac{1}{6} f_{Z_n|X_0=x}(i)\,.$$

Wir bemerken, dass $Z_n$ unter der Bedingung $X_0 = x$ eine Summe der unabhängigen Variablen $X_i$ für $1 \leq i \leq n-1$ ist. Damit folgt mit Konvolution $\otimes$ der bedingten Dichten $f_{X_i|X_0=x}$ nach Satz der Vorlesung

$$f_{Z_n|X_0=x} = f_{X_1|X_0=x} \otimes f_{X_2|X_0=x} \otimes \ldots f_{X_{n-1}|X_0=x}\,.$$

Prinzipiell kann nun $f_Z(i)$ per Programm approximativ bestimmt werden.