Algorithmen, Wahrscheinlichkeit und Information
Prof. Christian Cachin

FS 2019, Universität Bern
Institut für Informatik

Übung 9

9.1 Entropie (2pt)

Berechnen Sie folgende Grössen für die Verteilung P_{XY} von $X, Y \in \{0,1\}^2$ unten:

a) $H(X)$, $H(Y)$, $H(XY)$;

b) $H(X|Y)$, $H(Y|X)$;

c) $H(X) - H(X|Y)$, $H(Y) - H(Y|X)$.

X \ Y	0	1
0	0.2	0.5
1	0	0.3

9.2 Gegenseitige Information (3pt)

Für zwei Zufallsvariablen X und Y ist die *gegenseitige Information von X und Y* (engl., *mutual information*) definiert als

$$I(X;Y) \;=\; H(X) - H(X|Y).$$

Intuitiv gesprochen misst die gegenseitige Information die Reduktion der Unsicherheit von X welche erfolgt, wenn Y bekannt wird. Die Grösse ist symmetrisch, d.h., es gilt auch

$$I(X;Y) \;=\; H(Y) - H(Y|X) \;=\; I(Y;X).$$

Falls zusätzliche Information Z vorliegt, so ist die *bedingte gegenseitige Information von X und Y gegeben Z* (engl., *conditional mutual information*) definiert als

$$I(X;Y|Z) \;=\; H(X|Z) - H(X|YZ).$$

a) Zeigen Sie, dass

$$I(X;Y) \;=\; H(X) + H(Y) - H(XY);$$
$$I(X;X) \;=\; H(X).$$

b) Finden Sie einen WSK-Raum mit drei Zufallsvariablen X, Y und Z, für welche gilt

$$I(X;Y|Z) > I(X;Y).$$

Das heisst, zusätzliches Wissen (Z) kann die gegenseitige Information erhöhen.

c) Finden Sie einen WSK-Raum mit drei Zufallsvariablen X, Y und Z, für welche gilt

$$I(X;Y|Z) < I(X;Y).$$

Das heisst, zusätzliches Wissen (Z) kann die gegenseitige Information auch verkleinern!

9.3 Kettenregel für Information (2pt)

Die Kettenregel für Entropie besagt, dass $H(XY) = H(X) + H(Y|X)$. Zeigen Sie, dass auch die folgende *Kettenregel für Information* gilt

$$I(X;YZ) \;=\; I(X;Y) + I(X;Z|Y).$$

In Worten: Die Information zwischen X und YZ ist die Summe aus der Information zwischen X und Y plus der Information zwischen X und Z gegeben Y.

9.4 Entropie der geometrischen Verteilung (3pt)

Sei $G \in \mathbb{N}$ eine geometrisch verteilte Zufallsvariable mit Erfolgswahrscheinlichkeit p, d.h., $\mathrm{P}[G = n] = (1-p)^{n-1}\,p$. Zeigen Sie, dass

$$H(G) \;=\; \frac{-p\log p - (1-p)\log(1-p)}{p} \;=\; \frac{h(p)}{p}.$$

Hinweis: Wählen Sie eine der zwei Strategien im Umgang mit der geometrischen Verteilung, wie gezeigt anhand der Herleitung des Mittelwertes [MU17, 2.4]. Berechnen Sie entweder direkt die Entropie von G analog zu Lemma 2.8, wobei folgende Summen hilfreich sein können:

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}, \qquad \sum_{n=0}^{\infty} n r^n = \frac{r}{(1-r)^2}.$$

Alternativ dazu konstruieren Sie ein ähnliches Argument wie [MU17, S. 34], indem Sie eine Indikator-Zufallsvariable Z für das Ergebnis des ersten Versuchs einführen, danach $H(GZ)$ auf verschiedene Weise schreiben und ausnützen, dass G *memoryless* ist.

Referenzen

[MU17] M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis*. Cambridge University Press, 2nd edition, 2017.