ETH Zürich
Dept. Informatik

Prof. Markus Gross
WS 06/07

Informationstheorie
Lösung 3

3.1 Entropie

Die Entropie von X ist

$$\begin{aligned} H(X) &= H([1/2,\, 1/12+\epsilon,\, 5/12-\epsilon]) \\ &= -1/2\log_2(1/2) - (1/12+\epsilon)\log_2(1/12+\epsilon) - (5/12-\epsilon)\log_2(5/12-\epsilon) \\ &= 1/2 - (1/12+\epsilon)\log_2(1/12+\epsilon) - (5/12-\epsilon)\log_2(5/12-\epsilon) \end{aligned}$$

Als erstes sieht man, dass P_X nur für $-1/12 \leq \epsilon \leq 5/12$ eine Wahrscheinlichkeitsverteilung ist. Für die Intervallgrenzen ist die Entropie

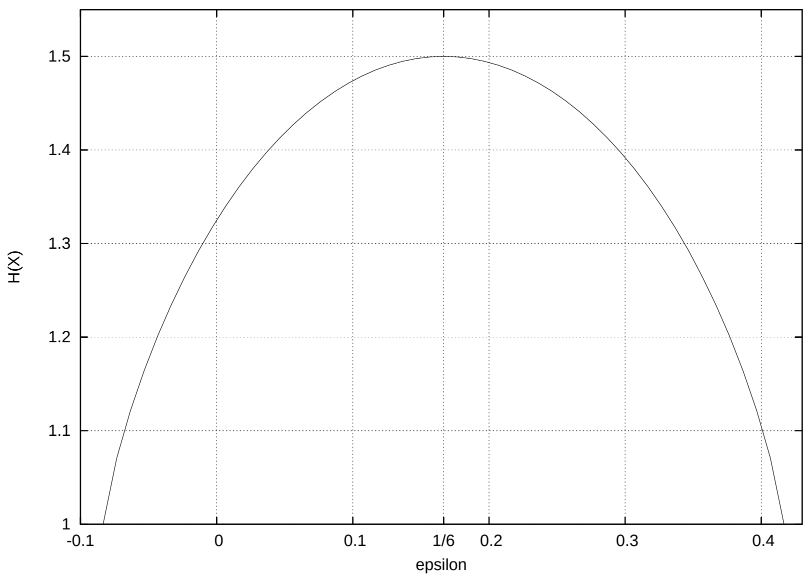
$$\begin{aligned} H(X|_{\epsilon=-1/12}) &= H([1/2, 0, 1/2]) = 1 \\ H(X|_{\epsilon=5/12}) &= H([1/2, 1/2, 0]) = 1 \end{aligned}$$

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

$$\frac{dH(X)}{d\epsilon} = -\log_2(1/12+\epsilon) + \log_2(5/12-\epsilon)$$

Dieser Ausdruck ist Null für $\epsilon = 1/6$. Das ist nicht weiter erstaunlich, da für dieses ϵ die Verteilung von X am nächsten an der uniformen Verteilung ist. Das Maximum ist

$$H(X|_{\epsilon=1/6}) = H([1/2, 1/4, 1/4]) = -1/2\log_2(1/2) - 1/4\log_2(1/4) - 1/4\log_2(1/4) = 3/2.$$



3.2 Forderungen an ein Unsicherheitsmass

a) Es gilt
$$\max\{p_1, \ldots, p_n, 0\} = \max\{p_1, \ldots, p_n\}$$
weil $p_i \geq 0$ für alle $i \in \{1, \ldots, n\}$ bei einer Wahrscheinlichkeitsverteilung. Daraus folgt sofort die behauptete Identität
$$H_\infty([p_1, \ldots, p_n, 0]) = H_\infty([p_1, \ldots, p_n]).$$

b) Die Verteilung von X ist gegeben durch $[p_1, \ldots, p_L] = [\frac{1}{L}, \ldots, \frac{1}{L}]$ und diejenige von Y durch $[q_1, \ldots, q_{L+1}] = [\frac{1}{L+1}, \ldots, \frac{1}{L+1}]$. Daher gilt
$$\max_{1 \leq i \leq L} p_i = \frac{1}{L} \qquad \max_{1 \leq i \leq L+1} q_i = \frac{1}{L+1}$$
und somit
$$\max_{1 \leq i \leq L} p_i > \max_{1 \leq i \leq L+1} q_i.$$
Aus der Tatsache, dass $-\log_2(\cdot)$ monoton fällt, folgern wir
$$-\log_2(\max_{1 \leq i \leq L} p_i) < -\log_2(\max_{1 \leq i \leq L+1} q_i),$$
was äquivalent zu $H_\infty(X) < H_\infty(Y)$ ist.

c) Die Verteilung eines Münzwurfes ist gegeben durch $[Pr(\text{Kopf}), Pr(\text{Zahl})] = [q, 1-q]$ wobei $q = 1/2$ für eine faire Münze und $q \in [0,1] \backslash \{1/2\}$ für eine unfaire Münze. Man kann sich leicht überlegen, dass
$$\max\{q, 1-q\} \geq 1/2 \qquad \text{für jedes } q \in [0,1],$$
wobei Gleichheit nur dann eintritt, wenn $q = 1/2$. Da $-\log_2(\cdot)$ streng monoton fällt, ist somit $H_\infty(X) = -\log_2(1/2) = 1$ maximal für einen fairen Münzwurf X und kleiner als 1 für jeden unfairen Münzwurf Y.

d) Für zwei Zufallsvariablen X und Y gilt
$$H_\infty(XY) = -\log_2(\max_{(x,y)} P_{XY}(x,y)).$$
Sind X und Y unabhängig, d.h. gilt $P_{XY}(x,y) = P_X(x)P_Y(y)$ für alle Paare (x,y), gilt
$$\begin{aligned} \max_{(x,y)} P_{XY}(x,y) &= \max_{(x,y)} P_X(x)P_Y(y) \\ &= (\max_x P_X(x))(\max_y P_Y(y)). \end{aligned}$$
Wenden wir $-\log_2(\cdot)$ auf diese Gleichung an und verwenden, dass $\log_2(ab) = \log_2(a) + \log_2(b)$, resultiert die behauptete Gleichung $H_\infty(XY) = H_\infty(X) + H_\infty(Y)$.

Eine weitere vernünftige Forderung, die an ein Unsicherheitsmass gestellt werden kann, hat folgende Form. Sei X die Zufallsvariable, die das Resultat eines (nicht notwendigerweise fairen) Würfelwurfes repräsentiert. Dann soll die verbleibende Unsicherheit über X, wenn man erfährt, ob X gerade oder ungerade ist, gleich dem Mittel aus der Unsicherheit über X wenn X gerade ist und der Unsicherheit über X gegeben dass X ungerade ist, betragen. Etwas formaler soll gelten

$$H(X) - h(p) = pH(X|X \text{ gerade}) + (1-p)H(X|X \text{ ungerade}),$$

wobei $p = Pr[X \in \{2,4,6\}]$. Dies ist für die Entropie H erfüllt, für die Min-Entropie H_∞ jedoch nicht.

3.3 Zufallsvariablen und Entropie

a) Die Elementarereignisse sind $[0,0]$, $[0,1]$, $[1,0]$ und $[1,1]$. Alle Elementarereignisse haben Wahrscheinlichkeit $1/4$. Die Verteilungen der einzelnen Zufallsvariablen sind:

A	0	1	2
P_A	1/2	1/4	1/4

B	[0,0]	[0,1]	[1,0]	[1,1]
P_B	1/4	1/4	1/4	1/4

C	0	1
P_C	1/2	1/2

D	[[0,0],0]	[[0,0],1]	[[0,1],0]	[[0,1],1]	[[1,0],0]	[[1,0],1]	[[1,1],0]	[[1,1],1]
P_D	1/4	0	0	1/4	0	1/4	1/4	0

b)

		X	
	P_{XY}	0	1
Y	0	1/4	1/4
	1	1/4	1/4

			A	
	P_{AC}	0	1	2
C	0	1/4	0	1/4
	1	1/4	1/4	0

c)

D	[[0,0],0]	[[0,0],1]	[[0,1],0]	[[0,1],1]	[[1,0],0]	[[1,0],1]	[[1,1],0]	[[1,1],1]
$P_{D\|C=1}$	0	0	0	1/2	0	1/2	0	0

B	[0,0]	[0,1]	[1,0]	[1,1]
$P_{B\|A\neq 2}$	1/3	1/3	1/3	0

d) A und C sind nicht statistisch unabhängig, denn es gilt zum Beispiel

$$P_A(1)P_C(0) \neq P_{AC}(1,0).$$

C und X sind statistisch unabhängig, denn

$$P_C(c)P_X(x) = P_{C|X}(c,x)P_X(x) = 1/4 = P_{CX}(c,x)$$

für alle c und x.

e)

$$\begin{aligned} H(A) &= H[1/2, 1/4, 1/4] = 1/2 \cdot 1 + 1/4 \cdot 2 + 1/4 \cdot 2 = 1.5 \\ H(C) &= H[1/2, 1/2] = 1/2 \cdot 1 + 1/2 \cdot 1 = 1 \\ H(AC) &= H[1/4, 1/4, 1/4, 1/4] = \log(4) = 2 \end{aligned}$$

f) Da die Verteilungen P_{XY} und P_B identisch sind, ist auch die Entropie gleich: $H(XY) = H(B)$. D ist eine Funktion von B und C, welche wiederum Funktionen von X und Y sind. Also gilt $H(D) \leq H(XY)$. Ein Ausrechnen der Entropien zeigt $H(D) = H(B) = H(XY)$.

3.4 Entropiediagramm