Informationstheorie, FS 2021 Prof. Joachim M. Buhmann
ETH Zürich, Dept. Informatik Fabian Laumer

Lösungen zu Übungsblatt 2

2.1

Sei X eine Zufallsvariable mit der folgenden Verteilung:

$$P_X(1) = \frac{1}{2}, \quad P_X(2) = \frac{1}{12} + \varepsilon, \quad P_X(3) = \frac{5}{12} - \varepsilon.$$

Skizzieren Sie $H(X)$ als Funktion von ε. Bestimmen Sie insbesondere das Maximum von $H(X)$.

Wir schreiben $H(p_1, \dots, p_n)$ für die Entropie einer Verteilung, deren Elementarereignisse die Wahrscheinlichkeiten $p_1, \dots, p_n$ haben (vgl. Aufgabe 2.1). Die Entropie von X ist

$$\begin{aligned} H(X) &= H(\tfrac{1}{2}, \tfrac{1}{12} + \varepsilon, \tfrac{5}{12} - \varepsilon) \\ &= -\tfrac{1}{2}\log(\tfrac{1}{2}) - (\tfrac{1}{12} + \varepsilon)\log(\tfrac{1}{12} + \varepsilon) - (\tfrac{5}{12} - \varepsilon)\log(\tfrac{5}{12} - \varepsilon) \\ &= \tfrac{1}{2} - (\tfrac{1}{12} + \varepsilon)\log(\tfrac{1}{12} + \varepsilon) - (\tfrac{5}{12} - \varepsilon)\log(\tfrac{5}{12} - \varepsilon) \end{aligned}$$

Zunächst sieht man, dass P_X nur für $-\frac{1}{12} \le \varepsilon \le \frac{5}{12}$ eine Wahrscheinlichkeitsverteilung ist. Für die Intervallgrenzen ist die Entropie

$$\begin{aligned} H(X|_{\varepsilon=-\frac{1}{12}}) &= H(\tfrac{1}{2}, 0, \tfrac{1}{2}) = 1 \\ H(X|_{\varepsilon=\frac{5}{12}}) &= H(\tfrac{1}{2}, \tfrac{1}{2}, 0) = 1 \end{aligned}$$

Um das Maximum zu finden, leiten wir diesen Ausdruck ab und finden die Nullstelle:

$$\tfrac{dH(X)}{d\varepsilon} = -\log(\tfrac{1}{12} + \varepsilon) + \log(\tfrac{5}{12} - \varepsilon)$$

Dieser Ausdruck ist Null für $\varepsilon = \frac{1}{6}$. Das ist intuitiv einleuchtend, da für dieses ε die Verteilung von X am nächsten an der uniformen Verteilung ist. Das Maximum ist

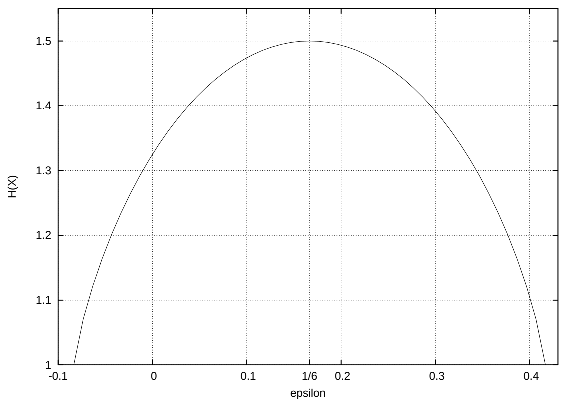
$$H(X|_{\varepsilon=\frac{1}{6}}) = H(\tfrac{1}{2}, \tfrac{1}{4}, \tfrac{1}{4}) = -\tfrac{1}{2}\log(\tfrac{1}{2}) - \tfrac{1}{4}\log(\tfrac{1}{4}) - \tfrac{1}{4}\log(\tfrac{1}{4}) = \tfrac{3}{2}.$$

2.2

Im Folgenden schreiben wir $H(p_1, p_2, \dots)$ für die Entropie einer (nicht unbedingt endlichen) Verteilung, deren Elementarereignisse die Wahrscheinlichkeiten $p_1, p_2, \dots$ haben. Zeigen Sie folgende "Zerlegungseigenschaft" der Entropie:

$$H(p_1, p_2, \dots) = H(p_1, 1 - p_1) + (1 - p_1)H\big(\tfrac{p_2}{1-p_1}, \tfrac{p_3}{1-p_1}, \dots\big)$$

Interpretieren Sie diese Eigenschaft.



Es gilt

$$\begin{aligned}
H(p_1, p_2, \dots) &= -\sum_{i\geq 1} p_i \log p_i \\
&= -p_1 \log p_1 - (1-p_1)\log(1-p_1) - \sum_{i\geq 2} p_i(\log p_i - \log(1-p_1)) \\
&= -p_1 \log p_1 - (1-p_1)\log(1-p_1) - (1-p_1)\sum_{i\geq 2} \frac{p_i}{1-p_1} \log \frac{p_i}{1-p_1} \\
&= H(p_1, 1-p_1) + (1-p_1) H(\frac{p_2}{1-p_1}, \frac{p_3}{1-p_1}, \dots)
\end{aligned}$$

Dabei haben wir im ersten Umformungsschritt den hinzugefügten Term $-(1-p_1)\log(1-p_1)$ in der Summe wieder abgezogen, unter Ausnutzung von $\sum_{i\geq 2} p_i = 1 - p_1$.

Interpretation: Wir können die Unsicherheit bzgl. des Wertes einer Zufallsvariablen mit der Verteilung $(p_1, p_2, \dots)$ aufspalten als Summe der Unsicherheit darüber, ob ein bestimmter Wert x_1 angenommen wird oder nicht, und der verbleibenden Unsicherheit über alle restlichen Werte, wenn wir wissen, dass x_1 nicht angenommen wurde.

Die Zerlegungseigenschaft ergibt sich übrigens auch unmittelbar aus der Kettenregel $H(X,Y) = H(X) + H(Y|X)$, wenn man für X die Indikatorvariable $I_{\{x_1\}}$ wählt.

2.3

Eine faire Münze wird geworfen, bis zum ersten Mal "Kopf" erscheint. X sei die Zufallsvariable, die angibt, beim wievielten Wurf das der Fall ist. Berechnen Sie $H(X)$ auf zwei Arten. 1): Direkte Berechnung der Entropie über die Defintion der Entropie einer Zufallsvariable. 2): Durch Anwednung der SZerlegeigenschaft"der Entropie.

Das Eintreten des Ereignisses $\{X = n\}$ bedeutet, dass bei den ersten $n-1$ Würfen "Zahl" fällt, und beim n-ten Wurf "Kopf". Da bei jedem Wurf "Kopf" und "Zahl" jeweils Wahrscheinlichkeit $\frac{1}{2}$ haben, gilt folglich

$$P(X = n) = (\tfrac{1}{2})^{n-1} \cdot \tfrac{1}{2} = \tfrac{1}{2^n}$$

Die Entropie von X ist daher

$$H(X) = -\sum_{i=1}^{\infty} \frac{1}{2^n} \log \frac{1}{2^n} = -\sum_{i=1}^{\infty} \frac{1}{2^n} \cdot n$$

a) Mit Hilfe der Identität $\sum_{n=1}^{\infty} n f^n = \frac{f}{(1-f)^2}$.
Mit Hilfe der angegeben Identität ergibt sich direkt $H(X) = 2$.

b) Mit Hilfe der Zerlegungseingeschaft aus der vorigen Aufgabe.
Wir können $H(X)$ auch schreiben als $H(X) = H(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots)$. Mit Hilfe der Zerlegungseigenschaft erhalten wir

$$H(X) = H(\tfrac{1}{2}, \tfrac{1}{4}, \tfrac{1}{8}, \dots) = H(\tfrac{1}{2}, \tfrac{1}{2}) + \tfrac{1}{2} H(\tfrac{1}{2}, \tfrac{1}{4}, \tfrac{1}{8}, \dots) = H(\tfrac{1}{2}, \tfrac{1}{2}) + \tfrac{1}{2} H(X).$$

Aufösen nach $H(X)$ liefert $H(X) = 2H(\frac{1}{2}, \frac{1}{2}) = 2 \cdot 1 = 2$.

2.4

Seien X und Y Zufallsvariablen mit $Y = f(X)$ für eine Funktion f. Können Sie allein mit dieser Information den Wert von $H(Y|X)$ bestimmen? Können Sie eine nötige und hinreichende Bedingung an X und Y dafür angeben, dass $H(Y|X)$ diesen Wert annimmt?

Für jedes x mit $P_X(x) > 0$ ist $P_{Y|X=x}(y) = 1$ falls $y = f(x)$ und $P_{Y|X=x}(y) = 0$ falls $y \neq f(x)$. Die auf $\{X = x\}$ bedingte Entropie von Y ist daher $H(Y|X = x) = H(1, 0) = 0$. Wegen $H(Y|X) = E_{P_X}(H(Y|X = x))$ ist damit auch $H(Y|X) = 0$. Das ist auch intuitiv klar: Da Y eine Funktion von X ist, besteht keine Unsicherheit mehr bzgl. des Werts von Y sobald der von X bekannt ist.

Nehmen wir nun an, dass $H(Y|X) = 0$ gilt. Dann ist für jedes x mit $P(x) > 0$ die bedingte Entropie $H(Y|X = x) = 0$, und daher gibt es ein eindeutig bestimmtes $y =: f(x)$, so dass $P_{Y|X=x}(y) = 1$ gilt (siehe Vorlesung). Mit der so definierten Funktion f betrachten wir die Menge $f(X(\Omega)) = \{y \mid \exists x : y = f(x)\}$. Es gilt

$$\begin{aligned}
P_Y(f(X(\Omega))) &= \sum_{y \in f(X(\Omega))} P_Y(y) \\
&= \sum_{x, P_X(x) > 0} P_X(x) \sum_{y \in f(X(\Omega))} P_{Y|X=x}(y) \\
&= \sum_{x, P_X(x) > 0} P_X(x) \\
&= 1.
\end{aligned}$$

Folglich gibt es für jedes y mit $P_Y(y) > 0$ ein x, so dass $y = f(x)$.

Eine nötige und hinreichende Bedingung für $H(Y|X) = 0$ ist deswegen die Existenz einer Funktion $f : \mathcal{X} \to \mathcal{Y}$, so dass für jedes $y \in \mathcal{Y}$ mit $P_Y(y) > 0$ ein $x \in \mathcal{X}$ existiert, so dass $y = f(x)$ ist. (Genau genommen ist diese Bedingung schwächer als die im ersten Teil geforderte Bedingung $Y = f(X)$; es ist aber leicht zu sehen, dass sie ebenfalls $H(Y|X) = 0$ impliziert.)

2.5

Wir betrachten zwei unabhängige Würfel. Seien die Zufallsvariablen X und Y die jeweils gewürfelten Augenzahlen. Weiter seien die Zufallsvariablen S und T definiert als $S = X + Y$ und $T = X + 2Y$. Bestimmen Sie folgende Ausdrücke mit möglichst geringem Rechenaufwand: $H(S)$, $H(T)$, $H(X|S)$, $I(X;S)$, $I(Y;S)$, $I(X;T)$ und $H(S|X)$.

Die Wahrscheinlichkeitsverteilungen der Zufallsvariablen S und T können aus den Wahrscheinlichkeitsverteilungen von X und Y ($P_X(i) = P_Y(i) = 1/6$ für alle $i \in \{1,\ldots,6\}$) abgeleitet werden:

s	2	3	4	5	6	7	8	9	10	11	12
$P_S(s)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

t	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$P_T(t)$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{3}{36}$	$\frac{3}{36}$	$\frac{3}{36}$	$\frac{3}{36}$	$\frac{3}{36}$	$\frac{3}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{1}{36}$

Es gilt $H(X,Y) = H(X) + H(Y)$ (da X und Y unabhängig sind) und $H(X) = H(Y) = H(\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6}) \approx 2.58$. Aus $S = X + Y$ und $Y = S - X$ folgt $H(X,Y) = H(X,S)$, denn aus X und Y lassen sich X und S berechnen und umgekehrt. Analog gilt auch $H(X,T) = H(X,Y)$. Mit diesen Überlegungen lassen sich die gefragten Grössen einfach berechnen.

- $H(S) = H(\frac{1}{36},\frac{2}{36},\frac{3}{36},\frac{4}{36},\frac{5}{36},\frac{6}{36},\frac{5}{36},\frac{4}{36},\frac{3}{36},\frac{2}{36},\frac{1}{36}) \approx 3.27$.
- $H(T) = H(\frac{1}{36},\frac{1}{36},\frac{2}{36},\frac{2}{36},\frac{3}{36},\frac{3}{36},\frac{3}{36},\frac{3}{36},\frac{3}{36},\frac{3}{36},\frac{3}{36},\frac{3}{36},\frac{2}{36},\frac{2}{36},\frac{1}{36},\frac{1}{36}) \approx 3.89$.
- $H(X|S) = H(X,S) - H(S) = H(X,Y) - H(S) = 2H(X) - H(S) \approx 1.90$.
- $I(X;S) = H(X) - H(X|S) \approx 0.69$.
- $I(Y;S) = I(X;S) \approx 0.69$.
- $I(X;T) = H(X) - H(X|T) = H(X) - H(X,T) + H(T) = H(X) - H(X,Y) + H(T) = H(X) - 2H(X) + H(T) = H(T) - H(X) \approx 1.31$.
- $H(S|X) = H(S,X) - H(X) = H(X,Y) - H(X) = H(X) \approx 2.58$.

2.6

Zeigen, dass die Kullback-Leibler-Divergenz $D_{KL}(P\|Q)$ zwischen zwei Verteilungen als eine Verallgemeinerung der wechselseitigen Information $I(X;Y)$ zwischen zwei Zufallsvariablen aufgefasst werden kann.

Die wechselseitige Information zweier Zufallsvariablen X und Y ist gegeben durch

$$\begin{aligned} I(X;Y) &= -\sum_{x,y} P_{XY}(x,y)\log\frac{P_X(x)P_Y(y)}{P_{XY}(x,y)} \\ &= D_{KL}(P_{XY}\|P_X P_Y), \end{aligned}$$

d.h. sie ist gleich der Kullback-Leibler-Divergenz zwischen der gemeinsamen Verteilung P_{XY} und $P_X P_Y$, dem Produkt der Marginalverteilungen von X und Y.

2.7

Beweisen Sie die Positivität der Kullback-Leibler-Divergenz (*Gibbs-Ungleichung*): Für alle Verteilungen $P, Q : \Omega \to [0,1]$ gilt $D_{KL}(P\|Q) \geq 0$, mit Gleichheit genau dann, wenn $P = Q$ ist. Folgern Sie daraus und aus dem Ergebnis der letzten Aufgabe, dass $I(X;Y) \geq 0$ ist mit Gleichheit genau dann, wenn X, Y unabhängig sind.

Es gilt

$$\begin{aligned} D_{KL}(P\|Q) &= -\sum_x P(x) \log \frac{Q(x)}{P(x)} \\ &= E_P\Big(-\log \frac{Q}{P}\Big) \\ &\geq -\log\Big(E_P \frac{Q}{P}\Big) \\ &= 0, \end{aligned}$$

wobei wir die Jensen-Ungleichung und die Konvexität der Funktion $z \mapsto -\log z$ verwendet haben. Da diese Funktion sogar *strikt* konvex ist, kann Gleichheit nur eintreten, wenn $-\log \frac{P}{Q}$ gleich einer Konstanten ist; da P und Q Verteilungen sind, muss diese Konstante gleich 1 sein, d.h. $P = Q$. Die Aussagen über $I(X;Y)$ ergeben sich direkt aus $I(X;Y) = D_{KL}(P_{XY}\|P_X P_Y)$.