TECHNISCHE UNIVERSITÄT DRESDEN

Fakultät Elektrotechnik und Informationstechnik – Institut für Nachrichtentechnik – Professur für Informationstheorie und maschinelles Lernen

Dr. Martin Mittelbach, Dr. Anne Wolf 13.04.2023

1. Hausaufgabe zur Lehrveranstaltung Informationstheorie

Die Aufgaben des 1. Tutoriums (= 1. Hausaufgabe) sollten sinnvollerweise nach der Beschäftigung mit der 1. Übung sowie den Folien der 1. Vorlesung und vor dem 1. Tutorium bearbeitet werden. Mit diesem Material sollten Sie in der Lage sein, die Aufgaben (überwiegend) eigenständig zu lösen.

Aufgabe 6: (Erwartungswert und Varianz transformierter Zufallsgrößen)

Gegeben seien die unabhängigen und identisch verteilten Zufallsgrößen $X_1, X_2, \ldots, X_n$ mit dem Alphabet $\mathcal{X}$ sowie die Funktion $g\colon\ \mathcal{X} \to \mathbb{R}$. Für $k = 1, 2, \ldots, n$ definieren wir die Zufallsgrößen Y_k durch $Y_k := g(X_k)$. Weiterhin sei die Zufallsgröße Z_n gegeben durch

$$Z_n := \frac{1}{n} \sum_{k=1}^{n} Y_k.$$

a) Berechnen Sie den Erwartungswert und die Varianz der Zufallsgröße Z_n und vereinfachen Sie die Ausdrücke so weit wie möglich.

   Hinweis: Die Zufallsgrößen $Y_1, Y_2, \ldots, Y_n$ sind unabhängig und identisch verteilt, da die Zufallsgrößen $X_1, X_2, \ldots, X_n$ unabhängig und identisch verteilt sind und jede Zufallsgröße Y_k durch die Transformation mit der selben Funktion g mit der jeweiligen Zufallsgröße X_k als einzigem Argument entsteht.

b) Berechnen Sie die Grenzwerte $\lim\limits_{n\to\infty} \mathbb{E}\left(Z_n\right)$ und $\lim\limits_{n\to\infty} \operatorname{var}\left(Z_n\right)$.

Aufgabe 7: (Eigenschaften der binären Entropiefunktion)

Gegeben sei eine Bernoulli-verteilte Zufallsgröße X, die Werte aus dem Alphabet $\mathcal{X} = \{0, 1\}$ annehmen kann. Für deren Wahrscheinlichkeitsfunktion p_X gilt: $p_X(1) = q$ und $p_X(0) = 1 - q$ mit $q \in (0, 1)$. Die sogenannte Entropie $H(X)$ dieser Zufallsgröße ist definiert als $H(X) := \mathbb{E}\left(-\log_2(p_X(X))\right)$. Da X nur zwei Werte annehmen kann, ist die Entropie $H(X)$ nur von dem Parameter q abhängig:

$$H(X) = -\sum_{x\in\mathcal{X}} p_X(x) \log_2 p_X(x) = -q \log_2 q - (1-q) \log_2(1-q) =: H_b(q).$$

Sie wird auch als *binäre Entropiefunktion* H_b bezeichnet – siehe Skript (§ 1.4). Einige Eigenschaften dieser binären Entropiefunktion sollen im Folgenden gezeigt werden.

a) Berechnen Sie die Grenzwerte der binären Entropiefunktion für $q \to 0$ und $q \to 1$.

b) Zeigen Sie, dass die Entropie der Zufallsgröße X nichtnegativ ist.

c) Zeigen Sie, dass die binäre Entropiefunktion konkav für $q \in (0, 1)$ ist.

d) Zeigen Sie, dass die maximale Entropie der Zufallsgröße X durch $q = \frac{1}{2}$ erreicht wird, und berechnen Sie diese maximale Entropie.

Aufgabe 8: (Zweistufiges Zufallsexperiment)

Wir betrachten ein zweistufiges Spiel bestehend aus einem Quiz mit anschließendem Münzwurf. Ein Spieler muss zunächst eine Frage beantworten. Beantwortet er die Frage richtig, wird Münze 1 geworfen. Bei falscher Antwort wird Münze 2 geworfen. Der Spieler gewinnt bei dem Ereignis ZAHL.

Beide Münzen sind unfair. Die Wahrscheinlichkeit für ZAHL beträgt $\frac{3}{4}$ bei Münze 1 und $\frac{1}{4}$ bei Münze 2, d. h. durch eine richtige Antwort erhöht der Spieler seine Gewinnchance. Die Wahrscheinlichkeit dafür, dass der Spieler die Frage richtig beantwortet, sei q.

Wir wollen das erste Zufallsexperiment (Frage) durch die Zufallsgröße X und das zweite Zufallsexperiment (Münzwurf) durch die Zufallsgröße Y beschreiben.

a) Bestimmen Sie die gemeinsame Wahrscheinlichkeitsfunktion $p_{X,Y}$ sowie die Randwahrscheinlichkeitsfunktionen p_X und p_Y.

b) Bestimmen sie folgende Erwartungswerte:

$$H(X) := \mathbb{E}\left(-\log_2(p_X(X))\right) \qquad H(Y) := \mathbb{E}\left(-\log_2(p_Y(Y))\right)$$
$$H(X,Y) := \mathbb{E}\left(-\log_2(p_{X,Y}(X,Y))\right).$$

Stellen Sie die Ergebnisse mit Hilfe der binären Entropiefunktion (siehe Aufgabe 7) dar.

c) Begründen Sie, warum

$$H(X,Y) \neq H(X) + H(Y)$$

für alle $0 < q < 1$ gilt.

Aufgabe 9: (Konkavität der Entropie)

Wir betrachten den Vektorraum $(\mathbb{R}^n, +, \cdot)$, d. h. die Menge $\mathbb{R}^n$ der n-dimensionalen reellen Vektoren zusammen mit der Vektoraddition '+' und der Multiplikation '$\cdot$' eines Vektors mit einem Skalar. Dies ist ein Spezialfall des im Skript in (A.8) betrachteten allgemeinen Vektorraums. Wir betrachten nun die Teilmenge

$$\mathcal{M} := \{p = (p_k)_{k=1}^n \in \mathbb{R}^n : p_k > 0 \text{ und } \sum_{k=1}^n p_k = 1\}.$$

a) Zeigen Sie, dass $\mathcal{M} \subseteq \mathbb{R}^n$ eine konvexe Menge ist.

b) Nun betrachten wir die Funktion

$$H: \quad \mathcal{M} \to \mathbb{R}, \quad p = (p_k)_{k=1}^n \mapsto H(p) := -\sum_{k=1}^n p_k \log_2 p_k.$$

Zeigen Sie, dass H eine konkave Funktion ist, d. h. dass für alle Vektoren $p, q \in \mathcal{M}$ und jede reelle Zahl $\lambda \in [0,1]$ gilt:

$$H(\lambda \cdot p + (1-\lambda) \cdot q) \geq \lambda\, H(p) + (1-\lambda)\, H(q).$$

Hinweis: Sie können dafür die folgende Ungleichung verwenden. Es gilt:

$$\sum_{k=1}^n a_k \log_2 a_k \geq \sum_{k=1}^n a_k \log_2 b_k,$$

falls $a_k > 0$, $b_k > 0$ und $\sum_{k=1}^n a_k = \sum_{k=1}^n b_k = 1$.

Eine äquivalente Formulierung dieser Ungleichung finden Sie im Skript in (§ 1.18).