Informations- und Codierungstheorie
7. Übungsblatt für den 20. November 2012

1. (cf. [MacKay, 2003, Exercise 8.6]) Die Zufallsvariablen X und Y haben folgende gemeinsame Verteilung:

$P[X=x,Y=y]$	$x=1$	$x=2$	$x=3$	$x=4$
$y=1$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$
$y=2$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$
$y=3$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
$y=4$	$\frac{1}{4}$	0	0	0

   (a) Berechnen Sie $H(X \otimes Y)$, $H(X)$, $H(Y)$.
   (b) Berechnen Sie für jedes $y \in \{1,2,3,4\}$ die Entropie von $X|_{Y=y}$.
   (c) Berechnen Sie $H(X|Y)$ und $H(Y|X)$.

2. (a) (cf.[MacKay, 2003, Exercise 8.1]) Consider three independent random variables U, V, W with entropies $H(U)$, $H(V)$, $H(W)$. Let $X := U \otimes V$, $Y := V \otimes W$. What is $H(X \otimes Y)$? What is $H(X|Y)$?
   (b) (cf. [MacKay, 2003, Exercise 8.2]) Confirm that it is possible for $H(X|_{Y=b_k})$ to exceed $H(X)$.

3. Wir würfeln einmal. Seien X, Y, Z Zufallsvariablen, die so definiert sind:
   - $X = 0$ genau dann, wenn die Augenzahl gerade ist, $X = 1$ sonst.
   - $Y = 0$ genau dann, wenn die Augenzahl höchstens 3 ist, $Y = 1$ sonst.
   - Z sei die Augenzahl des Wurfes.

   (a) Berechnen Sie $H(Z|X \otimes Y)$ und $H(X \otimes Y|Z)$.
   (b) Wieviel erspart das Wissen über $X \otimes Y$ für die Übertragung von Z? Wie groß ist also $H(Z) - H(Z|X \otimes Y)$?
   (c) Wieviel erspart das Wissen über Z für die Übertragung von $X \otimes Y$? Wie groß ist also $H(X \otimes Y) - H(X \otimes Y|Z)$?
   (d) Berechnen Sie $H(X) - H(X|Y)$!

4. (Teile dieses Beispiels sind [Ash, 1990, Beispiel 1.4 (b)]) Sei (Ω, P) ein endlicher Wahrscheinlichkeitsraum, und seien $X : \Omega \to M$, $Y : \Omega \to N$, $Z : \Omega \to R$ Zufallsvariablen. Zeigen Sie:
   (a) $H(Y \otimes Z|X) = H(Y|X) + H(Z|X \otimes Y)$.
   (b) $H(X) - H(X|Y) = H(Y) - H(Y|X)$.

5. [MacKay, 2003] Wir werfen eine unfaire Münze, die mit Wahrscheinlichkeit p $(0 < p < 1)$ auf Zahl fällt, N mal. Sei $\beta > 0$, und sei z die Anzahl der Würfe auf Zahl. Wir nennen diese Folge von N Würfen *typisch zum Parameter* β, wenn

$$|\frac{1}{N} \cdot \left(-\log_2 \left(p^z \cdot (1-p)^{N-z}\right)\right) - H(p, 1-p)| \leq \beta.$$

   Sei $q(N, \beta)$ die Wahrscheinlichkeit, dass eine Folge von N Münzwürfen mit Zahlwahrscheinlichkeit p typisch zum Parameter β ist. Zeigen Sie:
   (a) $\lim_{N\to\infty} q(N, \beta) = 1$.
   (b) Es gibt höchstens $2^{N\cdot(H(p,1-p)+\beta)}$ typische Folgen der Länge N.

Literatur

[Ash, 1990] Ash, R. B. (1990). *Information theory.* Dover Publications Inc., New York. Corrected reprint of the 1965 original.

[MacKay, 2003] MacKay, D. J. C. (2003). *Information theory, inference and learning algorithms.* Cambridge University Press, New York. The book can be viewed at `http://www.inference.phy.cam.ac.uk/mackay/itprnn/book.html`.