Informationstheorie, FS 2021
ETH Zürich, Dept. Informatik

Prof. Joachim M. Buhmann
Fabian Laumer

Lösungen zu Übungsblatt 1

1.1

Es seien A, B und C drei Ereignisse aus einer Menge Ω von Elementarereignissen. Drücken Sie mit Hilfe von Mengenoperationen (d.h. Vereinigung, Schnitt, Komplementbildung) die folgenden Ereignisse aus:

a) A_1 : Keines der drei Ereignisse A, B und C tritt ein.

b) A_2 : Mindestens eines tritt ein.

c) A_3 : Genau eines tritt ein.

d) A_4 : Mindestens zwei treten ein.

e) A_5 : Mindestens eines tritt nicht ein.

f) A_6 : Höchstens zwei treten ein.

- $A_1 = (A \cup B \cup C)^c$
- $A_2 = A \cup B \cup C$
- $A_3 = (A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C)$
- $A_4 = (A \cap B) \cup (A \cap C) \cup (B \cap C)$
- $A_5 = (A \cap B \cap C)^c$
- $A_6 = A_5$

1.2

Wir modellieren einen unfairen Würfel durch folgende Verteilung auf $\{1, 2, 3, 4, 5, 6\}$:

X	1	2	3	4	5	6
P_X	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$

Berechnen Sie die Erwartungswerte der Zufallsvariablen X, X^2 und $P_X(X)$.

- $E[Y_1] = E[X] = \sum_{x=1}^{6} x \cdot P_X(x) = \frac{17}{8}$
- $E[Y_2] = E[X^2] = \sum_{x=1}^{6} x^2 \cdot P_X(x) = \frac{55}{8}$
- $E[Y_3] = E[P_X(X)] = \sum_{x=1}^{6} P_X(x) \cdot P_X(x) = \frac{21}{64}$

1.3

Nehmen Sie Stellung zum Wahrheitsgehalt der folgenden Aussage: Beim dreimaligen Würfeln mit einem fairen Würfel sind die Ereignisse 'die Augensumme ist 11' und 'die Augensumme ist 12' gleich wahrscheinlich, denn beide Ereignisse können auf genau sechs verschiedene Arten dargestellt werden:

$$11 = 6+4+1 = 6+3+2 = 5+5+1 = 5+4+2 = 5+3+3 = 4+4+3,$$
$$12 = 6+5+1 = 6+4+2 = 6+3+3 = 5+5+2 = 5+4+3 = 4+4+4.$$

Die Aussage berücksichtigt nicht, dass die einzelnen Würfelergebnisse nicht gleich wahrscheinlich sind. Würde die Reihenfolge eine Rolle spielen, so wären bei dreifachem Würfeln 6^3 verschiedene Kombinationen möglich, so dass jeder die Wahrscheinlichkeit 6^{-3} zukäme. Da die Reihenfolge jedoch unerheblich ist, besitzen die Ergebnisse mit drei verschiedenen Augenzahlen die Wahrscheinlichkeit $6 \cdot 6^{-3}$, da sie in sechs verschiedenen Reihenfolgen vorkommen können, diejenigen mit genau zwei identischen Augenzahlen eine Wahrscheinlichkeit von $3 \cdot 6^{-3}$ und die Dreierpaschs die Wahrscheinlichkeit 6^{-3}. Insgesamt erhält man somit für das Ereignis "Augensumme 11" die Wahrscheinlichkeit $27 \cdot 6^{-3}$, für das Ereignis "Augensumme 12" dagegen die Wahrscheinlichkeit $25 \cdot 6^{-3}$.

1.4

Berechnen Sie die Wahrscheinlichkeit dafür, dass in einem gut gemischten Stapel von 32 Skatkarten die 4 Asse direkt aufeinander liegen. Geben Sie explizit den genutzen Wahrscheinlichkeitsraum an.

Als Wahrscheinlichkeitsraum kann man die Menge Σ_{32} der Permutationen (d.h. der bijektiven Selbstabbildungen) von $\{1, 2, \ldots, 32\}$ mit der Gleichverteilung wählen; es gilt dann $p(\sigma) = \frac{1}{32!}$ für jedes $\sigma \in \Sigma_{32}$, da $|\Sigma_{32}| = 32!$. Jede Permutation entspricht dabei einer möglichen Reihenfolge der 32 Karten. Das Ereignis "Alle 4 Asse folgen direkt aufeinander" tritt ein mit Wahrscheinlichkeit

$$\frac{4! \cdot 28! \cdot 29}{32!} = 29 \cdot \binom{32}{4}^{-1},$$

denn es gibt $4! \cdot 28!$ Permutationen, die vier gegebene Zahlen (entsprechend den vier Assen) auf einen gegebenen Block von vier aufeinanderfolgenden Zahlen abbilden, und 29 solcher Blöcke.

1.5

Wir betrachten ein Testverfahren für eine Krankheit, an der eine Person mit Wahrscheinlichkeit $p_K = 0.00001$ erkrankt, d.h. es tritt durchschnittlich ein Erkrankter auf 100.000 Einwohner auf. Das Verfahren hat eine Fehlerwahrscheinlichkeit von $p_F = 0.01$, d.h. das Testergebnis ist mit Wahrscheinlichkeit 0.99 korrekt (sowohl im positiven wie im negativen Fall).

a) Alice lässt sich testen und der Test fällt positiv aus. Mit welcher Wahrscheinlichkeit leidet Alice trotz des Testergebnisses nicht an der Krankheit?

b) Bobs Test fällt negativ aus. Mit welcher Wahrscheinlichkeit leidet Bob dennoch an der Krankheit?

Vergleichen Sie mit den jeweiligen Wahrscheinlichkeiten vor dem Test.

a) Wir definieren die Zufallsvariablen K und T: $K = 1$ falls Alice krank ist, sonst $K = 0$. $T = 1$ falls das Testergebnis von Alice positiv ist, sonst $T = 0$. Gesucht ist die bedingte Wahrscheinlichkeit

$$\begin{aligned} P(K=0|T=1) &= \frac{P(K=0,T=1)}{P(T=1)} = \frac{P(K=0)\cdot P(T=1|K=0)}{0.99999\cdot 0.01 + 0.00001\cdot 0.99} \\ &= \frac{0.99999\cdot 0.01}{0.0100098} = \frac{99999}{100098} \approx 0.99901097. \end{aligned}$$

Die A-priori-Wahrscheinlichkeit liegt bei 0.99999. Ein positives Testergebnis braucht Alice also nicht stark zu beunruhigen.

b) K und T seien wie oben für Bob definiert. Gesucht ist die bedingte Wahrscheinlichkeit

$$\begin{aligned} P(K=1|T=0) &= \frac{P(K=1,T=0)}{P(T=0)} = \frac{P(K=1)\cdot P(T=0|K=1)}{0.99999\cdot 0.99 + 0.00001\cdot 0.01} \\ &= \frac{0.00001\cdot 0.01}{0.9899902} = \frac{1}{9899902} \approx 0.0000001010111. \end{aligned}$$

Die A-priori-Wahrscheinlichkeit liegt bei 0.00001. Ein negatives Testergebnis erhöht Bobs Gelassenheit um etwas weniger als den Faktor 100.

1.6

Die folgende Tabelle zeigt die gemeinsame Wahrscheinlichkeitsverteilung der Zufallsvariablen $K =$ "Wetter in Kanada" und $S =$ "Wetter in der Schweiz". Beide Zufallsvariablen nehmen Werte in der Menge $\{h, w, k, e\}$ an ("heiss", "warm", "kühl","eiskalt").

Die folgende Tabelle gibt die gemeinsame Verteilung an:

		S			
	P_{KS}	h	w	k	e
	h	0.17	0.12	0.03	0
K	w	0.04	0.09	0.06	0.01
	k	0.01	0.06	0.10	0.03
	e	0	0.03	0.11	0.14

a) Ist P_{KS} eine Wahrscheinlichkeitsverteilung?

b) Berechnen Sie die Verteilungen P_S und P_K.

c) Sind die Ereignisse $\{K = k\}$ und $\{S = w\}$ unabhängig?

d) Sind die Zufallsvariablen K und S unabhängig?

e) Wie gross ist die Wahrscheinlichkeit, dass es in Kanada warm ist, gegeben dass es in der Schweiz heiss ist?

f) Angenommen Sie mögen warmes Wetter und wissen, dass es nicht eiskalt sein wird in Kanada. Wo sollten Sie ihre Ferien verbringen (Kanada oder Schweiz)?

a) Ja, weil die Werte der Tabelle sich auf 1 summieren.

b) Die Verteilungen von P_S und P_K sind durch die folgende Tabelle gegeben:

P_S	h	w	k	e
	0.22	0.3	0.3	0.18

P_K	h	w	k	e
	0.32	0.2	0.2	0.28

c) Die Ereignisse $K = k$ und $S = w$ sind unabhängig, weil

$$P[K = k, S = w] = P[K = k] \cdot P[S = w].$$

$$\begin{aligned} P[K = k, S = w] &= 0.06 \\ P[K = k] \cdot P[S = w] &= 0.2 \cdot 0.3 = 0.06 \end{aligned}$$

d) Die Zufallsvariablen K und S sind nicht unabhängig, weil

$$P[K = e, S = h] \neq P[K = e] \cdot P[S = h].$$

$$\begin{aligned} P[K = e, S = h] &= 0 \\ P[K = e] \cdot P[S = h] &= 0.28 \cdot 0.22 = 0.0616 \end{aligned}$$

e) Es gilt

$$P[K = w \,|\, S = h] = \frac{P[K = w, S = h]}{P[S = h]} = \frac{0.04}{0.22} \approx 0.1818$$

f) Sie sollten in der Schweiz bleiben, weil $P[S = w \,|\, K \neq e] > P[K = w \,|\, K \neq e]$

$$\begin{aligned} P[K = w \,|\, K \neq e] &= \frac{P[K = w, K \neq e]}{P[K \neq e]} \\ &= \frac{P[K = w]}{P[K \neq e]} \\ &= \frac{0.2}{1 - 0.28} \approx 0.2778 \end{aligned}$$

und

$$\begin{aligned} P[S = w \,|\, K \neq e] &= \frac{P[S = w, K \neq e]}{P[K \neq e]} \\ &= \frac{0.12 + 0.09 + 0.06}{1 - 0.28} = 0.375 \end{aligned}$$

1.7

Angenommen, wir wenden den $(r = 3)$-Wiederholcode für die Datenübertragung über den binären symmetrischen Kanal mit Fehlerwahrscheinlichkeit $0 < p < \frac{1}{2}$. Dessen Decoder $\{0,1\}^3 \to \{0,1\}$, $\mathbf{r} \mapsto \hat{s}(\mathbf{r})$, bildet $\mathbf{r}$ auf das häufiste in $\mathbf{r}$ vorkommende Symbol ab (siehe Vorlesung bzw. [MacKay, Kapitel 1]). Zeigen Sie, dass $\hat{s}(\mathbf{r}) = \mathrm{argmax}_s P(\mathbf{r}|s)$ gilt.

Wenn der Sender $s \in \{0,1\}$ sendet, ist die Wahrscheinlichkeit dafür, dass der Empfänger ein bestimmtes $\mathbf{r}$ empfängt, gegeben durch

$$P(\mathbf{r}|s) = p^{3-k}(1-p)^k,$$

wobei $k \in \{0,1,2,3\}$ die Anzahl der Bits r_i mit $r_i = s$ angibt. Wegen $p < \frac{1}{2} < 1 - p$ wird dieser Ausdruck bei gegebenem $\mathbf{r}$ maximiert durch das $s \in \{0,1\}$ mit dem grösseren k, d.h. das in $\mathbf{r}$ häufiger vorkommende s; nach Definition des Decodierers ist das gleich $\hat{s}(\mathbf{r})$, und damit folgt $\hat{s}(\mathbf{r}) = \mathrm{argmax}_s P(\mathbf{r}|s)$.