ETH Zürich
Dept. Informatik

Prof. Markus Gross
WS 05/06

Informationstheorie

Lösung 1

1.1 Erwartungswerte

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

1.2 Bedingte Wahrscheinlichkeiten

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.

1.3 Gemeinsame Wahrscheinlichkeitsverteilungen

$$\begin{array}{cc|cccc} & & & S & & \\ & P_{KS} & h & w & k & e \\ \hline & h & 0.17 & 0.12 & 0.03 & 0 \\ K & w & 0.04 & 0.09 & 0.06 & 0.01 \\ & k & 0.01 & 0.06 & 0.10 & 0.03 \\ & e & 0 & 0.03 & 0.11 & 0.14 \end{array}$$

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

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

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

Die Zufallsvariablen K und S sind nicht unabhängig, weil

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

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

d) Es gilt

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

e) 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

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

1.4 Kenngrössen bei Summen von Zufallsvariablen

a) Wir zeigen die Behauptung zuerst für $n = 2$:

$$\begin{aligned}
E[X_1 + X_2] &= \sum_{x_1,x_2} (x_1 + x_2) P_{X_1 X_2}(x_1, x_2) \\
&= \sum_{x_1,x_2} x_1 P_{X_1 X_2}(x_1, x_2) + \sum_{x_1,x_2} x_2 P_{X_1 X_2}(x_1, x_2) \\
&= \sum_{x_1} x_1 \sum_{x_2} P_{X_1 X_2}(x_1, x_2) + \sum_{x_2} x_2 \sum_{x_1} P_{X_1 X_2}(x_1, x_2) \\
&= \sum_{x_1} x_1 P_{X_1}(x_1) + \sum_{x_2} x_2 P_{X_2}(x_2) \\
&= E[X_1] + E[X_2].
\end{aligned}$$

Für $n > 2$ folgt die Behauptung aus der Tatsache, dass man $X_2 + \ldots + X_n$ als eine Zufallsvariable interpretieren und daher gemäss obigem Verfahren X_1 „abspalten“ kann.

b) Wir zeigen die Behauptung zuerst für $n = 2$:

$$\begin{aligned}
\mathrm{Var}[X_1 + X_2] &= \sum_{x_1,x_2} \left((x_1 + x_2) - E[X_1 + X_2]\right)^2 P_{X_1X_2}(x_1, x_2) \\
&= \sum_{x_1,x_2} \left(x_1 - E[X_1] + x_2 - E[X_2]\right)^2 P_{X_1}(x_1)P_{X_2}(x_2) \\
&= \sum_{x_1,x_2} \left(x_1 - E[X_1]\right)^2 P_{X_1}(x_1)P_{X_2}(x_2) \\
&\quad + \sum_{x_1,x_2} \left(x_2 - E[X_2]\right)^2 P_{X_1}(x_1)P_{X_2}(x_2) \\
&\quad + \sum_{x_1,x_2} 2\left(x_1 - E[X_1]\right)\left(x_2 - E[X_2]\right) P_{X_1}(x_1)P_{X_2}(x_2) \\
&= \sum_{x_1} \left(x_1 - E[X_1]\right)^2 P_{X_1}(x_1) \\
&\quad + \sum_{x_2} \left(x_2 - E[X_2]\right)^2 P_{X_2}(x_2) \\
&\quad + \sum_{x_1,x_2} 2x_1x_2 P_{X_1}(x_1)P_{X_2}(x_2) \\
&\quad + \sum_{x_1,x_2} 2E[X_1]E[X_2]P_{X_1}(x_1)P_{X_2}(x_2) \\
&\quad - \sum_{x_1,x_2} 2x_1E[X_2]P_{X_1}(x_1)P_{X_2}(x_2) \\
&\quad - \sum_{x_1,x_2} 2x_2E[X_1]P_{X_1}(x_1)P_{X_2}(x_2) \\
&= \mathrm{Var}[X_1] + \mathrm{Var}[X_2] \\
&\quad +2\sum_{x_1} x_1P_{X_1}(x_1) \sum_{x_1,x_2} x_2P_{X_2}(x_2) \\
&\quad +2E[X_1]E[X_2] \\
&\quad -\sum_{x_1} 2x_1E[X_2]P_{X_1}(x_1) \\
&\quad -\sum_{x_2} 2x_2E[X_1]P_{X_2}(x_2) \\
&= \mathrm{Var}[X_1] + \mathrm{Var}[X_2] \\
&\quad +2E[X_1]E[X_2] \\
&\quad +2E[X_1]E[X_2] \\
&\quad -2E[X_1]E[X_2] \\
&\quad -2E[X_1]E[X_2] \\
&= \mathrm{Var}[X_1] + \mathrm{Var}[X_2].
\end{aligned}$$

Für $n > 2$ folgt die Behauptung gemäss Teilaufgabe a.

1.5 Ausblick Codierung von Information

Es ist klar, dass der erste Text sich sehr schwer komprimieren lässt, da die zugrundeliegende Verteilung uniform ist. Der Unterschied zwischen dem zweiten und dritten Text ist weniger klar. Theoretisch ist der dritte Text ein bisschen besser komprimierbar, da seine sogenannte Entropie kleiner ist.

$$H(T_1) = 3 \;>\; H(T_2) = 1.91664 \;>\; H(T_3) = 1.91129$$