# Zusatzübung zur Theoretischen Informationstechnik I

Prof. Dr. Rudolf Mathar, Simon Görtzen, Christoph Schmitz, Ehsan Zandi

13.3.2014

**Aufgabe 1.** Für eine Krankheit X existiert ein klassisches Impfverfahren, bei dem 3 Injektionen vorgenommen werden, um einen möglichst guten Schutz bezüglich der Erkrankung herzustellen. Eine Reportage über Nebenwirkungen der Impfungen hat die Impfbereitschaft der Bevölkerung in den letzten Jahren beeinflusst, so dass einige Personen das Impfverfahren bereits nach weniger als 3 Injektionen abbrachen.

Zu einem späteren Zeitpunkt kommt es zu einem verbreiteten Ausbruch der Krankheit. Sie sollen nun die Auswirkungen des fehlenden Impfschutzes anhand folgender gegebener Daten analysieren.

Die Wahrscheinlichkeit, nach einer Injektion immun zu sein, beträgt 0.3, nach zwei Injektionen 0.6 und nach drei Injektionen 0.9. Erhält man keine Injektion, so ist eine Immunität unmöglich.

Zum Zeitpunkt des Ausbruchs der Krankheit X haben 10 % der Bevölkerung keine Injektion erhalten, 20 % haben eine Injektion erhalten und 30 % haben zwei Injektionen erhalten. Dagegen haben 40 % haben das klassische Impfverfahren absolviert und alle 3 vorgesehenen Injektionen erhalten.

Insgesamt erkranken 25 % der nicht Immunisierten.

a) Wieviel Prozent der Bürger sind immun?

b) Wieviel Prozent der Bevölkerung erkranken?

c) Wieviel Prozent der immunen Menschen haben alle 3 Injektionen erhalten?

d) Wieviel Prozent der Erkrankten gehören zur Gruppe der gänzlich Ungeimpften?

**Aufgabe 2.** Die Gesamtaufgabe besteht aus drei Teilaufgaben, welche **unabhängig** voneinander lösbar sind.

**Teil I**

Es sei $X$ eine absolut-stetige Zufallsvariable und $a, b \in \mathbb{R}$. Zeigen Sie die Gültigkeit von

**a)** $\mathrm{E}(aX + b) = a\,\mathrm{E}(X) + b$,

**b)** $\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$.

**Teil II**

Es sei $X = X_1 + \ldots + X_n$ die Summe von gemeinsam stochastisch unabhängigen Zufallsvariablen $X_1, \ldots, X_n$. Damit gilt die sogenannte Bienaymé-Gleichung

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i).$$

**a)** Es seien $X_i$, $i = 1, \ldots, n$ Bernoulli-verteilt mit dem Parameter $0 < p < 1$. Geben Sie die resultierende Zähldichte von $X$ an. Berechnen Sie den Erwartungswert und die Varianz von $X$.

**b)** Jetzt gilt $X_i \sim \mathrm{N}(\mu, \sigma^2)$ für $i = 1, \ldots, n$. Berechnen Sie den Erwartungswert und die Varianz von $X$.

**c)** Wie muss $X$ aus dem Unterpunkt **b)** transformiert werden, damit die daraus resultierende Zufallsvariable standardnormalverteilt ist (Zwischenschritte erforderlich)?

**Teil III**

Es seien $X$ bzw. $K$ reelle Zufallsvariablen, mit denen Gesprächsdauer bzw. Kosten eines Telefonanrufs beschrieben werden können. Dabei ist $X \sim \mathrm{Exp}(\lambda)$ mit $\lambda = 2$. Die Kosten eines Anrufes der Länge $y$ betragen

$$k(y) = \begin{cases} 1 & y \leq 4, \\ \frac{1}{4}y & y > 4. \end{cases}$$

Berechnen Sie die mittleren Kosten eines Telefonanrufes.

**Aufgabe 3.** Gegeben sind ein XOR- und ein UND-Gatter mit folgenden Wahrheitstabellen.


XOR-Gatter

| $X$ | $Z$ | $Y$ |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |

UND-Gatter

| $X$ | $Z$ | $Y$ |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 1 | 1 | 1 |

Als Modell der Übertragung dienen die Zufallsvariablen $X, Y \in \{0, 1\}$, wobei $X$ die Kanaleingabe und $Y$ die Kanalausgabe beschreibt. Weiterhin betrachten wir die Zufallsvariable $Z \in \{0, 1\}$ als Störterm, welcher stochastisch unabhänging von $X$ ist. Abhängig von der Verteilung von $Z$ beschreiben beide verrauschte Gatter einen binären, aber nicht notwendigerweise symmetrischen Kanal.

Hinweis: Verwenden Sie für die gesamte Aufgabe $\log_2()$. Außerdem gilt $0 \cdot \log_2 0 = 0$.

**a)** Es gelte $P(Z = 0) = (1 - \varepsilon)$. Wie lautet die Kapazität des verrauschten XOR-Gatters? Geben Sie die kapazitätserreichende Eingangsverteilung für $X$ an.

Für den Rest der Aufgabe wird nur noch das verrauschte UND-Gatter mit $\varepsilon = 1/2$ betrachtet. Zunächst gelte für die Eingangsverteilung $P(X = 1) = p$.

**b)** Berechnen Sie die bedingten Wahrscheinlichkeiten $P(Y = y|X = x)$, für $x, y \in \{0, 1\}$. Skizzieren Sie den dazu korrespondierenden Kanal.

**c)** Bestimmen Sie die Transinformation $I(X; Y) = H(Y) - H(Y|X)$ als Funktion abhängig von $p$.

**d)** Berechnen Sie die kapazitätserreichende Eingangsverteilung.

**Aufgabe 4.** Eine gedächtnislose Quelle produziere die Symbole aus dem Quellalphabet $\mathcal{X} = \{\mathrm{E}, \mathrm{I}, \mathrm{N}, \mathrm{O}, \mathrm{S}\}$ mit den folgenden relativen Häufigkeiten $p_j = \mathrm{P}(X = x_j)$:

| $x_j$ | E | I | N | O | S |
|---|---|---|---|---|---|
| $p_j$ | 39 % | 17 % | 21 % | 7 % | 16 % |

Um die von der Quelle produzierte Symbolfolge über einen binären symmetrischen Kanal (BSC) mit einer Fehlerwahrscheinlichkeit von $\varepsilon = 0.15$ zu übertragen, wird zunächst die folgende binäre Blockkodierung verwendet:

| $x_j$ | E | I | N | O | S |
|---|---|---|---|---|---|
| $\boldsymbol{c}_j = g(x_j)$ | 000 | 011 | 101 | 111 | 110 |

**a)** Welche Ausgabewörter können bei der Übertragung der Eingabe durch den BSC auftreten?

**b)** Die Übertragungswahrscheinlichkeiten des Kanals können als

$$\mathrm{P}(\boldsymbol{Y} = \boldsymbol{b} \mid \boldsymbol{X} = \boldsymbol{c}_j) = \varepsilon^{\Delta(\boldsymbol{c}_j, \boldsymbol{b})}(1-\varepsilon)^{3-\Delta(\boldsymbol{c}_j, \boldsymbol{b})} = \left(\frac{\varepsilon}{1-\varepsilon}\right)^{\Delta(\boldsymbol{c}_j, \boldsymbol{b})} (1-\varepsilon)^3$$

dargestellt werden. Dabei bezeichnet $\Delta(\boldsymbol{c}_j, \boldsymbol{b})$ die Hamming-Distanz zwischen dem Kodewort $\boldsymbol{c}_j$ und dem Ausgabewort $\boldsymbol{b}$, also die Anzahl der Stellen, an denen sich die beiden Wörter unterscheiden. Da die Fehlerwahrscheinlichkeit $\varepsilon$ kleiner als 0.5 ist, ist die Übertragungswahrscheinlichkeit streng monoton fallend in der Hamming-Distanz. Konstruieren Sie mit Hilfe dieses Zusammenhangs eine ML-Dekodierung.

**c)** Bei der Übertragung der Symbolfolge 'NOISE' werde das siebte Bit gestört. Welche Ausgabe liefert die ML-Dekodierung in diesem Fall?

Um die Übertragungskapazität des BSC besser auszunutzen, wird nun im Folgenden statt des Blockkodes ein Huffman-Kode verwendet.

**d)** Konstruieren Sie einen binären Huffman-Kode für die oben angegebene Eingabeverteilung, und geben Sie die mittlere Kodewortlänge $\overline{n}(g)$ an.

**e)** Ist es möglich, die mittlere Kodewortlänge pro Quellbuchstabe im Vergleich zu dem im vorherigen Aufgabenteil ermittelten Wert $\overline{n}(g)$ nochmals um 5 % zu verringern, indem man einen Huffman-Kode für längere Symbolblöcke konstruiert?

**f)** Nach der Übertragung des mit dem Huffman-Kode kodierten Worts 'NOISE' ist das siebte Bit gestört. Welches Wort über dem Alphabet $\{\mathrm{E}, \mathrm{I}, \mathrm{N}, \mathrm{O}, \mathrm{S}\}$ wird dekodiert?